24TH ANNUAL MEETING OF THE

VISION SCIENCES SOCIETY

MAY 17-22, 2024 ST. PETE BEACH, FLORIDA

Abstracts

ELSEVIER

National Eye Institute
NATIONAL INSTITUTES OF HEALTH

VPixx Vision Science Solutions
VPixx Technologies Inc.
www.vpixx.com

VISION SCIENCES SOCIETY

VPixx
Vision Science Solutions

VPixx is proud to be a gold-level sponsor of VSS 2024

Here's where you can find us at the conference

Satellite Session: A Multispectral Projector for Advanced Vision Science
A tour of the engineering and design of the PROPixx Tetracolour and its applications, featuring guest Dr. Pablo Barrionuevo (UNT-CONICET, Justus-Liebig Universitat Giessen)
May 18, 12:45 pm - 2:15 pm in Banyan/Citrus

Poster: A Recipe for a 4+ Primaries Projector
Design considerations of multispectral displays, featuring the PROPixx Tetracolour
May 19, 8:30 am - 12:30 pm in Banyan Breezeway

Demo Night: Magic Metamers and Saccadic Suppression, Hidden in Plain Sight
See our PROPixx Tetracolour and 1440 Hz mode in action. Find the hidden message on the display and win a free drink ticket!
May 20, 7:00 pm - 10:00 pm in Talk Room 1- 2

Exhibition Hall
May 18, 19, 21, 8:00 am - 5:30 pm in Pavillion
May 20, 8:00 am - 12:30 pm in Pavillion
Come and see our demos, grab some swag, and check out our promotion to win a free colorimeter!

See the full schedule:

vpixx.com/whats-new/

OPAM 32

November 21, 2024

NEW YORK CITY NY

Join the conference organised for and by graduate and postdoctoral researchers where we highlight outstanding early career research on the topics of Object Perception, Attention, and Memory (OPAM)

Visit www.opam.net for more details

***OPAM Social*: Monday, May 20th**

Salty's Tiki Bar starting at 12:30pm

Come chat with us! Swing by the Tiki Bar to meet past and present organizers at VSS. We're happy to answer any questions or just have you hang out with us!

Organized by: Han Zhang, Karolina Krzys, Ryan O'Donnell, Dock Duncan, Noah Britt, William Narhi-Martinez

Schedule Overview

Wednesday, May 15, 2024			
9:00 am - 6:00 pm	Computational and Mathematical Models in Vision (MODVIS)	Satellite	Blue Heron
Thursday, May 16, 2024			
9:00 am - 6:00 pm	Computational and Mathematical Models in Vision (MODVIS)	Satellite	Blue Heron
4:00 - 7:00 pm	Registration Open	Registration	Grand Palm Colonnade
Friday, May 17, 2024			
7:00 am - 6:00 pm	Registration Open	Registration	Grand Palm Colonnade
7:30 am - 9:30 pm	VSS Quiet Lounge	Lounge	Compass
7:30 am - 9:30 pm	VSS Social Lounge	Lounge	Royal Tern
8:00 - 9:00 am	Coffee Break	Break	Garden Courtyard
9:00 am - 12:00 pm	Computational and Mathematical Models in Vision (MODVIS)	Satellite	Blue Heron
9:00 am - 12:00 pm	The Vision Science of Digital Readability: Community-Building Workshop	Satellite	Snowy Egret
11:30 am - 12:00 pm	Coffee Break	Break	Garden Courtyard
12:00 - 2:00 pm	Neurodiversity in visual functioning: Moving beyond case-control studies	Symposium	Talk Room 1
12:00 - 2:00 pm	Large-scale visual neural datasets: where do we go from here?	Symposium	Talk Room 2
12:00 - 2:30 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard	Break	Garden Courtyard
2:00 - 2:30 pm	Coffee Break	Break	Garden Courtyard
2:30 - 4:30 pm	The temporal evolution of visual perception	Symposium	Talk Room 1
2:30 - 4:30 pm	Attention: accept, reject, or major revisions?	Symposium	Talk Room 2
4:30 - 5:00 pm	Coffee Break	Break	Garden Courtyard
5:00 - 7:00 pm	The Multifaceted effects of blindness and how sight might be restored	Symposium	Talk Room 1
5:00 - 7:00 pm	Using deep networks to re-imagine object-based attention and perception	Symposium	Talk Room 2
7:00 - 9:30 pm	Opening Night Reception	Social	Beach
8:30 - 10:00 pm	Visibility: A Gathering of LGBTQ+ Vision Scientists and Friends	Satellite	Garden Courtyard
Saturday, May 18, 2024			
7:30 am - 6:45 pm	Registration Open	Registration	Grand Palm Colonnade
7:30 am - 9:30 pm	VSS Quiet Lounge	Lounge	Compass
7:30 am - 9:30 pm	VSS Social Lounge	Lounge	Royal Tern
7:45 - 8:45 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard & Pavilion
8:00 am - 5:30 pm	Exhibits Open	Exhibits	Pavilion
8:15 - 9:45 am	Perceptual Organization	Talk Session	Talk Room 1
8:15 - 9:45 am	Object Recognition: Models	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Saturday Morning Posters Scene Perception: Miscellaneous Spatial Vision: Neural mechanisms Development: Clinical and high-level Plasticity and Learning: Models, neural mechanisms Face and Body Perception: Bodies Face and Body Perception: Neural mechanisms of social cognition Multisensory Processing: Neural coding Multisensory Processing: Development, clinical Motion: Higher-order	Poster Session	Banyan Breezeway

Saturday, May 18, 2024 …continued			
8:30 am - 12:30 pm	Saturday Morning Posters Visual Search: Attention, phenomena 1 Attention: Capture Attention: Neural mechanisms Attention: Inattention, attentional blindness, suppression Decision Making: Decision making and actions Decision Making: Perceptual decision making 1	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break	Break	Garden Courtyard & Pavilion
10:45 am - 12:30 pm	Eye Movements: Neural mechanisms	Talk Session	Talk Room 1
10:45 am - 12:30 pm	Color, Light and Materials: Neural mechanisms, models	Talk Session	Talk Room 2
12:30 - 2:30 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard	Break	Garden Courtyard
12:45 - 2:15 pm	Unveiling the Potential of AI in Understanding Human Vision with Ethical Integration Organized by the VSS Student-Postdoc Advisory Committee (SPC)	Workshop	Snowy Egret
12:45 - 2:15 pm	A Multispectral Projector for Advanced Vision Science Organized by VPixx Technologies	Satellite	Banyan/Citrus
12:45 - 2:15 pm	Enhancing Accessibility Workshop Organized by the VSS Student-Postdoc Advisory Committee (SPC) and VSS Diversity & Inclusion Committee	Workshop	Sabal/Sawgrass
1:00 - 2:00 pm	VSS Public Lecture given by Arthur Shapiro	Other	Offsite
2:30 - 4:15 pm	Attention: Selection, modulation, resource competition	Talk Session	Talk Room 1
2:30 - 4:15 pm	Face and Body Perception: Development, disorders, models, neural mechanisms	Talk Session	Talk Room 2
2:45 - 6:45 pm	Saturday Afternoon Posters Motion: Optic flow Temporal Processing: Neural mechanisms, models Eye Movements: Learning, expertise, context and faces Eye Movements: Saccades Visual Memory: Working memory and neural mechanisms Undergraduate Just-In-Time 1	Poster Session	Banyan Breezeway
2:45 - 6:45 pm	Saturday Afternoon Posters Multisensory Processing: Audiovisual behavior Multisensory Processing: Illusions, recognition Object Recognition: High-level features Object Recognition: Visual preference Color, Light and Materials: Surfaces, materials Color, Light and Materials: Art, cognition 3D Perception: Depth cue integration, neural mechanisms Scene Perception: Neural mechanisms	Poster Session	Pavilion
4:15 - 5:00 pm	Afternoon Coffee and Snack	Break	Garden Courtyard & Pavilion
5:15 - 6:45 pm	Perception of Relations, Intuitive Physics	Talk Session	Talk Room 1
5:15 - 6:45 pm	Visual Search 1	Talk Session	Talk Room 2
7:15 - 8:15 pm	Keynote Lecture given by Dora Biro Eye in the sky: visually-guided navigation in birds	Keynote	Talk Room 1-2
Sunday, May 19, 2024			
7:30 am - 6:45 pm	Registration Open	Registration	Grand Palm Colonnade
7:30 am - 9:30 pm	VSS Quiet Lounge	Lounge	Compass
7:30 am - 9:30 pm	VSS Social Lounge	Lounge	Royal Tern

Sunday, May 19, 2024… continued			
7:45 - 8:45 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard & Pavilion
8:00 am - 5:30 pm	Exhibits Open	Exhibits	Pavilion
8:00 am - 6:00 pm	Art of Memory Exhibition	Satellite	Glades
8:15 - 9:45 am	Motion	Talk Session	Talk Room 1
8:15 - 9:45 am	Scene Perception: Behaviour, psychophysics	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Sunday Morning Posters Plasticity and Learning: Electrophysiology, brain stimulation Decision Making: Perceptual decision making 2 Binocular Vision: Eye dominance and rivalry Color, Light and Materials: Neural mechanisms, models, disorders Development: Natural experience and eye movements Pre-Data-Collection Posters	Poster Session	Banyan Breezeway
8:30 am - 12:30 pm	Sunday Morning Posters Object Recognition: Neural mechanisms Attention: Features, objects 1 Attention: Features, objects 2 Attention: Spatial selection 1 Attention: Spatial selection 2 Eye Movements: Fixational eye movements	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break	Break	Garden Courtyard & Pavilion
10:45 am - 12:30 pm	Visual Memory: Working and behavior	Talk Session	Talk Room 1
10:45 am - 12:30 pm	Spatial Vision	Talk Session	Talk Room 2
12:30 - 2:30 pm	Canadian Vision Science Social Sponsored by the York Centre for Vision Research and VISTA	Satellite	Sabal/Sawgrass
12:30 - 2:30 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard	Break	Garden Courtyard
12:45 - 2:15 pm	Virtual Reality + Eye Tracking for Research Organized by WorldViz VR	Satellite	Blue Heron
1:00 - 2:00 pm	Career Transitions Workshop Organized by the VSS Student-Postdoc Advisory Committee (SPC)	Workshop	Snowy Egret
1:00 - 2:00 pm	US Funding Workshop	Workshop	Banyan/Citrus
2:30 - 4:15 pm	Attention: Tracking, shifting, capture	Talk Session	Talk Room 1
2:30 - 4:15 pm	Color, Light and Materials: Materials, integrated perception	Talk Session	Talk Room 2
2:45 - 6:45 pm	Sunday Afternoon Posters Object Recognition: Reading Object Recognition: Acquisition of categories Scene Perception: Virtual environments, intuitive physics Spatial Vision: Models Visual Search: Cueing, context, scene complexity, semantics Visual Search: Eye movements, suppression	Poster Session	Banyan Breezeway
2:45 - 6:45 pm	Sunday Afternoon Posters Face and Body Perception: Neural mechanisms 1 Face and Body Perception: Neural mechanisms 2 Face and Body Perception: Development, experience Visual Memory: Working memory and neural mechanisms, models, decision making Visual Memory: Imagery Visual Memory: Working memory and objects, features Action: Reach, grasp, track	Poster Session	Pavilion

Sunday, May 19, 2024... continued			
4:15 - 5:00 pm	Afternoon Coffee and Snack	Break	Garden Courtyard & Pavilion
5:15 - 7:15 pm	Plasticity and Learning	Talk Session	Talk Room 1
5:15 - 7:15 pm	Eye Movements: Early visual processing	Talk Session	Talk Room 2
7:15 - 9:15 pm	VISxVISION Workshop: Vision Science and Data Visualization Research	Satellite	Banyan/Citrus
Monday, May 20, 2024			
7:30 am - 12:30 pm	VSS Quiet Lounge	Lounge	Compass
7:30 am - 12:30 pm	VSS Social Lounge	Lounge	Royal Tern
7:45 - 8:45 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard & Pavilion
7:45 am - 1:30 pm	Registration Open	Registration	Grand Palm Colonnade
8:00 am - 12:30 pm	Exhibits Open	Exhibits	Pavilion
8:00 am - 12:30 pm	Art of Memory Exhibition	Satellite	Glades
8:15 - 9:45 am	Eye Movements: Perception and timing	Talk Session	Talk Room 1
8:15 - 9:45 am	Object Recognition: Neural mechanisms	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Monday Morning Posters Motion: Illusions Visual Memory: Encoding, retrieval Visual Memory: Working memory and development, individual differences, capacity, resolution Visual Memory: Capacity, long-term memory Color, Light and Materials: Appearance, categories Undergraduate Just-In-Time 2	Poster Session	Banyan Breezeway
8:30 am - 12:30 pm	Monday Morning Posters Data Visualization Visual Search: Neural mechanisms, clinical, applied Spatial Vision: Crowding, eccentricity Scene Perception: Categorization Attention: Tracking, shifting Attention: Temporal selection Attention: Divided, resource competition Plasticity and Learning: Disorders, atypical vision	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break	Break	Garden Courtyard & Pavilion
10:45 am - 12:15 pm	3D Perception	Talk Session	Talk Room 1
10:45 am - 12:15 pm	Development	Talk Session	Talk Room 2
12:30 - 2:00 pm	VSS Awards Session Davida Teller Award, Ken Nakayama Medal, Young Investigator Talks	Award	Talk Room 2
12:30 - 2:30 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard	Break	Garden Courtyard
2:00 - 3:30 pm	Psychophysics Software with MATLAB Organized by MathWorks and INCF	Satellite	Jasmine/Palm
2:00 - 5:00 pm	So you published your work – now what? Paths for vision scientists to drive societal impact	Satellite	Blue Heron
2:00 - 5:00 pm	How to stop worrying and love computational neuroimaging of the visual cortex	Satellite	Banyan/Citrus
2:15 - 3:15 pm	Undergraduate Meet & Greet	Student	Pirate Island
3:30 - 5:00 pm	Meet the Professors	Student	Banyan Breezeway
6:00 - 8:00 pm	Demo Night Beach BBQ	Social	Beach
7:00 - 10:00 pm	Demo Night	Social	Talk Room 1-2

Tuesday, May 21, 2024			
7:30 am - 9:30 pm	VSS Quiet Lounge	Lounge	Compass
7:30 am - 9:30 pm	VSS Social Lounge	Lounge	Royal Tern
7:45 - 8:45 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard & Pavilion
7:45 am - 6:45 pm	Registration Open	Registration	Grand Palm Colonnade
8:00 am - 5:30 pm	Exhibits Open	Exhibits	Pavilion
8:00 am - 6:00 pm	Art of Memory Exhibition	Satellite	Glades
8:15 - 9:45 am	Object Recognition: Categories and features	Talk Session	Talk Room 1
8:15 - 9:45 am	Attention: Neural mechanisms	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Tuesday Morning Posters Color, Light and Materials: Lightness, brightness Scene Perception: Ensembles, natural image statistics Visual Memory: Working memory and attention Visual Memory: Working memory and encoding, retrieval Visual Memory: Working memory and behavior, models Plasticity and Learning: Properties Binocular Vision: Clinical Action: Representation	Poster Session	Banyan Breezeway
8:30 am - 12:30 pm	Tuesday Morning Posters Eye Movements: Natural world and VR Motion: Detection Visual Search: Attention, phenomena 2 Decision Making: Perceptual decision making 3 Temporal Processing: Duration, atypical, timing perception 3D Perception: Size, shape, distance 3D Perception: Virtual and augmented reality	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break	Break	Garden Courtyard & Pavilion
10:45 am - 12:15 pm	Face and Body Perception	Talk Session	Talk Room 1
10:45 am - 12:15 pm	Visual Search 2	Talk Session	Talk Room 2
12:15 - 1:00 pm	VSS Business Meeting	Business	Talk Room 2
12:30 - 2:30 pm	phiVis: Philosophy of Vision Science Workshop	Satellite	Banyan/Citrus
12:30 - 2:30 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard	Break	Garden Courtyard
1:00 - 2:30 pm	Connect with Industry	Networking	Spotted Curlew
1:00 - 2:30 pm	Connect with Industry	Networking	Blue Heron
1:00 - 2:30 pm	Connect with Industry	Networking	Snowy Egret
1:00 - 2:30 pm	Negotiation: When To Do It and How To Do It Successfully Organized by Females of Vision et al (FoVea)	Satellite	Jasmine/Palm
2:30 - 4:15 pm	Visual Memory: Working and neural mechanisms	Talk Session	Talk Room 1
2:30 - 4:15 pm	Temporal Processing	Talk Session	Talk Room 2
2:45 - 6:45 pm	Tuesday Afternoon Posters Face and Body Perception: Emotion Face and Body Perception: Wholes, parts configurations, features Visual Search: Memory, search templates Visual Search: Mechanisms, models Perceptual Organization: Neural mechanisms, models Perceptual Organization: Parts, wholes, groups Pre-Data-Collection Posters	Poster Session	Banyan Breezeway

Tuesday, May 21, 2024 …continued			
2:45 - 6:45 pm	Tuesday Afternoon Posters Action: Locomotor, flow, steering Action: Clinical, neural Eye Movements: Clinical Object Recognition: Basic features Binocular Vision: Disparity, stereopsis and suppression Object Recognition: Structure of categories Spatial Vision: Machine learning, neural networks	Poster Session	Pavilion
4:15 - 5:00 pm	Afternoon Coffee and Snack	Break	Garden Courtyard & Pavilion
5:15 - 7:15 pm	Multisensory Processing	Talk Session	Talk Room 1
5:15 - 7:15 pm	Decision Making	Talk Session	Talk Room 2
10:00 pm - 2:00 am	Club Vision	Social	Talk Room 1-2
Wednesday, May 22, 2024			
7:30 am - 12:30 pm	VSS Quiet Lounge	Lounge	Compass
7:30 am - 12:30 pm	VSS Social Lounge	Lounge	Royal Tern
7:45 - 8:45 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard & Pavilion
7:45 am - 12:45 pm	Registration Open	Registration	Grand Palm Colonnade
8:00 am - 12:30 pm	Art of Memory Exhibition	Satellite	Glades
8:15 - 10:00 am	Action	Talk Session	Talk Room 1
8:15 - 10:00 am	Visual Memory	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Wednesday Morning Posters Perceptual Organization: Segmentation, shapes, objects Spatial Vision: Image statistics, neural mechanisms Motion: Neural mechanisms Eye Movements: Perception, cognition and memory Eye Movements: Accuracy, pursuit and eccentricity	Poster Session	Banyan Breezeway
8:30 am - 12:30 pm	Wednesday Morning Posters Object Recognition: Models Face and Body Perception: Models Face and Body Perception: Disorders, individual differences Face and Body Perception: Social cognition Attention: Reward, motivation, emotion Attention: Exogenous, endogenous, gaze	Poster Session	Pavilion
10:00 - 10:45 am	Coffee Break	Break	Garden Courtyard & Pavilion
11:00 am - 12:45 pm	Scene Perception: Neural mechanisms, representations	Talk Session	Talk Room 1
11:00 am - 12:45 pm	Binocular Vision	Talk Session	Talk Room 2

Symposium Sessions

Symposium: Friday, May 17, 2024, 12:00 – 2:00 pm, Talk Room 1

Neurodiversity in visual functioning: Moving beyond case-control studies

Organizers: Catherine Manning[1], Michael-Paul Schallmo[2];
[1]University of Reading, UK, [2]University of Minnesota
Presenters: Catherine Manning, Michael-Paul Schallmo, Victor Pokorny, Brian Keane, Beier Yao, Alice Price

Although vision science has a rich history of investigating atypical functioning in developmental and psychiatric conditions, these studies have tended to compare a single diagnosis against a normative comparison group (the case-control approach). However, by studying diagnoses in isolation, we cannot determine whether case-control differences are condition-specific, or instead reflect neural changes that occur across multiple conditions. A related challenge to the case-control approach is the growing recognition that categorical diagnoses are not biologically or psychologically discrete entities: multiple diagnoses commonly co-occur within individuals, considerable heterogeneity is found among individuals with the same diagnosis, and similarities are often found between diagnosed individuals and those with subclinical traits. Moreover, categorical diagnoses do not clearly map onto the underlying biology (e.g., genes, neural function). Accordingly, there has been a recent conceptual shift away from the traditional case-control approach towards considering continuous, transdiagnostic dimensions of neurodiversity, which might better reflect the underlying biology (c.f. NIH's Research Domain Criteria framework). By studying dimensions of visual functioning across conditions, we will elucidate the mechanisms implicated in cases of atypical visual functioning, while also helping to understand individual differences in the non-clinical population. This symposium will bring together cutting-edge research that goes beyond the traditional case-control approach to demonstrate this recent conceptual shift. Speakers representing diverse career-stages, scientific approaches and nationalities will present research encompassing a range of conditions (e.g., autism, dyslexia, schizophrenia, bipolar disorder, migraine) and methods (EEG, fMRI, psychophysics, computational modelling, questionnaires). Cathy Manning will first introduce the traditional case-control approach and its limitations, before presenting EEG and behavioural work identifying both convergence and divergence in autistic and dyslexic children's visual motion processing and decision-making. Second, Michael-Paul Schallmo will show that weaker surround suppression is shared by both adults with autism and schizophrenia, and linked to continuous dimensions of psychiatric symptoms. Third, Victor Pokorny will describe a recent meta-analysis that found surprisingly weak evidence for generally weakened use of visuospatial context in schizophrenia, bipolar disorder, and related sub-clinical populations, but stronger evidence for specific alterations in contrast perception. Fourth, Brian Keane will describe how functional connectivity involving a higher-order visual network is aberrant in psychosis patients, regardless of diagnosis. Fifth, Beier Yao will present a visuomotor mechanism that is altered across psychosis diagnoses and relates to positive symptoms. Finally, Alice Price will describe how factors of the visual Cardiff Hypersensitivity Scale differ across conditions and in the general population. We will finish with a panel discussion drawing out overall themes and covering theoretical and practical considerations for advancing investigations into neurodiversity in visual functioning. The symposium will inform a richer understanding within the VSS community of visual function in psychiatric and neurodevelopmental conditions, and individual differences more broadly. The presentations and discussion will benefit both junior and senior vision scientists by highlighting cutting-edge methods and emerging theories of neurodiversity. The symposium is timely not only because of the recent "transdiagnostic revolution" (Astle et al., 2022), but also due to the increasing prevalence of diagnoses (e.g., autism, mental health difficulties).

Talk 1

VISUAL PROCESSING AND DECISION-MAKING IN CHILDREN WITH AUTISM AND DYSLEXIA: INSIGHTS FROM CROSS-SYNDROME APPROACHES

Catherine Manning[1,2]; [1]University of Reading, UK, [2]University of Birmingham, UK

Atypical visual processing has been reported in a range of developmental conditions, including autism and dyslexia. One explanation for this is that certain neural processes are vulnerable to atypical development, leading to shared effects across developmental conditions. However, few studies make direct comparisons between developmental conditions, or use sensitive-enough methods, to conclude whether visual processing is affected differently in these conditions, or whether they are affected similarly, therefore reflecting a more general marker of atypical development. After evaluating the current state of the science, I will present findings from two sets of studies that apply computational modelling approaches (equivalent noise modelling and diffusion modelling) and measure EEG data in matched groups of autistic, dyslexic and typically developing children aged 6 to 14 years (n = ~50 per group). These methods help pinpoint the component processes involved in processing visual information and making decisions about it, while linking brain and behaviour. The results identify both areas of convergence and divergence in autistic and dyslexic children's visual processing and decision-making. For example, both autistic and dyslexic children show differences in late stimulus-locked EEG activity in response to coherent motion stimuli, which may reflect reduced segregation of signal-from-noise. However only dyslexic children (and not autistic children) show a reduced accumulation of sensory evidence which is reflected in a shallower build-up of activity in a centro-parietal EEG component. Therefore, while there may be some shared effects across conditions, there are also condition-specific effects, which will require refined theories.

TALK 2

WEAKER VISUAL SURROUND SUPPRESSION IN BOTH AUTISM SPECTRUM AND PSYCHOSIS SPECTRUM DISORDERS

Michael-Paul Schallmo[1]; [1]University of Minnesota

Issues with sensory functioning and attention are common in both autism spectrum and psychosis spectrum disorders. Despite important differences in symptoms and developmental time course, these conditions share a number of common features with regard to visual perception. One such phenomenon that we and others have observed in both populations is a reduced effect of surrounding spatial context during the perception of basic visual features such as contrast or motion. In this talk, we will consider whether these differences in visual function may have a common source. In a series of psychophysical, and brain imaging experiments, we found that young adults with ASD showed weaker visual surround suppression during motion perception, as compared to neurotypical individuals. This was reflected by differences in behavioral task performance and fMRI responses from area MT. Likewise, across multiple experiments in people with psychosis, we have found that individuals with schizophrenia show weaker behavioral and neural surround suppression during visual contrast perception. Recently, we used a divisive normalization model to show that narrower spatial attention may be sufficient to explain weaker surround suppression in ASD. This theory was subsequently given support by another group who showed weaker suppression for narrow vs. broad attention conditions in healthy adults. Previous studies have also found narrower spatial attention both in people with ASD and in schizophrenia. Thus, we suggest narrower attention may be a common sensory difference that is sufficient to account for weaker surround suppression across both ASD and schizophrenia, versus neurotypicals.

TALK 3

ATYPICAL USE OF VISUOSPATIAL CONTEXT IN SCHIZOPHRENIA, BIPOLAR DISORDER, AND SUBCLINICAL POPULATIONS: A META-ANALYSIS

Victor Pokorny[1], Sam Klein[1], Collin Teich[2], Scott Sponheim[1,2], Cheryl Olman[1], Sylia Wilson[1]; [1]University of Minnesota, [2]Minneapolis Veterans Affairs Health Care System

Visual perception in people with psychotic disorders is thought to be minimally influenced by surrounding visual elements (i.e. visuospatial context). Visuospatial context paradigms have unique potential to clarify the neural bases of psychotic disorders because a) the neural mechanisms are well-studied in both animal and human models and b) generalized cognitive deficits are unlikely to explain altered performance. However, the published literature on the subject is conflicting and heterogeneous such that a systematic consolidation and evaluation of the published evidence is needed. We conducted a systematic review and meta-analysis of 46 articles spanning over fifty years of research. Articles included behavioral, fMRI and EEG reports in schizophrenia, bipolar disorder, and subclinical populations. When pooling across all paradigm types, we found little evidence of reduced use of visuospatial context in schizophrenia (Hedges' g=0.20), and marginal evidence for bipolar disorder (g=0.25). The strongest evidence was observed for altered contrast perception paradigms in schizophrenia (g=0.73). With respect to subclinical populations, we observed immense heterogeneity in populations of interest, individual-difference measures, and study designs. Our meta-analysis provided surprisingly weak evidence for the prevailing view that psychotic disorders are associated with a general reduction in use of visuospatial context. Instead, we observed strongest evidence for a specific alteration in the effect of visuospatial context during contrast perception. We propose altered feedback to primary visual cortex as a potential neural mechanism of this effect.

TALK 4

A NOVEL SOMATO-VISUAL FUNCTIONAL CONNECTIVITY BIOMARKER FOR AFFECTIVE AND NON-AFFECTIVE PSYCHOSIS

Brian Keane[1], Yonatan Abrham[1], Michael Cole[2], Brent Johnson[1], Carrisa Cocuzza[3]; [1]University of Rochester, [2]The State University of New Jersey, [3]Yale University

People with psychosis are known to exhibit thalamo-cortical hyperconnectivity and cortico-cortical hypoconnectivity with sensory networks, however, it remains unclear if this applies to all sensory networks, whether it impacts affective and non-affective psychosis equally, or whether such differences could form the basis of a viable biomarker. To address the foregoing, we harnessed data from the Human Connectome Early Psychosis Project and computed resting-state functional connectivity (RSFC) matrices for healthy controls and affective/non-affective psychosis patients who were within 5 years of illness onset. Primary visual, secondary visual ("visual2"), auditory, and somatomotor networks were defined via a recent brain network partition. RSFC was determined for 718 regions (358 subcortical) via multiple regression. Both patient groups exhibited cortico-cortical hypoconnectivity and thalamo-cortical hyperconnectivity in somatomotor and visual2 networks. The patient groups were similar on every RSFC comparison. Across patients, a robust psychosis biomarker emerged when thalamo-cortical and cortico-cortical connectivity values were averaged across the somatomotor and visual2 networks, normalized, and subtracted. Four thalamic regions linked to the same two networks disproportionately drove the group difference (p=7e-10, Hedges' g=1.10). This "somato-visual" biomarker was present in antipsychotic-naive patients and discoverable in a 5 minute scan; it could differentiate psychosis patients from healthy or ADHD controls in two independent data sets. The biomarker did not depend on comorbidities, had moderate test-retest reliability (ICC=.59), and could predict patient status in a held-out sample (sensitivity=.66, specificity=.82, AUC=.83). These results show that-across psychotic disorder diagnoses- an RSFC biomarker can differentiate patients from controls by the early illness stages.

TALK 5

ABNORMAL OCULOMOTOR COROLLARY DISCHARGE SIGNALING AS A TRANS-DIAGNOSTIC MECHANISM OF PSYCHOSIS

Beier Yao[1,2,3], Martin Rolfs[4], Rachael Slate[5], Dominic Roberts[3], Jessica Fattal[6], Eric Achtyes[7,8], Ivy Tso[9], Vaibhav Diwadkar[10], Deborah Kashy[3], Jacqueline Bao[3], Katharine Thakkar[3]; [1]McLean Hospital, [2]Harvard Medical School, [3]Michigan State University, [4]Humboldt University, [5]Brigham Young University, [6]Northwestern University, [7]Cherry Health, [8]Western Michigan University Homer Stryker M.D. School of Medicine, [9]The Ohio State University, [10]Wayne State University

Corollary discharge signals (CD) are "copies" of motor signals sent to sensory areas to predict the corresponding input. Because they are used to distinguish actions generated by oneself versus external forces, altered CDs are a hypothesized mechanism for agency disturbances in psychosis (e.g., delusion of alien control). We focused on the visuomotor system because the CD relaying circuit has been identified in primates, and the CD influence on visual perception can be quantified using psychophysical paradigms. Previous studies have shown a decreased influence of CD on visual perception in (especially more symptomatic) individuals with schizophrenia. We therefore hypothesized that altered CDs may be a trans-diagnostic mechanism of psychosis. We examined oculomotor CDs (using the trans-saccadic localization task) in 49 participants with schizophrenia or schizoaffective disorder (SZ), 36 psychotic bipolar participants (BPP), and 40 healthy controls (HC). Participants made a saccade to a visual target. Upon saccade initiation, the target disappeared and reappeared at a horizontally displaced position. Participants indicated the direction of displacement. With intact CDs, participants can remap the pre-saccadic target and make accurate perceptual judgements. Otherwise, participants may use saccade landing site as a proxy of pre-saccadic target. We found that both SZ and BPP were less sensitive to target displacement than HC. Regardless of diagnosis, patients with more severe positive symptoms were more likely to rely on saccade landing site. These results suggest a reduced influence of CDs on visual perception in SZ and BPP and, thus, that altered CD may be a trans-diagnostic mechanism of psychosis.

TALK 6

THE FOUR FACTORS OF VISUAL HYPERSENSITIVITY: DEFINITION AND MEASUREMENT ACROSS 16 CLINICAL DIAGNOSES AND AREAS OF NEURODIVERSITY

Alice Price[1], Petroc Sumner[1], Georgie Powell[1]; [1]Cardiff University

Subjective sensitivity to visual stimuli, including repeating patterns and bright lights, is known to associate with several clinical conditions (e.g., migraine, anxiety, autism), and also occurs in the general population. Anecdotal reports suggest that people might be sensitive to different types of visual stimuli (e.g., to motion vs lights). The visual Cardiff Hypersensitivity Scale-Visual (CHYPS-V) was developed to define and measure the different factors of visual hypersensitivity, using questions which focus upon functional impact rather than affective changes. Across five samples (n > 3000), we found four highly replicable factors using bifactor modelling. These were brightness (e.g., sunlight), repeating patterns (e.g., stripes), strobing (e.g., light flashes), and intense visual environments (e.g., supermarkets). The CHYPS-V and its subscales show very good reliability (α > .80, ω > .80) and improved correlations with measures of visual discomfort. We also used the CHYPS-V to delineate how these factors may differentiate clinical diagnoses and areas of neurodiversity from each other, and from the general population. Differences from individuals reporting no clinical diagnoses were most pronounced for the intense visual environments subscale, with individuals reporting a diagnosis of autism, fibromyalgia, or persistent postural perceptual dizziness (PPPD) scoring highest. Whilst many conditions showed a similar pattern of visual sensitivity across factors, some conditions (e.g., migraine, PPPD) show evidence of condition specific sensitivities (e.g., to pattern, or to strobing). Further to identifying the factor structure of visual hypersensitivity, CHYPS-V can be used to help investigate underlying mechanisms which give rise to these differences in visual experience.

Symposium: Friday, May 17, 2024, 12:00 – 2:00 pm, Talk Room 2

Large-scale visual neural datasets: where do we go from here?

Organizers: Alessandro Gifford[1], Kendrick Kay[2]; [1]Freie Universität Berlin, [2]University of Minnesota
Presenters: Eline R. Kupers, Won Mok Shim, Ian Charest, Tomas Knapen, Jacob Prince, Alessandro T. Gifford

Vision science has witnessed an increase in worldwide initiatives collecting and publicly releasing large-scale visual neural datasets (LSVNDs). These initiatives have allowed thousands of vision scientists to readily harness LSVNDs, enabling new investigations and resulting in novel discoveries. This suggests vision science is entering a new era of inquiry characterized by big open data. The rapid growth in the collection and use of LSVNDs spawns urgent questions, the answers to which will steer the direction of the field. How can different researchers across the vision sciences spectrum benefit from these datasets? What are the opportunities and pitfalls of LSVNDs for theory formation? Which kinds of LSVNDs are missing, and what characteristics should future LSVNDs have to maximize their impact and utility? How can LSVNDs support a virtuous cycle between neuroscience and artificial intelligence? This symposium invites the VSS community to engage these questions in an interactive and guided community process. We will start with a short introduction (5 minutes), followed by six brief, thought-provoking talks (each 9 minutes plus 3 minutes for Q&A). Enriched by these perspectives, the symposium will then move to a highly interactive 30-minute discussion where we will engage the audience to discuss the most salient open questions on LSVNDs, generate and share insights, and foster new collaborations. Speakers from diverse career stages will cover a broad range of perspectives on LSVNDs, including dataset creators (Kupers, Shim), dataset users (Prince), and researchers playing both roles (Gifford, Charest, Knapen). Eline Kupers will expose behind-the-scenes knowledge on a particular LSVND that has received substantial traction in the field, the Natural Scenes Dataset (NSD), and will introduce ongoing efforts for a new large-scale multi-task fMRI dataset called Visual Cognition Dataset. Won Mok Shim will introduce the Naturalistic Perception Action and Cognition (NatPAC) 7T fMRI dataset, and discuss how this dataset allows investigation of the impact of goal-directed actions on visual representations under naturalistic settings. Ian Charest will present recent results on semantic representations enabled by NSD, as well as ongoing large-scale data collection efforts inspired by NSD. Tomas Knapen will demonstrate how combining LSVNDs with other datasets incites exploration and discovery, and will present ongoing large-scale data collection efforts in his group. Jacob Prince will provide a first-hand perspective on how researchers external to the data collection process can apply LSVNDs for diverse research aims across cognitive neuroscience, neuroAI, and neuroimaging methods development. Finally, Ale Gifford will highlight broad opportunities that LSVNDs offer to the vision sciences community, and present a vision for the future of large-scale datasets. This symposium will interest any VSS member interested in neural data as it will expose opportunities and limitations of LSVNDs and how they relate to smaller, more narrowly focused datasets. Our goal is to align the VSS community with respect to open questions regarding LSVNDs, and help incentivize and coordinate new large-scale data collection efforts. We believe this symposium will strengthen the impact of LSVNDs on the field of vision science, and foster a new generation of big-data vision scientists.

TALK 1

THE NATURAL SCENES DATASET: LESSONS LEARNED AND WHAT'S NEXT?

Eline R. Kupers[1,2], Celia Durkin[2], Clayton E Curtis[3], Harvey Huang[4], Dora Hermes[4], Thomas Naselaris[2], Kendrick Kay[2]; [1]Stanford University, [2]University of Minnesota, [3]New York University, [4]Mayo Clinic

Release and reuse of rich neuroimaging datasets have rapidly grown in popularity, enabling researchers to ask new questions about visual processing and to benchmark computational models. One highly used dataset is the Natural Scenes Dataset (NSD), a 7T fMRI dataset where 8 subjects viewed more than 70,000 images over the course of a year. Since its recent release in September 2021, NSD has gained 1700+ users and resulted in 55+ papers and pre-prints. Here, we share behind-the-scenes considerations and inside knowledge from the NSD acquisition effort that helped ensure its quality and impact. This includes lessons learned regarding funding, designing, collecting, and releasing a large-scale fMRI dataset. Complementing the creator's perspective, we also highlight the user's viewpoint by revealing results from a large anonymous survey distributed amongst NSD users. These results will provide valuable (and often unspoken) insights into both positive and negative experiences interacting with NSD and other publicly available datasets. Finally, we discuss ongoing efforts towards two new large-scale datasets: (i) NSD-iEEG, an intracranial electroencephalography dataset with extensive electrode coverage in cortex and sub-cortex using a similar paradigm to NSD and (ii) Visual Cognition Dataset, a 7T fMRI dataset that samples a large diversity of tasks on a common set of visual stimuli (in contrast to NSD which samples a large diversity of stimuli during a single task). By sharing these lessons and ideas, we hope to facilitate new data collection efforts and enhance the ability of these datasets to support new discoveries in vision and cognition.

TALK 2

EXPLORING NATURALISTIC VISION IN ACTION WITH THE 7T NATURALISTIC PERCEPTION, ACTION, AND COGNITION (NATPAC) DATASET

Won Mok Shim[1,2], Royoung Kim[1,2], Jiwoong Park[1,2]; [1]Institute of Basic Science, Republic of Korea, [2]Sungkyunkwan University

Large-scale human neuroimaging datasets have provided invaluable opportunities to examine brain and cognitive functions. Our recent endeavor, the 7T NatPAC project, is designed to provide high-resolution human MRI structural and functional datasets using moderately dense sampling (12–16 2-hr sessions per subject) across a broad range of tasks. While previous large-scale datasets have featured sparse sampling of cognitive functions, our goal is to encompass a more extensive spectrum of cognitive and affective processes through diverse tasks, spanning both structured and naturalistic paradigms. Notably, we incorporated naturalistic tasks to probe a variety of higher-order cognitive functions including watching movies, freely speaking, and interactive 3D video game playing within a Minecraft environment. Through a collection of innovative Minecraft-based games simulating real-world behaviors, we aim to investigate the neural mechanisms of perception, action, and cognition as an integrative process that unfolds in naturalistic contexts. In this talk, I will focus on a shepherding game, where participants engage in strategic planning with hierarchical subgoals and adaptively update their strategies while navigating a virtual world. In combination with high-precision eye tracking data corrected for head motion, we explore how visual responses, including population receptive field (pRF) mapping, are modulated in the visual cortex and frontoparietal regions during free viewing and complex goal-directed behaviors compared to passive viewing of game replays and conventional pRF experiments. I will discuss the broader implications of the impact of goal-directed actions on visual representations and how large-scale datasets enable us to examine such effects in naturalistic settings.

TALK 3

EXPLOITING LARGE-SCALE NEUROIMAGING DATASETS TO REVEAL NOVEL INSIGHTS IN VISION SCIENCE

Ian Charest[1,2], Peter Brotherwood[1], Catherine Landry[1], Jasper van den Bosch[1], Shahab Bakhtiari[1,2], Tim Kietzmann[3], Frédéric Gosselin[1], Adrien Doerig[3]; [1]Université de Montréal, [2]Mila - Québec AI Institute, [3]University of Osnabrück

Building quantitative models of neural activity in the visual system is a long-standing goal in neuroscience. Though this research program is fundamentally limited by the small scale and low signal-to-noise of most existing datasets, with the advent of large-scale datasets it has become possible to build, test, and discriminate increasingly expressive competing models of neural representation. In this talk I will describe how the scale of the 7T fMRI Natural Scenes Dataset (NSD) has made possible novel insights into the mechanisms underlying scene perception. We harnessed recent advancements in linguistic artificial intelligence to construct models that capture progressively richer semantic information, ranging from object categories to word embeddings to scene captions. Our findings reveal a positive correlation between a model's capacity to capture semantic information and its ability to predict NSD data, a feature then replicated with recurrent convolutional networks trained to predict sentence embeddings from visual inputs. This collective evidence suggests that the visual system, as a whole, is better characterized by an aim to extract rich semantic information rather than merely cataloging object inventories from visual inputs. Considering the substantial power of NSD, collecting additional neuroimaging and behavioral data using the same image set becomes highly appealing. We are expanding NSD through the development of two innovative datasets: an electroencephalography dataset called NSD-EEG, and a mental imagery vividness ratings dataset called NSD-Vividness. Datasets like NSD not only provide fresh insights into the visual system but also inspire the development of new datasets in the field.

TALK 4

FAREWELL TO THE EXPLORE-EXPLOIT TRADE-OFF IN LARGE-SCALE DATASETS

Tomas Knapen[1,2], Nick Hedger[3], Thomas Naselaris[4], Shufan Zhang[1,2], Martin Hebart[5,6]; [1]Vrije Universiteit, [2]Royal Dutch Academy of Arts and Sciences, [3]University of Reading, [4]University of Minnesota, [5]Justus Liebig University, [6]Max Planck Institute for Human Cognitive and Brain Sciences

LSVNDs are a very powerful tool for discovery science. Due to their suitability for exploration, large datasets synergize well when supplemented with more exploitative datasets focused on small-scale hypothesis testing that can confirm exploratory findings. Similar synergy can be attained when combining findings across datasets, where one LSVND can be used to confirm and extend discoveries from another LSVND. I will showcase how we have recently leveraged several large-scale datasets in unison to discover principles of topographic visual processing throughout the brain. These examples demonstrate how LSVNDs can be used to great effect, especially in combination across datasets. In our most recent example, we combined the HCP 7T fMRI dataset (a "wide" dataset with 180 participants, 2.5 hrs of whole-brain fMRI each) with NSD (a "deep" dataset with 8 participants, 40 hrs of whole-brain fMRI each) to investigate visual body-part selectivity. We discovered homuncular maps in high-level visual cortex through connectivity with primary somatosensory cortex in HCP, and validated the body-part tuning of these maps using NSD. This integration of wide and deep LSVNDs allows inference about computational mechanisms at both the individual and population levels. For this reason, we believe the field needs a variety of LSVNDs. I will briefly present ongoing work from my lab collecting new ‘deep’ LSVND contributions: a brief (2.5-s) video watching dataset and a retinotopic mapping dataset, each with up to 10 sessions of 7T fMRI in 8 subjects.

TALK 5

LARGE DATASETS: A SWISS ARMY KNIFE FOR DIVERSE RESEARCH AIMS IN NEUROAI

Jacob Prince[1], Colin Conwell[2], Talia Konkle[1]; [1]Harvard University, [2]Johns Hopkins University

This talk provides a first-hand perspective on how users external to the data collection process can harness LSVNDs as foundation datasets for their research aims. We first highlight recent evidence that these datasets help address and move beyond longstanding debates in cognitive neuroscience, such as the nature of category selective regions, and the visual category code more broadly. We will show evidence that datasets like NSD have provided powerful new insight into how items from well-studied domains (faces, scenes) are represented in the context of broader representational spaces for objects. Second, we will highlight the potential of LSVNDs to answer urgent, emergent questions in neuroAI – for example, which inductive biases are critical for obtaining a good neural network model of the human visual system? We will describe a series of controlled experiments leveraging hundreds of open-source DNNs, systematically varying inductive biases to reveal the factors that most directly impact brain predictivity at scale. Finally, for users interested in neuroimaging methods development, we will highlight how the existence of these datasets has catalyzed rapid progress in methods for fMRI signal estimation and denoising, as well as for basic analysis routines like PCA and computing noise ceilings. We will conclude by reflecting on both the joys and pain points of working with LSVNDs, in order to help inform the next generation of these datasets.

TALK 6

WHAT OPPORTUNITIES DO LARGE-SCALE VISUAL NEURAL DATASETS OFFER TO THE VISION SCIENCES COMMUNITY?

Alessandro T. Gifford[1], Benjamin Lahner[2], Pablo Oyarzo[1], Aude Oliva[2], Gemma Roig[3], Radoslaw M. Cichy[1]; [1]Freie Universität Berlin, [2]MIT, [3]Goethe Universität Frankfurt

In this talk I will provide three complementary examples of the opportunities that LSVNDs offer to the vision sciences community. First, LSVNDs of naturalistic (thus more ecologically valid) visual stimulation allow the investigation of novel mechanisms of high-level visual cognition. We are extensively recording human fMRI and EEG responses for short naturalistic movie clips; modeling results reveal that semantic information such as action understanding or movie captions is embedded in neural representations. Second, LSVNDs contribute to the emerging field of NeuroAI, advancing research in vision sciences through a symbiotic relationship between visual neuroscience and computer vision. We recently collected a large and rich EEG dataset of neural responses to naturalistic images, using it on the one hand to train deep-learning-based end-to-end encoding models directly on brain data, thus aligning visual representations in models and the brain, and on the other hand to increase the robustness of computer vision models by exploiting inductive biases from neural visual representations. Third, LSVNDs make possible critical initiatives such as challenges and benchmarks. In 2019 we founded the Algonauts Project, a platform where scientists from different disciplines can cooperate and compete in creating the best predictive models of the visual brain, thus advancing the state-of-the-art in brain modeling as well as promoting cross-disciplinary interaction. I will end with some forward-looking thoughts on how LSVNDs might transform the vision sciences.

Symposium: Friday, May 17, 2024, 2:30 – 4:30 pm, Talk Room 1

The temporal evolution of visual perception

Organizers: Lina Teichmann[1], Chris Baker[1]; [1]Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, USA
Presenters: Lina Teichmann, Iris I. A. Groen, Diana Dima, Tijl Grootswagers, Rachel Denison

The human visual system dynamically processes input over the course of a few hundred milliseconds to generate our perceptual experience. Capturing the dynamic aspects of the neural response is therefore imperative to understand visual perception. By bringing five speakers together who use a diverse set of methods and approaches, the symposium aims to elucidate the temporal evolution of visual perception from different angles. All five speakers (four female) are early-career researchers based in Europe, Australia, the US, and Canada. Speakers will be allotted 18 minutes of presentation time plus 5 minutes of questions after each talk. In contrast to a lot of the current neuroimaging work, the symposium talks will focus on temporal dynamics rather than localization. Collectively, the work presented will demonstrate that the complex and dynamic nature of visual perception requires data that matches its temporal granularity. In the first talk, Lina Teichmann will present data from a large-scale study focusing on how individual colour-space geometries unfold in the human brain. Linking densely-sampled MEG data with psychophysics, her work on colour provides a test case to study the subjective nature of visual perception. Iris Groen will discuss findings from intracranial EEG studies that characterize neural responses across the visual hierarchy. Applying computational models, her work provides fundamental insights into how the visual response unfolds over time across visual cortex. Diana Dima will speak about how responses evoked by observed social interactions are processed in the brain. Using temporally-resolved EEG data, her research shows how visual information is modulated from perception to cognition. Tijl Grootswagers will present on studies investigating visual object processing. Using rapid series of object stimuli and linking EEG and behavioural data, his work shows the speed and efficiency of the visual system to make sense of the things we see. To conclude, Rachel Denison will provide insights into how we employ attentional mechanisms to prioritize relevant visual input at the right time. Using MEG data, she will highlight how temporal attention affects the dynamics of evoked visual responses. Overall, the symposium aims to shed light on the dynamic nature of visual processing at all levels of the visual hierarchy. It will be a chance to discuss benefits and challenges of different methodologies that will allow us to gain a comprehensive insight into the temporal aspects of visual perception.

Talk 1

THE TEMPORAL DYNAMICS OF INDIVIDUAL COLOUR-SPACE GEOMETRIES IN THE HUMAN BRAIN

Lina Teichmann[1], Ka Chun Lam[2], Danny Garside[3], Amaia Benitez-Andonegui[4], Sebastian Montesinos[1], Francisco Pereira[2], Bevil Conway[3,5], Chris Baker[1,5]; [1]Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, USA, [2]Machine Learning Team, National Institute of Mental Health, Bethesda, USA, [3]Laboratory of Sensorimotor Research, National Eye Institute, Bethesda, USA, [4]MEG Core Facility, National Institute of Mental Health, Bethesda, USA, [5]equal contribution

We often assume that people see the world in a similar way to us, as we can effectively communicate how things look. However, colour perception is one aspect of vision that varies widely among individuals as shown by differences in colour discrimination, colour constancy, colour appearance and colour naming. Further, the neural response to colour is dynamic and varies over time. Many attempts have been made to construct formal, uniform colour spaces that aim to capture universally valid similarity relationships, but there are discrepancies between these models and individual perception. Combining Magnetoencephalography (MEG) and psychophysical data we examined the extent to which these discrepancies can be accounted for by the geometry of the neural representation of colour and their evolution over time. In particular, we used a dense sampling approach and collected neural responses to hundreds of colours to reconstruct individual fine-grained colour-space geometries from neural signals with millisecond accuracy. In addition, we collected large-scale behavioural data to assess perceived similarity relationships between different colours for every participant. Using a computational modelling approach, we extracted similarity embeddings from the behavioural data to model the neural signal directly. We find that colour information is present in the neural signal from approximately 70 ms onwards but that neural colour-space geometries unfold non-uniformly over time. These findings highlight the gap between theoretical colour spaces and colour perception and represent a novel avenue to gain insights into the subjective nature of perception.

Talk 2

DELAYED DIVISIVE NORMALISATION ACCOUNTS FOR A WIDE RANGE OF TEMPORAL DYNAMICS OF NEURAL RESPONSES IN HUMAN VISUAL CORTEX

Iris I. A. Groen[1], Amber Brands[1], Giovanni Piantoni[2], Stephanie Montenegro[3], Adeen Flinker[3], Sasha Devore[3], Orrin Devinsky[3], Werner Doyle[3], Patricia Dugan[3], Daniel Friedman[3], Nick Ramsey[2], Natalia Petridou[2], Jonathan Winawer[4]; [1]Informatics Institute, University of Amsterdam, Amsterdam, Netherlands, [2]University Medical Center Utrecht, Utrecht, Netherlands, [3]New York University Grossman School of Medicine, New York, NY, USA, [4]Department of Psychology and Center for Neural Science, New York University, New York, NY, USA

Neural responses in visual cortex exhibit various complex, non-linear temporal dynamics. Even for simple static stimuli, responses decrease when a stimulus is prolonged in time (adaptation), reduce to stimuli that are repeated (repetition suppression), and rise more slowly for low contrast stimuli (slow dynamics). These dynamics also vary depending on the location in the visual hierarchy (e.g., lower vs. higher visual areas) and the type of stimulus (e.g., contrast pattern stimuli vs. real-world object, scenes and face categories). In this talk, I will present two intracranial EEG (iEEG) datasets in which we quantified and modelled the temporal dynamics of neural responses across the visual cortex at millisecond resolution. Our work shows that many aspects of these dynamics are accurately captured by a delayed divisive normalisation model in which neural responses are normalised by recent activation history. I will highlight how fitting this model to the iEEG data unifies multiple disparate temporal phenomena in a single computational framework, thereby revealing systematic differences in temporal dynamics of neural population responses across the human visual hierarchy. Overall, these findings suggest a pervasive role of history-dependent delayed divisive normalisation in shaping neural response dynamics across the cortical visual hierarchy.

TALK 3

HOW NATURAL ACTION PERCEPTION UNFOLDS IN THE BRAIN

Diana Dima[1], Yalda Mohsenzadeh[1]; [1]Western University, London, ON, Canada

In a fraction of a second, humans can recognize a wide range of actions performed by others. Yet actions pose a unique complexity challenge, bridging visual domains and varying along multiple perceptual and semantic features. What features are extracted in the brain when we view others' actions, and how are they processed over time? I will present electroencephalography work using natural videos of human actions and rich feature sets to determine the temporal sequence of action perception in the brain. Our work shows that action features, from visual to semantic, are extracted along a temporal gradient, and that different processing stages can be dissociated with artificial neural network models. Furthermore, using a multimodal approach with video and text stimuli, we show how conceptual action representations emerge in the brain. Overall, these data reveal the rapid computations underlying action perception in natural settings. The talk will highlight how a temporally resolved approach to natural vision can uncover the neural computations linking perception and cognition.

TALK 4

DECODING RAPID OBJECT REPRESENTATIONS

Tijl Grootswagers[1], Amanda K. Robinson[2]; [1]The MARCS Institute for Brain, Behaviour and Development, School of Computer, Data and Mathematical Sciences, Western Sydney University, Sydney, NSW, Australia, [2]Queensland Brain Institute, The University of Queensland, Brisbane, QLD, Australia

Humans are extremely fast at recognising objects, and can do this very reliably. Information about objects and object categories emerges within 200 milliseconds in the human visual system, even under difficult conditions such as occlusion or low visibility. These neural representations can be highly complex and multidimensional, despite relying on limited visual information. Understanding emerging object representations necessitates time-resolved neuroimaging methods with millisecond precision, such as EEG and MEG. Recent time-resolved neuroimaging work has used decoding methods in rapid serial visual presentation designs to show that relevant object-information about multiple sequentially presented objects is robustly encoded by the brain. This talk will highlight recent research on the time course of object representations in rapid image sequences, focusing on three key findings: (1) object representations are highly automatic, with robust representations emerging even with fast-changing visual input. (2) emerging object representations are highly robust to changes in context and task, suggesting strong reliance on feedforward processes. (3) object representational structures are highly consistent across individuals, to the extent that neural representations are predictive of independent behavioural judgments on a variety of tasks. Together, these findings suggest that the first sweep of information through the visual system contains highly robust information that is readily available for read-out in behavioural decisions.

TALK 5

ISOLATING NEURAL MECHANISMS OF VOLUNTARY TEMPORAL ATTENTION

Rachel Denison[1,2], Karen Tian[1,2], Jiating Zhu[1], David Heeger[2], Marisa Carrasco[2]; [1]Boston University, Department of Psychological and Brain Sciences, USA, [2]New York University, Department of Psychology and Center for Neural Science, USA

To handle the continuous influx of visual information, temporal attention prioritizes visual information at task-relevant moments in time. We first introduce a probabilistic framework that clarifies the conceptual distinction and formal relation between temporal attention, linked to timing relevance, and temporal expectation, linked to timing predictability. Next, we present two MEG studies in which we manipulated temporal attention while keeping expectation constant, allowing us to isolate neural mechanisms specific to voluntary temporal attention. Participants were cued to attend to one of two sequential grating targets with predictable timing, separated by a 300 ms SOA. The first study used time-resolved steady-state visual evoked responses (SSVER) to investigate how temporal attention modulates anticipatory visual activity. In the pre-target period, visual activity (measured with a background SSVER probe) steadily ramped up as the targets approached, reflecting temporal expectation. Furthermore, we found a low-frequency modulation of visual activity, which shifted approximately 180 degrees in phase according to which target was attended. The second study used time-resolved decoding and source reconstruction to examine how temporal attention affects dynamic target representations. Temporal attention to the first target enhanced its orientation representation within a left fronto-cingulate region ~250 ms after stimulus onset, perhaps protecting it from interference from the second target within the visual cortex. Together these studies reveal how voluntary temporal attention flexibly shapes pre-target periodic dynamics and post-target routing of stimulus information to select a task-relevant stimulus within a sequence.

Symposium: Friday, May 17, 2024, 2:30 – 4:30 pm, Talk Room 2

Attention: accept, reject, or major revisions?

Organizers: Alon Zivony[1]; [1]University of Sheffield
Presenters: Britt Anderson, Ruth Rosenholtz, Wayne Wu, Sarah Shomstein, Alon Zivony

Is attention research in crisis? After more than a century, we have come full circle from the intuition that "everybody knows what attention is" (James, 1890) to the conclusion that "nobody knows what attention is" (Hommel et al., 2019). It has been suggested that attention is an incoherent and sterile concept, or unsuitable for scientific research. And yet, attention research continues as strongly as ever with little response to these critiques. Is the field ignoring glaring theoretical problems, or does the current conception of attention merely require some revisions? In this symposium, our speakers bring different perspectives to examine this critical question. Rather than merely raising issues with the concept of attention, each also suggests practical and theoretical solutions, which can hopefully inform future research. Each speaker will present either a critical view or defence of the concept of attention, and suggest whether attention should be abandoned, kept as is, or redefined. Our first two speakers will argue that scientists may be better off without the concept of attention. Britt Anderson will criticize the use of attention as an explanation of observed phenomena. He will suggest that the common usage is non-scientific and results in circular logic. He offers in its place an attention-free account of so-called attention effects. Ruth Rosenholtz argues that recent work, for example on peripheral vision, calls into question many of the basic tenets of attention theory. She will talk about her year of banning 'attention' in order to rethink attention from the ground up. The second group of speakers will question common understanding of attention but will argue in favour of it as a scientific concept. Wayne Wu will suggest that our shared methodology of studying attention commits us to the Jamesonian functional conceptualization of attention. He will argue that attention can and should be retained if we locate it in the right level analysis in cognitive explanation. Sarah Shomstein will discuss "attentional platypuses", empirical abnormalities that do not fit into current attention research. These abnormalities reveal the need for a new way of thinking about attention. Alon Zivony will argue that many of the conceptual problems with attention stem from the standard view that equates attention with selection. Moving away from this definition will allow us to retain attention but will also require a change in our thinking. Each talk will conclude with a take-home message about what attention is and isn't, a verdict of whether it should be abandoned or retained, and suggestions of how their understanding of attention can be applied in future research. We will conclude with a panel discussion.

Talk 1

ATTENTION: IDOL OF THE TRIBE

Britt Anderson[1]; [1]Dept of Psychology and Centre for Theoretical Neuroscience, University of Waterloo

The term 'attention' has been a drag on our science ever since the early days of experimental psychology. Our frequent offerings and sacrifices (articles and the debates they provoke), and our unwillingness to abandon our belief in this reified entity indicates the aptness of the Jamesian phrase "idol of the tribe." While causal accounts of attention are empty, attention might be, as suggested by Hebb, a useful label. It could be used to indicate that some experimental observable is not immediately explained by the excitation of receptor cells. However, labeling of something as 'attention' means there is something to be explained; not that something has been explained. Common experimental manipulations used to provoke visual selective attention: instructions, cues, and reward are in fact the guide to explaining away 'attention'. The observations provoked by such manipulations frequently induce behavioral performance differences not explainable in terms of differences in retinal stimulation. These manipulations are economically summarized as components of a process in which base rates, evidence, value, and plausibility combine to determine perceptual experience. After briefly reviewing the history of how attention has been confusing from the start, I will summarize the notion of conceptual fragmentation and show how it applies. I will then review how the traditional conditions of an attentional experiment provide the basis for a superior, attention free, account of the phenomena of interest, and I will present some of the opportunities for the use of more formal descriptions that should lead to better theoretically motivated experimental investigations.

Talk 2

ATTENTION IN CRISIS

Ruth Rosenholtz[1]; [1]NVIDIA Research

Recent research on peripheral vision has led to a paradigm-shifting conclusion: that vision science as a field must rethink the concept of visual attention. Research has uncovered significant anomalies not explained by existing theories, and some methods for studying attention may instead have uncovered mechanisms of peripheral vision. Nor can a summary statistic representation in peripheral vision solve these problems on its own. A year of banning "attention" in my lab allowed us to rethink attention from the ground up; this talk will conclude with some of the resulting insights.

Talk 3

ATTENTION UNIFIED

Wayne Wu[1]; [1]Department of Philosophy and Neuroscience Institute, Carnegie Mellon University

For over a century, scientists have expressed deep misgivings about attention. A layperson would find this puzzling, for they know what attention is as well as those with sight know what seeing is. People visually attend all the time. Attention is real, we know what it is, and we can explain it. I shall argue that the problem of attention concerns the conceptual and logical structure of the scientific theory of attention. Because of shared methodology, we are committed to a single functional conception of attention, what William James articulated long ago. I show how this shared conception provides a principle of unification that links empirical work. To illustrate this, I show how two cueing paradigms tied to "external" and "internal" attention, spatial cueing and retro-cueing, are instances of the same kind of attention. Against common skepticism, I demonstrate that we are all committed to the existence of attention as a target of explanation. Yet in step with the skeptic, I show that attention is not an explainer in the sense that it is not a neural mechanism. Locating attention at the right level of analysis in cognitive explanation is key to understanding what it is and how science has made massive progress in understanding it.

TALK 4

WHAT DOES A PLATYPUS HAVE TO DO WITH ATTENTION?

Sarah Shomstein[1]; [1]Department of Psychological and Brain Sciences, George Washington University

Decades of research on understanding the mechanisms of attentional selection have focused on identifying the units (representations) on which attention operates in order to guide prioritized sensory processing. These attentional units fit neatly to accommodate our understanding of how attention is allocated in a top-down, bottom-up, or historical fashion. In this talk, I will focus on attentional phenomena that are not easily accommodated within current theories of attentional selection. We call these phenomena attentional platypuses, as they allude to an observation that within biological taxonomies the platypus does not fit into either mammal or bird categories. Similarly, attentional phenomena that do not fit neatly within current attentional models suggest that current models need to be revised. We list a few instances of the ‘attentional platypuses’ and then offer a new approach, that we term Dynamically Weighted Prioritization, stipulating that multiple factors impinge onto the attentional priority map, each with a corresponding weight. The interaction between factors and their corresponding weights determine the current state of the priority map which subsequently constrains/guides attention allocation. We propose that this new approach should be considered as a supplement to existing models of attention, especially those that emphasize categorical organizations.

TALK 5

IT’S TIME TO REDEFINE ATTENTION

Alon Zivony[1]; [1]Department of Psychology, University of Sheffield

Many models of attention assume that attentional selection takes place at a specific moment in time which demarcates the critical transition from pre-attentive to attentive processing of sensory inputs. In this talk, I will argue that this intuitively appealing assumption is not only incorrect, but it is also the reason behind the conceptual confusion about what attention is, and how it should be understood in psychological science. As an alternative, I will offer a “diachronic” framework that views attention as a modulatory process that unfolds over time, in tandem with perceptual processing. This framework breaks down the false dichotomy between pre-attentive and attentive processing, and as such, offers new solutions to old problems in attention research (the early vs. late selection debate). More importantly, by situating attention within a broader context of selectivity in the brain, the diachronic account can provide a unified and conceptually coherent account of attention. This will allow us to keep the concept of attention but will also require serious rethinking about how we use attention as a scientific concept.

SYMPOSIUM: FRIDAY, MAY 17, 2024, 5:00 – 7:00 PM, TALK ROOM 1

The Multifaceted effects of blindness and how sight might be restored

Organizer: Ella Striem-Amit[1]; [1]Georgetown University
Presenters: Lara Coelho, Santani Teng, Woon Ju Park, Elizabeth J. Saccone, Ella Striem-Amit, Michael Beyeler

Congenital blindness illustrates the developmental roots of visual cortex functions. Here, a group of early-career researchers will present various perspectives on the multifaceted effects of blindness on the brain and behavior. To start off the symposium, Coelho will describe the effect of sight loss on multisensory properties, and the reliance on vision to develop an intact multisensory body representation. This presentation will highlight the dependence across modalities, revealing rich interactions between vision and body representations. Discussing a unique manifestation of compensation in blindness, Teng will discuss how echolocation functions in naturalistic settings and its properties of active sensing. Continuing the theme of integration across senses and diving into visual cortical reorganization, Park will argue for partial dependence and partial independence on vision for the development of motion processing in hMT+. Saccone will show evidence for a functional takeover of language over typically face-selective FFA in blindness, showing plasticity beyond sensory representations. Together, these two talks will highlight different views of brain plasticity in blindness. Adding to our discussion of the multifaceted nature of plasticity, Striem-Amit will discuss whether plasticity in the visual cortex is consistent across different blind individuals, showing evidence for divergent visual plasticity and stability over time in adulthood. The last speaker will discuss the challenges and potential for sight restoration using visual prostheses. Beyeler will discuss how some of the challenges of sight restoration can be addressed through perceptual learning of implant inputs. This talk highlights how understanding plasticity in the visual system and across the brain has direct applications for successfully restoring sight. Together, the symposium will bring different theoretical perspectives to illustrate the effects of blindness, revealing the extent and diversity of neural plasticity, and clarify the state-of-the-art capacities for sight restoration.

TALK 1

IMPLICATIONS OF VISUAL IMPAIRMENT ON BODY REPRESENTATION

Lara Coelho[1], Monica Gori; [1]Unit for visually impaired people, Italian Institute of Technology, Genova, Italy

In humans, vision is the most accurate sensory modality for constructing our representation of space. It has been shown that visual impairment negatively influences daily living and quality of life. For example, spatial and locomotor skills are reduced in this population.

One possibility is that these deficiencies arise from a distorted representation of the body. Body representation is fundamental for motor control, because we rely on our bodies as a metric guide for our actions. While body representation is a by-product of multisensory integration, it has been proposed that vision is necessary to construct an accurate representation of the body. In the MySpace project, we are investigating the role of visual experience on haptic body representations in sighted and visually impaired (VI) participants. To this end, we employ a variety of techniques to investigate two key aspects of body representation 1) size perception, and 2)the plasticity of the proprioceptive system. These techniques include landmark localization, psychophysics, and the rubber hand illusion. Our results in sighted participants show distortions in haptic but not visual body representation. In the VI participants there are distortions when estimating forearm, hand, and foot size in several different haptic tasks. Moreover, VI children fail to update their perceived body location in the rubber hand illusion task. Collectively, our findings support the hypothesis that vision is necessary to reduce distortions in haptic body representations. Moreover, we propose, that VI children may develop with impaired representations of their own bodies. We discuss possible opportunities for reducing this impairment.

TALK 2

ACOUSTIC GLIMPSES: THE ACCUMULATION OF PERCEPTUAL INFORMATION IN BLIND ECHOLOCATORS

Santani Teng[1]; [1]Smith-Kettlewell Eye Research Institute

Blindness imposes constraints on the acquisition of sensory information from the environment. To mitigate those constraints, some blind people employ active echolocation, a technique in which self-generated sounds, like tongue “clicks,” produce informative reflections. Echolocating observers integrate over multiple clicks, or samples, to make perceptual decisions that guide behavior. What information is gained in the echoacoustic signal from each click? Here, I will draw from similar work in eye movements and ongoing studies in our lab to outline our approaches to this question. In a psychoacoustic and EEG experiment, blind expert echolocators and sighted control participants localized a virtual reflecting object after hearing simulated clicks and echoes. Left-right lateralization improved on trials with more click repetitions, suggesting a systematic precision benefit to multiple samples even when each sample delivered no new sensory information. In a related behavioral study, participants sat in a chair but otherwise moved freely while echoacoustically detecting, then orienting toward a reflecting target located at a random heading in the frontal hemifield. Clicking behavior and target size (therefore sonar strength) strongly influenced the rate and precision of orientation convergence toward the target, indicating a dynamic interaction between motor-driven head movements, click production, and the resulting echoacoustic feedback to the observer. Taken together, modeling these interactions in blind expert practitioners suggests similar properties, and potential shared mechanisms, between active sensing behavior in visual and echoacoustic domains.

TALK 3

CONSTRAINTS OF CROSS-MODAL PLASTICITY WITHIN HMT+ FOLLOWING EARLY BLINDNESS

Woon Ju Park[1], Kelly Chang, Ione Fine; [1]Department of Psychology, University of Washington

Cross-modal plasticity following early blindness has been widely documented across numerous visual areas, highlighting our brain’s remarkable adaptability to changes in sensory environment. In many of these areas, functional homologies have been observed between the original and reorganized responses. However, the mechanisms driving these homologies remain largely unknown. Here, we will present findings that aim to answer this question within the area hMT+, which responds to visual motion in sighted individuals and to auditory motion in early blind individuals. Our goal was to examine how the known functional and anatomical properties of this area influence the development of cross-modal responses in early blind individuals. Using a multimodal approach that encompasses psychophysics, computational modeling, and functional and quantitative MRI, we simultaneously characterized perceptual, functional, and anatomical selectivity to auditory motion within early blind and sighted individuals. We find that some anatomical and functional properties of hMT+ are inherited, while others are altered in those who become blind early in life.

TALK 4

VISUAL EXPERIENCE IS NECESSARY FOR DISSOCIATING FACE- AND LANGUAGE-PROCESSING IN THE VENTRAL VISUAL STREAM

Elizabeth J. Saccone[1], Akshi[1], Judy S. Kim[2], Mengyu Tian[3], Marina Bedny[1]; [1]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA, [2]Center for Human Values, Princeton University, Princeton, NJ, USA, [3]Center for Educational Science and Technology, Beijing Normal University at Zhuhai, China

The contributions of innate predispositions versus experience to face-selectivity in vOTC is hotly debated. Recent studies with people born blind suggest face specialization emerges regardless of experience. In blindness the FFA is said to process face shape, accessed through touch or sound, or maintain its behavioral role in person recognition by specializing for human voices. We hypothesized instead that in blind people the anatomical location of the FFA responds to language. While undergoing fMRI, congenitally blind English speakers (N=12) listened to spoken language (English), foreign speech (Russian, Korean, Mandarin), non-verbal vocalizations (e.g., laughter) and control non-human scene sounds (e.g., forest sounds) during a 1-back repetition task. Participants also performed a ‘face localizer’ task by touching 3D printed models of faces and control scenes and a language localizer (spoken words > backwards speech, Braille > tactile shapes). We identified individual-subject ROIs inside a FFA mask generated from sighted data. In people born blind, the anatomical location of the FFA showed a clear preference for language over all other sounds, whether human or not. Responses to spoken language were higher than to foreign speech or non-verbal vocalizations, which were not different from scene sounds. This pattern was observed even in parts of vOTC that responded more to touching faces. Specialization for faces in vOTC is influenced by experience. In the absence of vision, lateral vOTC becomes implicated in language. We speculate that shared circuits that evolved for communication specialize for either face recognition or language depending on experience.

Talk 5

INDIVIDUAL DIFFERENCES OF BRAIN PLASTICITY IN EARLY VISUAL DEPRIVATION

Ella Striem-Amit[1]; [1]Department of Neuroscience, Georgetown University Medical Center, Washington, DC 20057, USA

Early-onset blindness leads to reorganization in visual cortex connectivity and function. However, this has mostly been studied at the group level, largely ignoring differences in brain reorganization across early blind individuals. To test whether plasticity manifests differently in different blind individuals, we studied resting-state functional connectivity (RSFC) from the primary visual cortex in a large cohort of blind individuals. We find increased individual differences in connectivity patterns, corresponding to areas that show reorganization in blindness. Further, using a longitudinal approach in repeatedly sampled blind individuals, we showed that such individual patterns of organization and plasticity are stable over time, to the degree of decoding individual participant identity over 2 years. Together, these findings suggest that visual cortex reorganization is not ubiquitous, highlighting the potential diversity in brain plasticity and the importance of harnessing individual differences for fitting rehabilitation approaches for vision loss.

Talk 6

LEARNING TO SEE AGAIN: THE ROLE OF PERCEPTUAL LEARNING AND USER ENGAGEMENT IN SIGHT RESTORATION

Michael Beyeler[1]; [1]University of California, Santa Barbara

Retinal and cortical implants show potential in restoring a rudimentary form of vision to people living with profound blindness, but the visual sensations (“phosphenes”) produced by current devices often seem unnatural or distorted. Consequently, the ability of implant users to learn to make use of this artificial vision plays a critical role in whether some functional vision is successfully regained. In this talk, I will discuss recent work detailing the potential and limitations of perceptual learning in helping implant users learn to see again. Although the abilities of visual implant users tend to improve with training, there is little evidence that this is due to distortions becoming less perceptually apparent, but instead may be due to better interpretation of distorted input. Unlike those with natural vision, implant recipients must accommodate various visual anomalies, such as inconsistent spatial distortions and phosphene fading. Furthermore, perceptual measures such as grating acuity and motion discrimination, which are often used with the intention of objectively assessing visual function, may be modulated via gamification, highlighting the importance of user engagement in basic psychophysical tasks. Gamification may be particularly effective at engaging reward systems in the brain, potentially fostering greater plasticity through more varied stimuli and active attentional engagement. However, the effectiveness of such gamified approaches varies, suggesting a need for personalized strategies in visual rehabilitation.

Symposium: Friday, May 17, 2024, 5:00 – 7:00 pm, Talk Room 2

Using deep networks to re-imagine object-based attention and perception

Organizers: Hossein Adeli[1], Seoyoung Ahn[2], Gregory Zelinsky[2]; [1]Columbia University, [2]Stony Brook University
Presenters: Patrick Cavanagh, Frank Tong, Paolo Papale, Alekh Karkada Ashok, Hossein Adeli, Melissa Le-Hoa Võ

What can Deep Neural Network (DNN) methods tell us about the brain mechanisms that transform visual features into object percepts? Using different state-of-the-art models, the speakers in this symposium will reexamine different cognitive and neural mechanisms of object-based attention (OBA) and perception and consider new computational mechanisms for how the visual system groups visual features into coherent object percepts. Our first speaker, Patrick Cavanagh, helped create the field of OBA and is therefore uniquely suited to give a perspective on how this question, essentially the feature-binding problem, has evolved over the years and has been shaped by paradigms and available methods. He will conclude by outlining his vision for how DNN architectures create new perspectives on understanding OBA. The next two speakers will review the recent behavioral and neural findings on object-based attention and feature grouping. Frank Tong will discuss the neural and behavioral signatures of OBA through the utilization of fMRI and eye tracking methods. He will demonstrate how the human visual system represents objects across the hierarchy of visual areas. Paolo Papale will discuss neurophysiological evidence for the role of OBA and grouping in object perception. Using stimuli systematically increasing in complexity from lines to natural objects (against cluttered backgrounds) he shows that OBA and grouping are iterative processes. Both talks will also include discussions of current modeling efforts, and what additional measures may be needed to realize more human-like object perception. The following two talks will provide concrete examples of how DNNs can be used to predict human behavior during different tasks. Lore Goetschalckx will focus on the importance of considering the time-course of grouping in object perception and will discuss her recent work on developing a method to analyze dynamics of different models. Using this method, she shows how a deep recurrent model trained on an object grouping task predicts human reaction time. Hossein Adeli will review modeling work on three theories of how OBA binds features into objects: one that implements object-files, another that uses generative processes to reconstruct an object percept, and a third model of spreading attention through association fields. In the context of these modeling studies, he will describe how each of these mechanisms was implemented as a DNN architecture. Lastly, Melissa Võ will drive home the importance of object representations and how they collectively create an object context that humans use to control their attention behavior in naturalistic settings. She shows how GANs can be used to study the hidden representations underlying our perception of objects. This symposium is timely because the advances in computational methods have made it possible to put old theories to the test and to develop new theories of OBA mechanisms that engage the role played by attention in creating object-centric representations.

TALK 1

THE ARCHITECTURE OF OBJECT-BASED ATTENTION

Patrick Cavanagh[1], Gideon P. Caplovitz[2], Taissa K. Lytchenko[2], Marvin R. Maechler[3], Peter U. Tse[3], David R. Sheinberg[4]; [1]Glendon College, York University, [2]University of Nevada, Reno, [3]Dartmouth College, [4]Brown University

Evidence for the existence of object-based attention raises several important questions: what are objects, how does attention access them, and what anatomical regions are involved? What are the "objects" that attention can access? Several studies have shown that items in visual search tasks are only loose collections of features prior to the arrival of attention. Nevertheless, findings from a wide variety of paradigms, including unconscious priming and cuing, have overturned this view. Instead, the targets of object-based attention appear to be fully developed object representations that have reached the level of identity prior to the arrival of attention. Where do the downward projections of object-based attention originate? Current research indicates that the control of object-based attention must come from ventral visual areas specialized in object analysis that project downward to early visual areas. If so, how can feedback from object areas accurately target the object's early locations and features when the object areas have only crude location information? Critically, recent work on autoencoders has made this plausible as they are capable of recovering the locations and features of the target objects from the high level, low dimensional codes in the object areas. I will outline the architecture of object-based attention, the novel predictions it brings, and discuss how it works in parallel with other attention pathways.

TALK 2

BEHAVIORAL AND NEURAL SIGNATURES OF OBJECT-BASED ATTENTION IN THE HUMAN VISUAL SYSTEM

Frank Tong[1], Sonia Poltoratski[1], David Coggan[1], Lasyapriya Pidaparthi[1], Elias Cohen[1]; [1]Vanderbilt University

How might one demonstrate the existence of an object representation in the visual system? Does objecthood arise preattentively, attentively, or via a confluence of bottom-up and top-down processes? Our fMRI work reveals that orientation-defined figures are represented by enhanced neural activity in the early visual system. We observe enhanced fMRI responses in the lateral geniculate nucleus and V1, even for unattended figures, implying that core aspects of scene segmentation arise from automatic perceptual processes. In related work, we find compelling evidence of object completion in early visual areas. fMRI response patterns to partially occluded object images resemble those evoked by unoccluded objects, with comparable effects of pattern completion found for unattended and attended objects. However, in other instances, we find powerful effects of top-down attention. When participants must attend to one of two overlapping objects (e.g., face vs. house), activity patterns from V1 through inferotemporal cortex are biased in favor of the covertly attended object, with functional coupling of the strength of object-specific modulation found across brain areas. Finally, we have developed a novel eye-tracking paradigm to predict the focus of object-based attention while observers view two dynamically moving objects that mostly overlap. Estimates of the precision of gaze following suggest that observers can entirely filter out the complex motion signals arising from the task-irrelevant object. To conclude, I will discuss whether current AI models can adequately account for these behavioral and neural properties of object-based attention, and what additional measures may be needed to realize more human-like object processing.

TALK 3

THE SPREAD OF OBJECT ATTENTION IN ARTIFICIAL AND CORTICAL NEURONS

Paolo Papale[1], Matthew Self[1], Pieter Roelfsema[1]; [1]Netherlands Institute for Neuroscience

A crucial function of our visual system is to group local image fragments into coherent perceptual objects. Behavioral evidence has shown that this process is iterative and time-consuming. A simple theory suggested that visual neurons can solve this challenging task relying on recurrent processing: attending to an object could produce a gradual spread of enhancement across its representation in the visual cortex. Here, I will present results from a biologically plausible artificial neural network that can solve object segmentation by attention. This model was able to identify and segregate individual objects in cluttered scenes with extreme accuracy, only using modulatory top-down feedback as observed in visual cortical neurons. Then, I will present comparable results from large-scale electrophysiology recordings in the macaque visual cortex. We tested the effect of object attention with stimuli of increasing complexity, from lines to natural objects against cluttered backgrounds. Consistent with behavioral observations, the iterative model correctly predicted the spread of attentional modulation in visual neurons for simple stimuli. However, for more complex stimuli containing recognizable objects, we observed asynchronous but not iterative modulation. Thus, we produced a set of hybrid stimuli, combining local elements of two different objects, that we alternated with the presentation of stimuli of intact objects. By doing so, we made local information unreliable, forcing the monkey to solve the task iteratively. Indeed, we observed that this set of stimuli induced iterative attentional modulations. These results provide the first systematic investigation on object attention in both artificial and cortical neurons.

TALK 4

TIME TO CONSIDER TIME: COMPARING HUMAN REACTION TIMES TO DYNAMICAL SIGNATURES FROM RECURRENT VISION MODELS ON A PERCEPTUAL GROUPING TASK

Alekh Karkada Ashok[1], Lore Goetschalckx[1], Lakshmi Narasimhan Govindarajan[1], Aarit Ahuja[1], David Sheinberg[1], Thomas Serre[1]; [1]Brown University

To make sense of its retinal inputs, our visual system organizes perceptual elements into coherent figural objects. This perceptual grouping process, like many aspects of visual cognition, is believed to be dynamic and at least partially reliant on feedback. Indeed, cognitive scientists have studied its time course through reaction time measurements (RT) and have associated it with a serial spread of object-based attention. Recent progress in biologically-inspired machine learning, has put forward convolutional recurrent neural networks (cRNNs) capable of exhibiting and mimicking visual cortical dynamics. To understand how the visual routines learned by cRNNs compare to humans, we need ways to extract meaningful dynamical signatures from a cRNN and study temporal human-model alignment. We introduce a framework to train, analyze, and interpret cRNN dynamics. Our framework triangulates insights from attractor-based dynamics and evidential learning theory. We derive a stimulus-dependent metric, ξ, and directly compare it to existing human RT data on the same task: a grouping task designed to study object-based attention. The results reveal a "filling-in" strategy learned by the cRNN, reminiscent of the serial spread of object-based attention in humans. We also observe a remarkable alignment between ξ and human RT patterns for diverse stimulus manipulations. This alignment emerged purely as a byproduct of the task constraints (no supervision on RT). Our framework paves the way for testing further hypotheses on the mechanisms supporting perceptual grouping and object-based attention, as well as for inter-model comparisons looking to improve the temporal alignment with humans on various other cognitive tasks.

TALK 5

THREE THEORIES OF OBJECT-BASED ATTENTION IMPLEMENTED IN DEEP NEURAL NETWORK MODELS

Hossein Adeli[1], Seoyoung Ahn[2], Gregory Zelinsky[2], Nikolaus Kriegeskorte[1]; [1]Columbia University, [2]Stony Brook University

Understanding the computational mechanisms that transform visual features into coherent object percepts requires the implementation of theories in scalable models. Here we report on implementations, using recent deep neural networks, of three previously proposed theories in which the binding of features is achieved (1) through convergence in a hierarchy of representations resulting in object-files, (2) through a reconstruction or a generative process that can target different features of an object, or (3) through the elevation of activation by spreading attention within an object via association fields. First, we present a model of object-based attention that relies on capsule networks to integrate features of different objects in the scene. With this grouping mechanism the model is able to learn to sequentially attend to objects to perform multi-object recognition and visual reasoning. The second modeling study shows how top-down reconstructions of object-centric representations in a sequential autoencoder can target different parts of the object in order to have a more robust and human-like object recognition system. The last study demonstrates how object perception and attention could be mediated by flexible object-based association fields at multiple levels of the visual processing hierarchy. Transformers provide a key relational and associative computation that may be present also in the primate brain, albeit implemented by a different mechanism. We observed that representations in transformer-based vision models can predict the reaction time behavior of people on an object grouping task. We also show that the feature maps can model the spreading of attention in an object.

TALK 6

COMBINING GENERATIVE ADVERSARIAL NETWORKS (GANS) WITH BEHAVIOR AND BRAIN RECORDINGS TO STUDY SCENE UNDERSTANDING

Melissa Le-Hoa Võ[1], Aylin Kallmayer[1]; [1]Goethe University Frankfurt

Our visual world is a complex conglomeration of objects that adhere to semantic and syntactic regularities, a.k.a. scene grammar according to which scenes can be decomposed into phrases – i.e, smaller clusters of objects forming conceptual units – which again contain so-called anchor objects. These usually large and stationary objects further anchor predictions regarding the identity and location of most other smaller objects within the same phrase and play a key role in guiding attention and boosting perception during real-world search. They therefore provide an important organizing principle for structuring real-world scenes. Generative adversarial networks (GANs) trained on images of real-world scenes learn the scenes' latent grammar to then synthesize images that mimic images of real-world scenes increasingly well. Therefore GANs can be used to study the hidden representations underlying object-based perception serving as testbeds to investigate the role that anchor objects play in both the generation and understanding of scenes. We will present some recent work in which we presented participants with real and generated images recording both behavior and brain responses. Modelling behavioral responses from a range of computer vision models we found that mostly high-level visual features and the strength of anchor information predicted human scene understanding of generated scenes. Using EEG to investigate the temporal dynamics of these processes revealed initial processing of anchor information which generalized to subsequent processing of the scene's authenticity. These new findings imply that anchors pave the way to scene understanding and that models predicting real-world attention and perception should become more object-centric.

Talk Sessions

Talk Session: Saturday, May 18, 2024, 8:15 – 9:45 am, Talk Room 1

Perceptual Organization

Moderator: Gideon Caplovitz, University of Nevada

Talk 1, 8:15 am, 21.11

DOES PERSPECTIVE-DISTORTION MODULATE THE TEMPORAL TUNING OF SYMMETRY RESPONSES?

Nikan Movahedi[1,2] (movahedinikan@gmail.com), Shenoa Ragavaloo1,2[1,2], Peter J. Kohler[1,2]; [1]Department of Psychology, York University, Toronto, ON, [2]Center for Vision Research, York University, Toronto, ON

Symmetry is a highly salient feature in both natural and man-made environments. Numerous species are sensitive to symmetry, and symmetry is thought to be an important cue for visual tasks, including viewpoint-invariant representation of objects, detection of regularity and structure, and mate selection. However, although symmetries are common in natural and artificial objects and scenes, they are subject to perspective-distortion and thus rarely give rise to symmetrical patterns on the retina during natural vision. Here, we build on previous studies showing that perspective-distortion makes symmetry responses weaker and more task-dependent (Makin et al., 2014; Keefe et al., 2018) by investigating the effect of perspective-distortion on the temporal tuning of symmetry responses. We used novel, naturalistic 3D objects that had reflection symmetry over a vertical axis. The objects were procedurally generated along with well-matched control objects without any symmetries and then rendered to produce images in which object symmetries are either present in the image-plane or perspective-distorted. We measured visual system responses to image-plane and perspective-distorted symmetry using high-density EEG with a Steady-State Visual Evoked Potentials (SSVEPs) paradigm in which images of symmetrical objects alternate with images of control objects. This makes it possible to isolate symmetry-specific brain activity in the odd harmonics of the stimulation frequency. To investigate the temporal tuning of these responses, we used seven different stimulation frequencies in different conditions, between 1 and 10 Hz. We collected data from 30 participants with normal or corrected-to-normal visual acuity. We found that for both image-plane and perspective-distorted symmetry, responses peak at 2 Hz and are much reduced at higher frequencies across electrodes over occipital and temporal cortex. Response amplitudes were generally higher for image-plane symmetry, but surprisingly, the spatial tuning was not strongly modulated by perspective-distortion. Further investigations will determine how distinct visual regions may contribute to these results.

This work was supported by the Vision Science to Applications (VISTA) program funded by the Canada First Research Excellence Fund (CFREF, 2016–2023) and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada awarded to PJK.

Talk 2, 8:30 am, 21.12

MAGNITUDE OF MIDDLE TEMPORAL N300 REFLECTS CONTOUR FIDELITY IN CONTOUR INTEGRATION

Dongcheng He[1] (hedch1995@berkeley.edu), Angeline Yang[1], Daniel R. Coates[2], Haluk Ogmen[3], Susana Chung[1]; [1]University of California Berkeley, [2]University of Houston, [3]University of Denver

Contour integration (CI) reflects the ability of the visual system to bind individual elements into a global coherent shape. Previous neuroimaging studies of CI compared behavioral and neural data using stimuli with and without contours. However, these studies had multiple potential confounding variables, including the degree of visual awareness, temporal prediction, and task-relevance. To address these, we presented an array of 27x15 line-segments (each 2x0.5°) that changed their orientations independently and randomly at a rate of 15 Hz. Each trial comprised 25 frames (1.67 s). During one of these frames, the orientations of 12 line-segments were aligned to form a contour (outline of a box with 3 segments on each side), at 6° to the right or left of fixation. Subjects (n=10) fixated on a central dot throughout the trial and indicated whether the contour was on the right or left. EEG signals were collected and analyzed along with task performance. In Experiment 1, we varied the onset timing of the contour to test the effects of temporal prediction. Results showed that ERPs including Frontal P200, Occipital N200 and P400, Parietal P400 and contralateral N200, as well as Middle Temporal (MT) N300 were synchronized with the onset timing of the contour. In Experiment 2, we fixed the onset timing of the contour but varied its fidelity by adding various levels of random orientation jitter to the line segments that formed the contour. Results showed that among all the ERPs observed in Experiment 1, only the magnitudes of MT N300 and Parietal P400 were dependent on contour fidelity. As reported in the literature, Frontal P200 and Posterior P400 are likely to reflect awareness/attention and task-related-efforts, respectively, and Posterior N200 is likely to be a correlate of visual phenomenal consciousness. We therefore propose MT N300 as a neural correlate of contour integration.

NIH Grant EY030253

Talk 3, 8:45 am, 21.13

PERCEPTUAL GROUPING WITH LATENT NOISE

Ben Lonnqvist[1] (ben.lonnqvist@epfl.ch), Zhengqing Wu[1], Michael H. Herzog[1]; [1]EPFL

Humans effortlessly group elements into objects and segment them from the background and other objects without supervision. For example, the black and white stripes of a zebra are grouped together despite vastly different colors. A thorough theoretical and empirical account of perceptual grouping is still missing – Deep Neural Networks (DNNs), which are considered leading models of the visual system still regularly fail at simplistic perceptual grouping tasks. Here, we propose a counterintuitive unsupervised computational approach to perceptual grouping and segmentation: that they arise because of neural noise, rather than in spite of it. We show that adding noise in a DNN enables the network to separate objects and to segment images even though it was never trained on any segmentation labels. To test whether the models exhibit perceptual grouping, we introduce the Good Gestalt (GG) datasets – six datasets based on a century of Gestalt principles specifically designed to test perceptual grouping. These include illusory contours, closure, continuity, proximity, and occlusion. Our DNN using neural noise finds the correct perceptual groups while other control models, including state-of-the-art segmentation models, fail at these critical tests. We further show that our model performs well with remarkably low levels of noise, and requires only few successive time steps to compute. Using simplifying but realistic assumptions from optics, we are also able to mathematically link our model's perceptual grouping performance to image statistics. Together, our results suggest a novel unsupervised segmentation method requiring few assumptions, a new explanation for the formation of perceptual grouping, and a novel benefit of neural noise.

BL was supported by the Swiss National Science Foundation grant n. 176153 "Basics of visual processing : from elements to figures".

Talk 4, 9:00 am, 21.14

FORGING A HEAD: HOW INTERNAL AXES AND EXTERNAL VISUAL ELEMENTS DETERMINE A SHAPE'S PERCEIVED FACING DIRECTION

Jiangxue Valentina Ning[1] (ningj@newschool.edu), Benjamin van Buren[1]; [1]The New School

Human perceivers are very sensitive to which way others are facing, with head and gaze cues capturing attention (when directed at us), orienting attention (when directed elsewhere), and even influencing downstream judgments about others' social traits. But what causes us to see a shape as directed in the first place? Does the perception of a shape's facing direction depend mainly on its intrinsic structure — or might it also be influenced by spatial context? In Experiment 1, observers briefly viewed a randomly oriented oval, and afterward used a circular slider to report which way they saw it facing. A dot was always drawn near the oval — aligned with either its long or short symmetry axis. Observers were biased to see the oval as facing toward the dot, but this effect was much stronger when the dot was aligned with the oval's long (vs. short) symmetry axis, indicating that external elements interact with a shape's internal structure to determine its perceived facing direction. How automatic is this association between long-axis alignment and 'towardness'? In Experiment 2, participants saw the same displays, but now made speeded keypresses to indicate whether the oval's long or short axis was aligned with the dot. In one block of trials, they pressed an anterior (further forward) key to report long-axis alignment, and a posterior (further back) key to report short-axis alignment. In another block, they responded with opposite key-mappings. Participants responded faster in the block where an anterior key was paired with long-axis alignment and a posterior key with short-axis alignment, suggesting an automatic bias to see long-axis alignment as facing towards. We conclude that the perception of facing direction is driven by the interaction of internal structure and external context, in a way which indicates the particular salience of the long symmetry axis.

Talk 5, 9:15 am, 21.15

PHENOMENOLOGICAL CONTRACTION DOES NOT DEPEND ON EXPLICIT CUES TO OCCLUSION

Tess White[1] (tessw@unr.edu), Andrew Piotrowski[1], Jacob Crull[1], Chloe Leroy[1], Elka Piotter[1], Gideon Caplovitz[1]; [1]University of Nevada Reno

Just as our visual system enables us to experience the things we are looking at, so too does it allow us to experience aspects of those things that are not visible, as is the case when an object is partially occluded from view. Our research focuses on understanding how this 'seeing what is not there' is accomplished. While amodal completion is a well-studied process that allows us to experience partially occluded objects as complete, a lesser known phenomenon is that amodally completed parts of objects tend to appear smaller than identically sized counterparts that are fully visible. This phenomenological contraction was first described by Kanizsa and has not been studied nearly as much as the mechanisms behind amodal completion itself. We developed a paradigm using a stimulus composed of two triangles, one partially occluding the other that allowed us to quantify phenomenological contraction by measuring the mislocalization of the occluded triangle's vertex. In our previous work we found the contraction is independently influenced by the size of the occluded object, corner angle, and the level of occlusion. Our current research seeks to determine whether such mislocalization/contraction is dependent on the presence of explicit occlusion. Across three experiments our approach was to replicate our previous experiments while presenting only partial contours of a single triangle without an explicit occluder. We again found independent influences of corner angle and level of occlusion (how much contour was visible) even when the partial contour consisted of only two line segments (lacking the base of the triangle). Based on our data we conclude that the phenomenological contraction arises from mechanisms of interpolation and/or extrapolation that are largely independent of explicit cues to occlusion and as such be considered as separate from amodal completion.

Talk 6, 9:30 am, 21.16

INVESTIGATING THE BREADTH AND STRENGTH OF PERCEPTUAL CONTROL OF ILLUSORY APPARENT MOTION

Allison Allen[1] (alkallen@ucsc.edu), Matthew Jacobs[2], Nicolas Davidenko[1]; [1]University of California, Santa Cruz, [2]Queen's University

Recently, a stimulus called Illusory Apparent Motion (IAM) was discovered by Davidenko et al. (2017) wherein pixel textures randomly refreshing at a rate of 1.5 Hz generate the appearance of coherent apparent motion. IAM is a maximally ambiguous multistable stimulus that observers may perceive as moving coherently in a countless number of patterns (e.g., translation, shear, rotation, expansion-contraction). The current set of studies explores observers' ability to perceptually control the appearance of IAM. The first two experiments used paradigms similar to those used with other multistable stimuli. Experiment 1 (n = 99) used a motion-priming persistence task, based on the methods of Davidenko et al. (2017), while experiment 2 (n = 76) used a dynamic report task with no priming, based on the methods of Kohler et al. (2008). In both experiments, participants successfully controlled translational motion by 'changing' or 'holding' their percepts, indicating that observers are capable of perceptually controlling IAM, similar to other multistable stimuli. Having established this, Experiment 3 (n = 43) explored the breadth of participants' ability to perceive and control motion in IAM by testing them on 14 types of translational, shear, rotating, and expanding-contracting motion patterns. Participants were able to perceive a wide variety of motion patterns but were limited in the motion patterns they could control. Finally, Experiment 4 (n = 82) aimed to quantify the influence of perceptual control in biasing perceptions of IAM by presenting participants with a motion nulling signal (at above and below each participant's perceptual threshold) while they attempted to control the motion. We were successful in quantifying the strength of perceptual control of IAM relative to low-level motion signals. Collectively, these studies provide evidence for the breadth and strength of observers' ability to perceptually control IAM.

Talk Session: Saturday, May 18, 2024, 8:15 – 9:45 am, Talk Room 2

Object Recognition: Models

Moderator: Leila Wehbe, Carnegie Mellon University

Talk 1, 8:15 am, 21.21

BUILDING BETTER MODELS OF BIOLOGICAL VISION BY SEARCHING FOR MORE ECOLOGICAL DATA DIETS AND LEARNING OBJECTIVES

Drew Linsley[1] (drew_linsley@brown.edu), Akash Nagaraj[1], Alekh Ashok[1], Francis Lewis[1], Peisen Zhou[1], Thomas Serre[1]; [1]Brown University

The many successes of deep neural networks (DNNs) over the past decade have been driven by data and computational scale rather than biological insights. However, as DNNs have continued to improve on benchmarks like ImageNet, they have worsened as models of biological brains and behavior. For instance, recent DNNs with human-level object classification accuracy are no better at predicting human perception or image-evoked responses in primate inferotemporal (IT) cortex than DNNs from a decade ago (e.g., Linsley et al., 2023). Here, we build better DNN models of biological vision by finding data diets and objective functions that more closely resemble those that shape biological brains. We began by building a platform for searching through naturalistic data diets and objective functions for training a standardized DNN architecture at scale. Each DNN's data diet was sampled from our rendering engine, which generates life-like videos of objects in real-world scenes. In parallel, each model's objective function was sampled from a parametrized space of image reconstruction objectives, which made it possible to train models to learn combinations of causal and acausal recognition strategies over space or space and time. We evaluated the ability of hundreds of DNNs trained on this platform to predict human performance on a novel "Greebles" object recognition task (Ashworth et al., 2008). We found that DNNs trained to capture the causal structure of data were significantly more predictive of human decisions and reaction times than any other DNN tested. Moreover, these causal DNNs learned strong equivariance to out-of-plane variations in pose, recapitulating classical theory on the foundations of object constancy (Sinha & Poggio,1996) despite no explicit constraints to do so. Our work identifies key limitations in how DNNs are trained today and introduces a better approach for building DNN-based models of human vision that can ultimately advance perceptual science.

This work was supported by ONR (N00014-19-1-2029), NSF (IIS-1912280 and EAR-1925481), DARPA (D19AC00015), and NIH/NINDS (R21 NS 112743), Cloud TPU hardware through the TensorFlow Research Cloud (TFRC) program as well as computing hardware supported by NIH Office of the Director grant S10OD025181.

Talk 2, 8:30 am, 21.22

GENERATING OBJECTS IN PERIPHERAL VISION USING ATTENTION-GUIDED DIFFUSION MODELS

Ritik Raina[1] (ritik.raina@stonybrook.edu), Seoyoung Ahn[1], Gregory Zelinsky[1]; [1]Stony Brook University

Despite the majority of our visual field being blurry in the periphery, with only the central ~2 degrees offering high-resolution inputs, we have no difficulty perceiving and interacting with objects around us. We hypothesize that the human perception of a stable visual world is mediated by an active generation of objects from blurred peripheral vision. Furthermore, we hypothesize that this active peripheral generation is task-dependent, guided by information extracted from fixations, with the goal of constructing a relevant object and scene context for the current task. We test these hypotheses by using latent diffusion models and evaluating the influence of fixated image information on generating objects in the blurred periphery. We ask this question in the context of an object referral task, in which participants hear a spoken description of the search target (e.g., "right white van"). We recorded eye movements from participants (n=220) as they viewed 1,619 images and attempted to localize the referred targets. The model received high-resolution input only from fixated regions, mimicking foveated vision, and generated high-resolution objects in the originally blurred peripheral areas. We found that using foveated-image inputs corresponding to observed behavioral fixations led to the model generating target objects in the periphery with greater fidelity compared to randomly located fixations, as measured by squared pixel difference (Human Fixation SSE = 178.27; Random Fixation SSE = 212.42; averaged over the first 20 fixations). This fixation-driven advantage specifically applied to the reconstruction of task-relevant objects, such as objects of the same referred category, and did not extend to non-targets or background elements. Our findings support the idea that human perception actively generates relevant objects in the blurry periphery as a means of building a stable object context, which is guided by goal-directed attention control mechanisms.

This work was supported in part by NSF IIS awards 1763981 and 2123920 to G.Z.

TALK 3, 8:45 AM, 21.23

LEARNING TO DISCRIMINATE BY LEARNING TO GENERATE: ZERO-SHOT GENERATIVE MODELS INCREASE HUMAN OBJECT RECOGNITION ALIGNMENT

Robert Geirhos[1], Kevin Clark[1], Priyank Jaini[1]; [1]Google DeepMind

How does the human visual system recognize objects---through discriminative inference (fast but potentially unreliable) or using a generative model of the world (slow but potentially more robust)? The question of how the brain combines the best of both worlds to achieve fast and robust inference has been termed "the deep mystery of vision" (Kriegeskorte 2015). Yet most of today's leading computational models of human vision are simply based on discriminative inference, such as convolutional neural networks or vision transformers trained on object recognition. In contrast, we here revisit the concept of vision as generative inference. This idea dates back to the notion of vision as unconscious inference proposed by Helmholtz (1867), who hypothesized that the brain uses a generative model of the world to infer probable causes of sensory input. In order to build a generative model capable of recognizing objects, we take some of the world's most powerful generative text-to-image models (Stable Diffusion, Imagen and Parti) and turn them into zero-shot image classifiers using Bayesian inference. We then compare those generative classifiers against a broad range of discriminative classifiers and against human psychophysical object recognition data from the "model-vs-human" toolbox (Geirhos et al. 2021). We discover four emergent properties of generative classifiers: They show a record-breaking human-like shape bias (99% for Imagen), near human-level accuracy on challenging distorted images, and state-of-the-art alignment with human classification errors. Last but not least, generative classifiers understand certain perceptual illusions such as the famous bistable rabbit-duck illusion or Giuseppe Arcimboldo's portrait of a man's face composed entirely of vegetables, speaking to their ability to discern ambiguous input and distinguish local from global information. Taken together, our results indicate that while the current dominant paradigm for modeling human object recognition is discriminative inference, zero-shot generative models approximate human object recognition data remarkably well.

TALK 4, 9:00 AM, 21.24

OUT-OF-DISTRIBUTION GENERALIZATION BEHAVIOR OF DNN-BASED ENCODING MODELS FOR THE VISUAL CORTEX

Spandan Madan[1,3], Mingran Cao[2], Will Xiao[1], Hanspeter Pfister[1], Gabriel Kreiman[1,3]; [1]Harvard University, [2]The Francis Crick Institute, [3]Boston Children's Hospital

Deep Neural Networks (DNNs) trained for object classification have remarkably similar internal feature representations to neural representations in the primate ventral visual stream. This has led to the widespread use of encoding models of the visual cortex utilizing linear combinations of pre-trained DNN unit activities. However, DNNs struggle with generalization under distribution shifts, particularly when faced with out-of-distribution (OOD) samples. While DNNs excel at interpolating between training data points, they perform poorly when extrapolating beyond the bounds of the training data (e.g., Hasson et al., 2020). We characterized the generalization capabilities of DNN-based encoding models when predicting neuronal responses from the primate visual ventral stream. Using a large-scale dataset of neuronal responses from the macaque inferior temporal cortex to over 100,000 images, we simulated the effect of OOD neural activity prediction by dividing the images into multiple training and test sets, holding out subsets of the data to introduce different OOD domain shifts. This includes OOD low-level image features like contrast, hue, and size; OOD high-level features like animate vs inanimate, food vs non-food, different semantic object categories; and OOD K-means clusters in the distributed representations of ResNet features and neural data. For each feature, an OOD test set was constructed by defining a parametric value for that feature, and withholding from training a subset of possible values for testing. Overall, models performed much worse when predicting out-of-distribution image responses compared to standard cross-validation. Prediction on an IID test set with no distribution shift had an r^2 = 0.5, while OOD prediction ranges from 0.48 (images with OOD contrast shift) to as low as 0.1 (images with OOD hue). This indicates a deep problem in modern models of the visual cortex—the promise of current image-computable models remains limited to the training image distribution.

This work has been partially supported by NSF grant IIS-1901030.

TALK 5, 9:15 AM, 21.25

HIGHER VISUAL AREAS ACT LIKE DOMAIN-GENERAL FILTERS WITH STRONG SELECTIVITY AND FUNCTIONAL SPECIALIZATION

Meenakshi Khosla[1] (meenakshik1993@gmail.com), Leila Wehbe[2]; [1]University of California, San Diego, [2]Carnegie Mellon University

Modeling neural responses to naturalistic stimuli has been instrumental in advancing our understanding of the visual system. Dominant computational modeling efforts have been deeply rooted in preconceived hypotheses. Here, we develop a hypothesis-neutral computational methodology which brings neuroscience data directly to bear on the model development process. We demonstrate the effectiveness of this technique in modeling as well as systematically characterizing voxel tuning properties. We leverage the unprecedented scale of the Natural Scenes Dataset to constrain parametrized neural models of higher-order visual systems with brain response measurements and achieve novel predictive precision, outperforming the predictive success of state-of-the-art models. Next, we ask what kinds of functional properties emerge spontaneously in these response-optimized models? We examine trained networks through structural and functional analysis by running `virtual' fMRI experiments on large-scale probe datasets. Strikingly, despite no category-level supervision, since the models are optimized for brain response prediction from scratch, the units in the networks after optimization act strongly as detectors for semantic concepts like `faces' or `words', thereby providing one of the strongest evidences for categorical selectivity in these areas. Importantly, this selectivity is maintained when training the networks without images that contain the

preferred category, strongly suggesting that selectivity is not domain-specific machinery, but sensitivity to generic patterns that characterize preferred categories. Beyond characterizing tuning properties, we study the transferability of representations in response-optimized networks on different perceptual tasks. We find that the sole objective of reproducing neural targets, without any task-specific supervision, grants different networks intriguing functionalities. Finally, our models show selectivity only for a limited number of categories, all previously identified, suggesting that the essential categories are already known. Together, this new class of response-optimized models combined with novel interpretability techniques reveal themselves as a powerful framework for probing the nature of representations and computations in the brain.

Talk 6, 9:30 AM, 21.26

EMERGENCE OF ILLUSORY CONTOURS IN ROBUST DEEP NEURAL NETWORKS BY ACCUMULATION OF IMPLICIT PRIORS

Tahereh Toosi[1] (tahereh.toosi@columbia.edu), Kenneth Miller[1]; [1]Columbia University

Deep neural networks (DNNs), trained for object recognition, exhibit similarities to neural responses in the monkey visual cortex and are currently considered the best models of the primate visual system. It remains unclear whether psychophysical effects, such as illusory contours perceived by humans, also emerge in these models. Utilizing the invertibility properties of robustly trained feedforward neural networks, we demonstrated that illusory contours and shapes emerge when the network integrates its learned implicit priors. Our visual system is believed to store perceptual priors, with visual information learned and embedded in neural connections across all visual areas. This stored information is harnessed when required, for instance, during occlusion resolution or visual imagination generation. While the significance of feedback connections in these processes is well recognized, the precise neural mechanism that aggregates dispersed information throughout the visual cortex remains elusive. In this study, we leverage a ResNet50 neural network, conventionally used in image recognition, to shed light on the neural basis of illusory contour perception through its inherent feedback mechanism during error backpropagation. By iteratively accumulating the gradients of the loss with respect to an input—a Kanizsa Square—within an adversarially trained network, we observed the emergence of edge-like patterns in the area of the perceived 'white square'. This process, which unfolds over multiple iterations, echoes the time-dependent emergence of illusory contours in the visual cortices of rodents and primates as seen in experimental studies. Notably, the ResNet50 employed in this study was neither specifically enhanced with feedback capabilities nor optimized to detect or decode these illusory contours; it was merely trained for robust object recognition against adversarial examples. These findings highlight a compelling parallel, suggesting that the ability to perceive illusory contours might be an incidental consequence of the network's ability to handle adversarial noise during its training regime.

T.T. is supported by NIH 1K99EY035357-01. This work was also supported by NIH RF1DA056397, NSF 1707398, and Gatsby Charitable Foundation GAT3708.

Talk Session: Saturday, May 18, 2024, 10:45 AM – 12:30 PM, Talk Room 1

Eye Movements: Neural mechanisms

Moderator: J. Patrick Mayo, University of Pittsburgh

Talk 1, 10:45 AM, 22.11

A TRANSIENT SIGNAL IN FOVEAL SUPERIOR COLLICULUS NEURONS FOR JUMPSTARTING PERIPHERAL SACCADIC ORIENTING

Tong Zhang[1], Ziad Hafed; [1]University of Tuebingen

The superior colliculus (SC) is critical for saccade generation. Recent work has shown that, despite bursting at times other than saccades, SC population activity at the time of saccade motor bursts is more temporally aligned than for visual bursts (Jagadisan & Gandhi, 2022). Similarly, population activity in motor bursts resides in different subspaces to visual bursts (Baumann et al., 2023), and even the sensory signal embedded in SC motor bursts (Baumann et al., 2023) is transformed relative to visual bursts, such that the same individual neurons "prefer" different visual features in the two bursting epochs. However, how might such a transformation from a visual regime to a motor regime be realized? Here we first show that when a planned saccade is finally released with a go signal (removal of a fixation spot), peripheral SC neurons (representing the saccade target location) exhibit a robust, short-latency pause in spiking, before the motor bursts eventually erupt. This pause starts within ~50 ms from the go signal, and it is stimulus-dependent (e.g. having a stronger firing rate dip for a salient peripheral stimulus). Additionally, this pause still occurs, to a weaker extent, with saccades to a small spot or blank. When we then recorded from foveal SC neurons in similar tasks, we found that these neurons actually burst after the go signal, rather than paused. Remarkably, these foveal bursts occurred (and peaked) several milliseconds earlier than the pauses in the peripheral neurons, and they were not explained by offset responses to the removal of the fixation spot. Foveal bursts also occurred when releasing memory-guided saccades (with no peripheral visual targets), and they were not sensitive to peripheral target appearance. Thus, we found a transient foveal SC signal jumpstarting peripheral saccadic orienting, likely facilitating a necessary representational transformation needed for saccade motor bursts to occur.

Talk 2, 11:00 AM, 22.12

MIXED SELECTIVITY FOR TARGET SELECTION BIASES IN THE SUPERIOR COLLICULUS

Abe Leite[1] (abrahamjleite@gmail.com), Hossein Adeli[3], Robert M. McPeek[2], Gregory J. Zelinsky[1]; [1]Stony Brook University, [2]SUNY College of Optometry, [3]Columbia University

How does the brain flexibly integrate the multiple sources of information needed to control arbitrary goal-directed behavior? Mixed selectivity theory argues that this cognitive flexibility is achieved through flexible neural representations, with most neurons encoding nonlinear (and in some articulations dynamic) combinations of the stimulus factors. In this view, only fundamental computations underlying many behaviors merit neurons dedicated specifically to them. Despite its importance, the question of how mixed representations shape behavior in an attention-demanding task remains open. Our study applies mixed selectivity theory to visual attention by analyzing three factors known to bias saccade target selection during search: bottom-up feature contrast, top-down target guidance, and the history of previous object fixation (inhibitory tagging). We analyzed how single neuron responses in the rhesus superior colliculus encode these three attention-guiding properties of an object landing in the response field during eye movements in visual search, then determined mixed selectivity using two methods: standard nested GLM and our extension of an application of partial information decomposition (PID) to this behavior. We found that (1) Our application of PID, in contrast to standard GLM analyses, captures the dynamics of neural selectivity over time and the subtleties of how a neuron mixes multiple variables. (2) There is ample evidence for cells that sustain their encoding of multiple factors, and also cells whose selectivity varies over the time course of target selection. (3) In addition to these mixed selectivity neurons, a substantial group of neurons is uniquely selective to whether stimuli were previously fixated while searching, suggesting that inhibitory tagging may be a fundamental computation supporting overt visual attention. We conclude that both static and dynamic forms of mixed selectivity are used to represent attention biases in the superior colliculus, and that the colliculus may participate in a neural circuit dedicated to inhibitory tagging.

This project is based upon work supported by the National Institutes of Health under Grant No. 5R01EY030669-05 and work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2234683.

Talk 3, 11:15 am, 22.13

NEURONAL POPULATION ESTIMATES OF SPATIAL ATTENTION ARE ROBUST TO THE PRESENCE OF MICROSACCADES

Shawn Willett[1,2] (smw146@pitt.edu), Patrick Mayo[1,2]; [1]University of Pittsburgh Department of Ophthalmology, [2]Center for the Neural Basis of Cognition

Neurons in visual area V4 exhibit attention-related changes in firing rate. Recent work (Lowet et. al., 2018) proposed that attentional modulation of V4 activity occurred only after a microsaccade towards the attended location, suggesting that microsaccades gate attention-related effects. However, other work (Gongchen, et. al., 2022) reported that attentional modulation of neuronal activity in the superior colliculus (SC) occurred before and after microsaccades, and in the absence of microsaccades. Thus, microsaccades may not contribute to attention-related effects in the SC. To determine if these contrasting findings emerge because of differences in brain structure or task demands, we investigated population measures of attention in V4 aligned to microsaccade onset while monkeys performed a visual-spatial attention task (Mayo and Maunsell, 2016), similar to the task used in prior SC work. Monkeys detected an orientation change in one of two simultaneously presented oriented Gabors. We cued attention to one target using an 80% valid visual cue on instruction trials that occurred prior to each block of test trials. During each trial, monkeys fixated until they reported a change in orientation by saccading to the changed target. We recorded over 3500 V4 units from two bilaterally implanted Utah arrays across 54 sessions. We used de-mixed principal component analysis (dPCA; Kobak et. al, 2016) to extract an attention-related latent axis from our high-dimensional neuronal activity. We projected our microsaccade aligned neural population activity onto this attention-related axis and found that attention-related population activity was flat aligned to microsaccade onset, suggesting that attention modulates V4 activity regardless of microsaccades. Trials in which microsaccades occurred appeared identical to the unchanging activity observed in trials without microsaccades. Our results indicate that the modulation of V4 neural activity by attention and microsaccades is largely separable, and that attention modulates V4 activity regardless of the occurrence of a microsaccade.

Talk 4, 11:30 am, 22.14

PREFRONTAL NEURAL ACTIVITY PREDICTS AND MITIGATES SPATIAL UNCERTAINTY IN A GAZE TASK

Vishal Bharmauria[1] (bhav2501@yorku.ca), Adrian Schütz[2], Xiaogang Yan[1], Hongying Wang[1], Frank Bremmer[2], John Douglas Crawford[1]; [1]Center for Vision Research & Vision: Science to Applications (VISTA), York University, [2]Department of Neurophysics, Phillips Universität Marburg and Center for Mind, Brain and Behavior – CMBB, Philipps-Universität Marburg and Justus-Liebig-Universität Giessen

To predict the future, the brain must integrate past with current sensory information. Research has suggested that prefrontal cortex predicts the timing of events (Fu et al., 2023). Here, we investigated if it also predicts spatial uncertainty. To do this, we recorded neural activity in the frontal (FEF) and supplementary (SEF) eye fields of two rhesus macaques, trained to saccade toward remembered visual targets (T) in presence of a landmark (L) that was surreptitiously shifted to a new position (L') by a fixed amplitude in one of eight randomized directions arranged circularly around L. Previously, we showed that this results in retrospective shifts in FEF/SEF memory and gaze signals (Bharmauria et al., 2020, 2021). Here, we examined the period from the initial visual response to 300 ms after the landmark shift in 147/68 spatially tuned FEF/SEF neurons for prospective coding of this shift. We used a model-fitting technique to test memory delay coding along a T-T' continuum. Remarkably, just before the landmark shift, SEF coded a shift toward T'. Since this direction was randomized, we hypothesized that SEF might be 'guessing' the direction of the shift. We tested this using a 2D analysis with the real shift (T') rotated to the right and seven other imaginary shifts circularly arranged. Shortly after the visual response, response fields developed a donut-like prediction in all directions. This did not occur in shuffled controls and could not be accounted for by attraction toward landmark position (TL) or gaze error (TG). Eventually, after the real shift, this predictive 'donut' code shifted toward the actual L'. These data suggest that after thousands of training trials, the monkey brain, specifically SEF, created a guessing strategy based on learned probabilities and anticipation. This might allow the brain to optimize behavior and mitigate spatial uncertainty in the surrounding world.

Canadian Institutes for Health Research (CIHR); Vision: Science to Applications (VISTA) Program: Deutsche Forschungsgemeinschaft

(DFG)

TALK 5, 11:45 AM, 22.15

BEHAVIORAL AND NEURAL CORRELATES OF IMPAIRED SCENE PERCEPTION FOLLOWING SACCADIC EYE MOVEMENTS

Yong Min Choi[1] (choi.1696@osu.edu), Tzu-Yao Chiu[1], Julie D. Golomb[1]; [1]Department of Psychology, The Ohio State University

The visual input projected to the retina shifts drastically across saccadic eye movements. Although we are not aware of it, perception of simple visual stimuli presented around the time of a saccade is impaired (Burr et al., 1994). Meanwhile, the extent to which the post-saccadic impairment influences high-level visual scene perception remains unclear. We conducted behavioral and fMRI experiments examining processing for scene images containing different spatial frequency content presented at different delays following a saccade. First, subjects performed a 6-way scene categorization task (beach, mountain, etc.) on images presented 5, 16, 50, 158, or 500 ms after saccade completion. We found lower scene categorization accuracy at 5 ms and 16 ms post-saccadic delays compared to longer delays, for both low- and high-spatial frequency filtered images, suggesting broadly impaired scene perception lasting less than 50 ms after a saccade offset. To further investigate what visual information is impaired, and in the absence of an explicit categorization task, we conducted an fMRI experiment where subjects performed a 1-back task on scene images while making saccades. Short and long post-saccadic delay trials were sorted post-hoc using eye-tracking data. Using RSA-based decoding analysis, we assessed scene category information (urban vs nature) in scene-selective brain areas, and low-level visual information (high vs low spatial frequency) in the early visual cortex. We found decreased scene category information in the posterior parahippocampal place area on short versus long post-saccadic delay trials, consistent with the behavioral impairment. Interestingly, lower-level visual information of a scene image was less impaired; spatial frequency information in the early visual cortex was not significantly different between short and long post-saccadic delay trials. Taken together, the current study presents novel evidence for impaired processing of complex scenes following saccades that may be driven by selectively interrupted neural representations of high-level scene content.

NIH R01-EY025648 (JG)

TALK 6, 12:00 PM, 22.16

STEERING, OPTIC FLOW, AND COMPENSATORY EYE MOVEMENTS IN CORTICALLY BLIND DRIVERS

Arianna P. Giguere[1] (apg7742@rit.edu), Matthew R. Cavanaugh[2,3], Brett R. Fajen[4], Duje Tadin[2], Krystel R. Huxlin[2,3], Gabriel J. Diaz[1,2]; [1]Rochester Institute of Technology Center for Imaging Science, [2]University of Rochester Center for Visual Science, [3]Flaum Eye Institute, University of Rochester Medical Center, [4]Rensselaer Polytechnic Institute Department of Cognitive Science

It is well known that the control of steering (e.g., when driving) is reliant on visual information from optic flow (Kountouriotis et al. 2016). Because optic flow is spatially correlated and accurate heading judgments can be made using a sparse and partial flow field (Warren and Kurtz 1992), it is surprising that drivers with cortical blindness (CB) across ¼ to ½ of their visual field demonstrate more variable lane positioning than their visually intact counterparts (Bowers et al. 2010). We hypothesized that this deficit arises because residual noise introduced in the "blind" field affects optic flow processing in service of steering. To test this hypothesis, we analyzed steering behavior in 10 CB drivers and 5 visually-intact controls immersed in a virtual reality steering task. Participants were asked to maintain a center-lane position while traveling at 19 m/s on a procedurally generated one lane road. Turn direction (left/right) and turn radius (35, 55, or 75 m) were manipulated. Additionally, optic flow density was indirectly manipulated through variation in environmental texture density (low, medium, high). Analysis of the average distance from the inner road edge revealed that all CB drivers were biased away from their blind side, but only controls and those with right-sided deficits decreased their distance to the inner road edge on medium and high optic flow density trials. The difference between these groups and the steering behavior of left-sided CBs, who showed no impact of optic flow, could not be attributed to age differences, time since stroke, or sparing in the central 10° of the visual field. Our results suggest that left-sided CBs place less weight on optic flow than right-sided CBs and controls. Preliminary analysis of gaze data suggests the insensitivity to variations in optic flow might also be attributed to compensatory gaze behavior.

NIH 1R15EY031090 and Research to Prevent Blindness' Low Vision Research Award

TALK 7, 12:15 PM, 22.17

A RARE CASE OF BILATERAL DAMAGE TO CORTICAL MOTION PROCESSING AREAS 40 YEARS AFTER PATIENT L.M.

Miriam Spering[1] (mspering@mail.ubc.ca), Philipp Kreyenmeier[1], Juana Ayala Castañeda[1], Jason Barton[1]; [1]University of British Columbia

In 1983, Zihl and colleagues reported the case of patient L.M., who had suffered bilateral damage to the lateral temporal-occipital cortex and showed a "disturbance of movement vision in a rather pure form" [Zihl, von Cramon, & Mai, Brain 1983; p.313], manifesting in selective deficits in motion perception, smooth pursuit and manual tracking of moving targets, particularly at higher speeds. These findings suggest that a human homologue of the middle temporal area (MT, or area V5) was located in this region. Here we present 19-year-old female patient C.C., who suffered encephalitis at age 3, with recent MRI showing bilateral damage to lateral occipitotemporal and medial occipitoparietal cortex. Similar to L.M., patient C.C. reports feeling overwhelmed in crowded areas, struggling with ball sports, and inaccuracy with fine motor tasks that involve moving objects. We tested C.C.'s smooth pursuit eye movements to visible and occluded targets and her ability to track and rapidly intercept objects that moved unpredictably. Compared to healthy young adults, C.C.'s smooth pursuit had a reduced velocity gain (.75) and was frequently interrupted by catch-up saccades, even in response to slow (10°/s) targets. When the target was temporarily occluded (ramp-occlusion-ramp for 800 ms each), pursuit dropped to zero velocity during occlusion and did not predictively accelerate before target reappearance. These motion prediction deficits extended to the patient's performance in naturalistic interception tasks. Whereas pointing accuracy was high (interception error 0.8°) for objects moving along simple, horizontal trajectories, performance degraded significantly for complex flyball (M = 2.4°) and occluded trajectories (M = 3.6°) with almost no ability to discriminate different trajectory types. These findings provide neuropsychological evidence for a role of C.C.'s damaged areas in the control of predictive eye and hand movements to moving objects and show that there is little compensation for these deficits.

Talk Session: Saturday, May 18, 2024, 10:45 am – 12:30 pm, Talk Room 2

Color, Light and Materials: Neural mechanisms, models

Moderator: Bevil Conway, National Eye Institute, NIH

Talk 1, 10:45 am, 22.21

NATURAL RETINAL CONE DISTRIBUTIONS EMERGE FROM OPTICAL AND NEURAL LIMITS TO VISION

Yazhou Zhao[1,2] (nebula@connect.hku.hk), Zeyu Yun[2], Ruichang Sun[2], Dasheng Bi[2]; [1]University of Hong Kong, [2]University of California, Berkeley

Biological visual systems develop highly efficient solutions in response to physical limitations. In particular, the human retinal cone mosaic supports both high spatial and color vision acuity in an imperfect optical environment. Here, we show that naturalistic cone distributions can emerge from simple constraints such as chromatic aberration (CA) and naturalistic behavioral task performance. We model key components of the visual system with a CA-constrained optical simulation, learnable cone mosaic sampling and a state-of-the-art deep artificial neural network. We also designed a custom dataset, ImageNet-Bird, by selecting images that require both high spatial acuity and color acuity for correct classification. By training our model to perform this visual task, we show that the model's emerged cone mosaic resembles a cone mosaic found in humans. One important characteristic is the relative deficiency of S cones compared to M and L cones. Moreover, in a performance comparison experiment using fixed cone mosaics with different S cone ratios, we showed that the performance is consistently better when the S cone ratio is lower. Finally, we also observed that artificial neural networks have a different set of limitations from biological neural systems due to inductive biases imposed by the network architecture; for example, the model requires stationarity in the cone mosaic to achieve a good performance. More generally, our results serve as a concrete instance in which the functional organization of vision is driven by inherent optical and neural limitations, and may provide a new framework for understanding observed statistics of the visual system.

Talk 2, 11:00 am, 22.22

MIDGET RETINAL GANGLION CELL SURROUNDS IN MACAQUE: CONE-SELECTIVE OR NOT?

Nicolas Cottaris[1] (cottaris@upenn.edu), Brian Wandell[2], David Brainard[1]; [1]University of Pennsylvania, [2]Stanford University

Despite decades-long study of macaque midget retinal ganglion cells (mRGC), significant knowledge gaps exist regarding their receptive field (RF) properties. One example is the controversy regarding cone pooling in mRGC surrounds. Anatomy and in-vitro physiology, the latter in peripheral retina, indicate that L- and M-cones contribute non-selectively to mRGC RF surrounds, whereas in-vivo physiology in more central retina indicates that the RF surrounds are highly cone-type selective. To better understand the mRGCs, we developed a model of their linear spatiochromatic RFs. We model the cone inputs to the mRGCs based on anatomical and physiological data, taking into account the impact of physiological optics. Knowledge of these factors allows us to model the mRGC RFs across a large part of the visual field. We use the model to compute responses of synthetic mRGCs to cone-isolating grating and m-sequence stimuli, matched to those that have been employed by in-vivo physiological studies. Simulation enables us to compute the expected in-vivo responses for mRGCs with different surround L- to M-cone ratios. We perform the simulations over a range of eccentricities, taking into account the eccentricity dependence of the physiological optics, the cone fundamentals used to derive cone-isolating stimuli, and the mRGC RF structure. Our results reveal that near the fovea, where centers receive one or two cone inputs, physiological optics significantly enlarges the stimulus-referred RF center, thereby attenuating the antagonistic responses from surround-cones of the same type as the center cone. For this reason, the surround measured in vivo can appear heavily biased toward selective pooling of cones of the non-center cone type. In particular, this happens for models in which the simulated RF surrounds draw indiscriminately on L- and M-cones. This phenomenon, which we observed with both m-sequence and drifting grating simulations, provides a plausible explanation for the discrepancy in conclusions across studies.

Funded by AFOSR FA9550-22-1-0167

Talk 3, 11:15 am, 22.23

NEUROPHYSIOLOGICAL MECHANISMS OF VISION AT THE CENTER-OF-GAZE IN MACAQUE V1

Felix Bartsch[1,2] (felixbartsch@gmail.com), Ramon Bartolo Orozco[2], Ethan Lott[1], Jacob Yates[3], Daniel Butts[1], Bevil Conway[2]; [1]Department of Biology, University of Maryland, [2]Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, [3]Herbert Wertheim School of Optometry and Vision Science, UC Berkeley

Almost nothing is known about the spatial structure of V1 receptive fields at the center-of-gaze, yet these neurons are the building blocks for high acuity vision. This knowledge gap is due to the technical challenges in measuring receptive fields (RFs) at the resolution of single-cone inputs present in the retina (~1 arcminute resolution). Given the distinct anatomical and physiological characteristics of foveal retina, it is possible that RFs in foveal V1 are not simply finer versions of parafoveal cells and may have distinctive aspects of visual processing. Here, we developed a model-based approach leveraging neurophysiological data to refine eye position estimates beyond the limits of hardware-based eye trackers using chronic array recordings of foveal RFs in awake fixating macaque. We presented spatiochromatic noise stimuli while recording across cortical layers, using acute laminar arrays sampling many different columns. We applied data-driven nonlinear models to investigate how subcortical inputs are integrated in V1. We recovered detailed spatial RF structure from 429 cells spanning the very center of gaze up to 2o eccentricity. Foveal V1 cells showed a diversity of RF types; some cells were unmodulated by cone-opponent signals ("luminance-only"), and others were modulated by both luminance and cone-opponent signals. RFs were as small as 4 arcminutes, with features subtending ~2 arcminutes. Luminance-only cells had the finest spatial structure RFs (median RF width = 6 arcmin) and typically showed nonlinear responses; cells modulated by cone-opponent signals integrated over larger areas (median RF width = 12 arcmin) and were more likely to be linear. Even within individual cells, the luminance component had finer spatial RF structure than the cone-opponent components, suggesting limits on spatial acuity for cone-opponent computations. Our measurements offer the first detailed observations of spatiochromatic processing in foveal V1 and offer clues to how V1 RFs are constructed from the photoreceptor mosaic.

NIH Intramural Program; NSF IIS-2113197

Talk 4, 11:30 AM, 22.24

OCTOPUS ELECTROENCEPHALOGRAPHY PERMITS DETECTION OF LIGHT-INDUCED STEADY STATE VISUALLY EVOKED POTENTIALS

Peter Tse[1] (peter.tse@dartmouth.edu), Jay Vincelli[2], Gideon Caplovitz[3], Walt Besio[1]; [1]Dartmouth College, [2]University of Rhode Island School of Engineering, [3]University of Nevada, Reno

We describe the first underwater electroencephalography (EEG) involving octopuses. We can detect stimulus frequency-dependent correlates of flickering LEDs in electrodes placed on their skin under saltwater. This activity is biological in origin, because control experiments rule out that this activity is generated by the flickering light alone in the absence of an octopus. Rather than place electrodes onto the skin of an octopus, we place the octopus between two layers of fixed electrodes under saltwater, while presenting the octopus with visual input from outside its enclosure. We flickered light at various fixed frequencies outside of a transparent enclosure that held individual octopus bimaculoides, loosely sandwiched between two plates containing embedded EEG tripolar electrodes. Neural activity entrained to displayed frequencies can be detected as potential from electrodes situated on or near the midpoint between the two eyes of the octopus, but not from electrodes situated below the brain of the octopus. We are able to detect SSVEPs at multiple tested frequencies. EEG offers the major advantage that it is not invasive, so octopuses need not be fixed in place or anaesthetized. The parallel alignment of neurons in their brain's vertical lobe allows summation of the Local Field Potential (LFP) that emerges from MSF/MIF axons synapsing on the AM cells of the vertical lobe. There is also a similar co-alignment of neurons in some of the other octopus brain lobes, which may permit LFP summation of signals. Conclusion: An electrode placed near the center of the two eyes is able to detect neural responses to light flickering at various frequencies. Octopus EEG may eventually prove to be as fruitful as human EEG has proven to be in deciphering the neural correlates of complex cognition.

This work was supported by NSF grant 2122962

Talk 5, 11:45 AM, 22.25

CAN WE IMPROVE LUMINANCE? ONLINE AND LAB EXPERIMENTS

Shuchen Guan[1] (shuchen.guan@psychol.uni-giessen.de), Robert Ennis[1], Karl Gegenfurtner[1]; [1]Justus-Liebig Universität, Gießen

Luminance has served as the standard measure for light intensity for 100 years. Nevertheless, it has long been known that it has substantial flaws in predicting the perceived intensity of lights with different spectral distributions. Here, we wanted to evaluate potential improvements in the weighting of the cone inputs for heterochromatic brightness perception. To reach a large number of observers, we made measurements in the lab and online, using the same observers. We presented 144 patches encompassing 12 hues and 12 intensities in RGB space. Each trial involved 12 patches varying in both hue and intensity, and 43 observers ranked them based on their perceived brightness in 66 trials. Observers completed the experiment online on personal displays and in a well-controlled lab environment on an OLED. They also brought their personal displays to the lab for calibration. In the lab session, testing observers with a calibrated sRGB display revealed that luminance predicted 76.3% of observer rankings correctly. Radiance predicted more accurately (78.5%), and a non-linear weighted maxRGB model performed best (84.2%). The optimal weights fitted to RGB were [0.40, 0.45, 0.15]. Compared to Vlambda, the contributions of L- and S-cones were increased for heterochromatic brightness. The test-retest reliability, measured with a subset of 20 observers, was 83.9% in these lab-based experiments. For the home session, we first investigated stimuli consistency across displays. The patches presented on the lab OLED had a consistency ranging from 90% to 97% across all predictors on observers' displays based on the calibration data. Intra-observer response consistency across online and lab sessions was 80.8%, inter-observer consistency was 77.6%. Again, the maximum-weighted RGB model consistently outperformed luminance. We conclude that luminance systematically underestimates the contributions of L- and S-cones to heterochromatic brightness. Our results also indicate that online color experiments may be feasible for certain paradigms.

Talk 6, 12:00 PM, 22.26

DECIPHERING VISUAL REPRESENTATIONS BEHIND SUBJECTIVE PERCEPTION USING RECONSTRUCTION METHODS

Fan L. Cheng[1,2], Tomoyasu Horikawa[1,3], Kei Majima[2], Misato Tanaka[1,2], Mohamed Abdelhack[2,4], Shuntaro C. Aoki[1,2], Jin Hirano[2], Yukiyasu Kamitani[1,2]; [1]ATR Computational Neuroscience Laboratories, [2]Kyoto University, [3]NTT Communication Science Laboratories, [4]Krembil Centre for Neuroinformatics

Reconstruction techniques have been widely used to recover physical sensory inputs from brain signals. Numerous studies have consistently refined methods to achieve image reconstruction that faithfully mirrors the presented image at the pixel level. An intriguing extension of these techniques is their potential application to subjective mental contents, a domain that has proven to be especially challenging. Here, we introduce a general framework that can be used to reconstruct subjective perceptual content. This framework translates or decodes brain activity into deep neural network (DNN) representations, and then converts them into images using a generator. Through our research on visual illusions—a classic form of subjective perception defined by a discrepancy between sensory inputs and actual perception— we demonstrate how we successfully reconstructed visual features that were absent in the sensory inputs. Our work shows the potential of reconstruction techniques as invaluable tools for delving into visual mechanisms. The use of natural images as training data and the choice of DNNs were key in obtaining successful reconstruction. While extensive research has probed the neural underpinnings of visual illusions using qualitative hypotheses, our approach materializes mental content into formats amenable to visual interpretation and quantitative analysis. Reconstructions from individual brain areas shed light on the strength of illusory representation and its shared representations with real features at different levels of processing stages, which provides a means to decipher the visual representations underlying illusory perceptions.

Japan Society for the Promotion of Science KAKENHI grant JP20H05705 and JP20H05954, Japan Science and Technology Agency CREST grant JPMJCR18A5 and JPMJCR22P3 and SPRING grant JPMJSP2110, New Energy and Industrial Technology Development Organization Project JPNP20006

Talk 7, 12:15 pm, 22.27

INTEGRATED GRADIENT CORRELATION: A METHOD FOR THE INTERPRETABILITY OF FMRI DECODING DEEP MODELS

Pierre Lelièvre[1] (contact@plelievre.com), Chien-Chung Chen[1];
[1]Visual Neuroscience Lab, Department of Psychology, National Taiwan University, Taipei, Taiwan

Deep learning has reached the domain of visual perception with artificial models trained on image classification tasks, interestingly expressing some degree of similarity with human mechanisms. Currently, encoding/decoding operations of fMRI activation to features of interest usually stick to individual linear regressions per voxel/vertex. Modelers mitigate associated limitations with carefully hand-crafted linearizing features, however, the multidimensionality and intrinsic non-linearities of artificial neural networks could further improve domain adaptation, and even capture brain area interactions. One explanation of favoring simple models is the lack of interpretability of deep learning, i.e. the ability to compare informational content between different brain areas, for one feature, and across different features. We overcome this issue by introducing a new method called Integrated Gradient Correlation, IGC, completing the original IG attribution method. We also demonstrate the relevancy of our approach by investigating the representation of image statistics using the NSD dataset: a public fMRI dataset consisting of 70k BOLD activations acquired during a long term image recognition task. We particularly focused on surface-based data (fsaverage), limited to visual cortex ROIs (e.g. V1-V4, bodies, places). Statistics under scrutiny encompassed three first moments of image luminance distributions usually associated with human texture perception (i.e. mean luminance, contrast, and skewness), as well as a higher level statistic related to spatial luminance distributions (i.e. 1/f slope). Then, we evaluated several decoding models: traditional individual linear regressors, multidimensional linear models trained per ROI and on the whole visual cortex, and finally different deep architectures (sequences of fully connected layers, and/or graph convolutional layers). IGC results show that deep models provide significantly more accurate decoding predictions, and more informative/selective brain activation patterns, coherent with the literature. Consequently, our method could find applications beyond visual neuroscience, and become beneficial to any scientific inquiry using deep models.

Supported by NSTC.

Talk Session: Saturday, May 18, 2024, 2:30 – 4:15 pm, Talk Room 1

Attention: Selection, modulation, resource competition

Moderator: Stefan Van der Stigchel, Utrecht University

Talk 1, 2:30 pm, 24.11

WHEN PROCESSING RELATIONSHIPS, VISUAL PROCESSING CAPACITY IS FAR LESS THAN FOUR

Steven Franconeri[1] (franconeri@northwestern.edu), Tal Boger[2];
[1]Northwestern University, [2]Johns Hopkins University

Vision can provide rapid and powerful processing for some tasks, and encounter strong capacity constraints for others, with a typical limit of processing 4 objects at once. But some evidence suggests an even lower capacity limit when processing relationships between objects. We asked people to explore data visualizations with only 4 values, and found that *half* of viewers easily missed surprising improbable relationships (e.g., a child's height *decreasing* over time, or a better product costing *less*) in these trivially small datasets. The graph's design used spatial grouping cues to implicitly deprioritize an improbable relationship, and when the design instead implicitly prioritized those relationships, they were noticed 1.8x-3.4x more often. These demonstrations support an emerging view of a divide between capacity limits on visual processing: When tracking or memorizing a set of objects, capacity hovers around 4. But when computing relationships that require linking features (e.g., object heights or verbal labels) to particular objects, estimated capacity drops to 1-2. The present experiment is consistent with models that predict that surprisingly low level of ability. On the practical side, the results provide immediate guidance to the scientific community (as well as those in education and in organizations) as producers and consumers of data visualization. Graphs should be designed so that certain relationships are more intuitively recovered, and that 'data storytelling' techniques – highlighting and annotating data visualizations to help viewers quickly see the 'right' pattern – are critical, even within visualizations of trivially simple datasets.

NSF IIS-CHS-1901485

TALK 2, 2:45 PM, 24.12

EFFORT MINIMIZATION DRIVES SACCADE SELECTION

Christoph Strauch[1] (c.strauch@uu.nl), Damian Koevoet[1], Laura Van Zantwijk[1], Sebastiaan Mathôt[2], Marnix Naber[1], Stefan Van der Stigchel[1]; [1]Utrecht University, [2]University of Groningen

What determines where we move our eyes? We here hypothesized saccade costs to determine saccade selection. We first mapped saccade costs across directions by cueing participants to make a saccade towards a specific direction but to withhold the saccade until after the cue had disappeared. During this phase, we measured pupil size - an indicator of noradrenaline release and mental effort - to index cost. Next, we mapped saccade preferences by presenting any two of the previously cost-mapped saccade directions as a two alternative free choice task. For the first time, we here demonstrate that this cost critically underpins saccade selection: When participants chose between the two possible directions, low-effort options were strongly preferred ($R2=0.58$). Notably, saccades curved away from high-cost directions, suggesting an active weighing of costs and inhibition of costly alternatives. This general principle held when participants searched in natural scenes: cost remained a predictor of saccade direction preferences. Strikingly, effortful saccade directions were disproportionally avoided as soon as overall load was increased by introducing a secondary auditory counting task ($R2=0.50$). This implies that cognitive resources are flexibly (dis)allocated from and to oculomotor processes as resource demands change. Together, this shows that even the most subtle differences in cost are actively weighed to tune for resource-efficient behavior. Beyond stimulus material and goals, we therefore argue that eye-movement behavior is largely determined by a distinct and equally fundamental factor: effort.

TALK 3, 3:00 PM, 24.13

RECOGNITION MEMORY FLUCTUATES WITH THE FLOODLIGHT OF ATTENTIONAL STATE

Anna Corriveau[1] (corriveaual@uchicago.edu), Alfred Chao[1], Megan T. deBettencourt[1,2], Monica D. Rosenberg[1,2,3]; [1]Department of Psychology, The University of Chicago, [2]Institute for Mind and Biology, The University of Chicago, [3]Neuroscience Institute, The University of Chicago

Attentional state fluctuates across time and influences what we remember. However, it is not yet understood whether fluctuations in attention affect memory for task-relevant and task-irrelevant information similarly. One possibility is that increased attentional state heightens the roving spotlight of selective attention, resulting in better filtering of irrelevant stimuli. Alternatively, better attentional state may act like a flickering floodlight, with increased attentional capacity allowing for greater processing of irrelevant stimuli. These hypotheses make opposite predictions for the subsequent memory of irrelevant stimuli. We collected two online samples ($N1=188$; $N2=185$) in which participants viewed a stream of trial-unique stimuli (500 trials) consisting of face images superimposed on scene images and were asked to perform a category judgment on either the faces (males vs. females) or scenes (indoors vs. outdoors) by pressing one key for frequent-category images (e.g., males, 90%) and a different key for infrequent images (e.g., females, 10%). Critically, the other category (scenes or faces) was completely irrelevant for the task. Following the sustained attention task, a surprise test probed recognition memory for both relevant and irrelevant stimuli using a 4-point scale. Logistic models tested whether sustained attention measures predicted memory accuracy. Attention lapses (errors to infrequent stimuli) were preceded by established RT signatures of sustained attention, speed ($b1=.640$, $b2=.617$) and variance ($b1=-.296$, $b2=-.223$; all $ps<.001$). As expected, memory was better for task-relevant items ($b1=.722$, $b2=1.37$; all $ps<.001$). Furthermore, correct performance on infrequent trials predicted memory for both task-relevant ($b1=.134$, $p<.001$; $b2=.201$, $p<.001$) and task-irrelevant ($b1=.127$, $p<.001$; $b2=.111$, $p=.033$) stimuli in both experiments. These results support the flickering floodlight view of attentional state, such that moments of high attention improve memory of relevant and irrelevant stimuli.

National Science Foundation BCS-2043740 (M.D.R.)

TALK 4, 3:15 PM, 24.14

SPATIOTEMPORAL REGULARITIES GUIDE MOTOR PREDICTIONS IN A DYNAMIC VISUAL SEARCH

Nir Shalev[1,2] (shalenir@gmail.com), Noam Tzionit[3], Danielle Filmon[3], Anna C. Nobre[2,4], Ayelet N. Landau[3]; [1]Haifa University, Israel, [2]University of Oxford, UK, [3]Hebrew University of Jerusalem, IL, [4]Yale University, USA

Attention allows us to prioritise relevant information and ignore distraction in our sensory environment. Since natural scenes are constantly changing, it is important for us to adapt our attentional priorities accordingly. Predictable signals, like traffic lights, allow for anticipation and help us control attention in time and space. In this study, we explore how prediction-led attention affects how we guide the motor and oculomotor system in time and space. We used a dynamic variation of a visual search task, with trials lasting 14 seconds. Each trial included eight targets that faded in and out of the display, among visual distractors. Participants moved their eyes freely and used the mouse pointer to click on the targets. Critically, we embedded in each trial spatiotemporal regularities by presenting four out of eight targets at the same time and approximate location throughout the experiment. The remaining four targets could appear at any time and location. We also manipulated the distraction load by varying the number of irrelevant stimuli appearing in each trial. Our results offer a detailed description of the learning dynamics and prediction formation. Participants were faster and more accurate at detecting predictable targets compared to unpredictable ones. In line with the visual search literature, we also found that increasing the number of visual distractors reduced accuracy and slowed down responses. By tracking mouse and eye movements, we discovered that predictions enabled earlier and faster movements towards targets. Interestingly, we also observed earlier and more pronounced movements of the hand and eyes away from predictable targets once they were selected. These findings enhance our understanding of the real-time impact of prediction formation. In our presentation, we will provide a detailed description of these patterns under varying levels of visual distraction and discuss how they emerge during the task as a consequence of learning.

ANL: JSMF Scholar Award, ISF (958/16 & 1899/21), TIMECODE ERC Starting Grant (852387), Joy Ventures, and the Product Academy Award. NS: Daniel Turnberg Fellowship, Academy of Medical Sciences. ACN: Wellcome Investigator Award (104571/Z/14/Z) and the JSMF Collaborator Award (220020448)

TALK 5, 3:30 PM, 24.15

ATTENTION ROBUSTLY DISSOCIATES OBJECTIVE PERFORMANCE AND SUBJECTIVE VISIBILITY REPORTS

Karen Tian[1] (ktian@bu.edu), Brian Maniscalco[2], Michael Epstein[1], Angela Shen[2], Olenka Graham Castaneda[2], Giancarlo Arzu[2], Taiga Kurosawa[2], Jennifer Motzer[1], Emil Olsson[2], Lizbeth Romero[2], Emily Russell[1], Meghan Walsh[1], Juneau Wang[1], Tugral Bek Awrang Zeb[2], Richard Brown[3], Victor Lamme[4], Hakwan Lau[5], Biyu He[6], Jan Brascamp[7], Ned Block[6], David Chalmers[6], Megan Peters[2], Rachel Denison[1]; [1]Boston University, [2]University of California Irvine, [3]City University of New York, [4]University of Amsterdam, [5]RIKEN Center for Brain Science, [6]New York University, [7]Michigan State University

Background: Findings of subjective inflation, in which subjective reports of unattended, peripheral stimuli are stronger than the accuracy of sensory processing would suggest, have motivated higher-order theories of consciousness. However, empirical tests of subjective inflation have been surprisingly limited. Generally they have used a single pair of near-threshold stimulus strengths–weaker for attended and stronger for unattended–to equate objective performance, leaving it unclear whether inflation arises from decision biases and whether inflation extends beyond threshold perception. Goal: In a preregistered adversarial collaboration, we rigorously tested whether attention dissociates subjective reports and objective performance across a range of stimulus strengths and types. Methods: In three experiments, human observers (n=30/experiment) performed a spatial attentional cueing task. On each trial, observers viewed up to four peripheral targets, which varied independently across 7 stimulus strengths. A central precue (60% valid, 20% neutral, 20% invalid) directed attention to one or all target locations. A response cue instructed observers to simultaneously make 1) an objective orientation report and 2) a subjective visibility report. Targets were texture-defined figure-ground ovals (Experiments 1 and 2) or contrast-defined gratings (Experiment 3), presented at threshold (Experiments 1 and 3) or suprathreshold (Experiment 2) stimulus strengths. To assess subjective inflation, we developed an area-under-the-curve approach to quantitatively relate objective and subjective reports across stimulus strengths for matched levels of orientation discriminability. Results: We found strong and consistent subjective inflation under inattention across all experiments. Across a range of threshold and suprathreshold stimulus strengths, and different stimulus types, subjective visibility was reported as higher for unattended vs. attended stimuli when orientation discriminability was equated. Conclusion: Inattention robustly inflates subjective visibility reports, and inflation is not confined to threshold vision. Whether sensory signals are sufficient for explaining subjective visibility reports when they come apart from objective performance may help arbitrate between competing theories of consciousness.

Templeton World Charity Foundation Accelerating Research on Consciousness initiative TWCF 0567 (to BH, JB, NB, DC, RD, MP)

TALK 6, 3:45 PM, 24.16

PUPIL SIZE REVEALS PRESACCADIC ATTENTIONAL SHIFTS UP AND DOWNWARD: A POSSIBLE DISSOCIATION BETWEEN THE WHERE AND HOW OF ATTENTION

Damian Koevoet[1], Christoph Strauch[1], Marnix Naber[1], Stefan Van der Stigchel[1]; [1]Utrecht University

Humans frequently move their eyes to foveate relevant information in the world. It is dominantly assumed that attentional shifts must precede saccades to prepare the brain for postsaccadic retinal input, allowing for perceptual continuity across eye movements. A recent surge of studies have investigated visual anisotropies around the visual field, including presaccadic attention. Such studies demonstrated benefits of presaccadic attention on task performance for horizontal and downward, but not for upward saccades. This contrasts the dominant view: if attention is not moved prior to upward saccades, presaccadic attention may not be necessary to facilitate perceptual continuity. Here we capitalized on the fact that the pupil light response robustly tracks attention to investigate whether presaccadic attention moves up and downwards. We crucially manipulated whether presaccadic attention could shift toward the background brightness of the ensuing saccade target by presenting the brightness throughout the trial, or by presenting the brightness upon saccade onset. In two experiments, we observed acceleration of the onset of the pupil light response for both upward and downward saccades when the landing brightness could be prepared prior to the saccade. This shows that presaccadic attention is deployed, and can facilitate perceptual continuity along the vertical meridian. In combination with previous work, these results suggest that presaccadic attention can be shifted in space without enhancing specific facets (e.g. contrast sensitivity) of visual processing at the deployed location. The known underrepresentation of the upper visual field in early visual cortex may underlie the dissociation between where attention is deployed and how it affects visual processing. However, more work is necessary to identify when, and how, such dissociations occur.

TALK 7, 4:00 PM, 24.17

ATTENTIONAL SAMPLING BETWEEN EYE CHANNELS

Ayelet Landau[1] (ayelet.landau@gmail.com), Daniele Re[1]; [1]The Department of Psychology and the Department of Cognitive and Brain Sciences, The Hebrew University of Jerusalem

Eye channels refer to processing of visual information from each eye before integration in V1 (Hubel & Wiesel, 1977). During development, inputs from both eyes initially overlap in the visual cortex. However, through competitive interactions between neurons with different ocular preferences, the inputs become segregated into distinct columns (Hensch, 2005). This process involves competition between neurons

representing the left and right eyes (Tagawa et al., 2005). When the visual system processes several inputs, competitive and suppressive interactions are foundational to the neuronal response. Attention, which is the biasing selection towards relevant parts a scene, was previously found to be implemented through rhythmic brain activity. Similar to brain rhythms, also performance, fluctuates over time. Specifically, when more than one object is attended, objects are selected in alternation. In this study we sought to investigate whether this phenomenon, called attentional sampling, also emerges in the unconscious selection process among eye channels. We presented a display with a single object to both eyes and manipulated the presentation of a cue and a detection target to either both eyes or to the different eyes. We assume that presenting a cue to one eye biases the selection process to content presented to that eye. Target detection fluctuated at 8 Hz under the binocular condition, and at 4 Hz when the dominant eye was cued. This is consistent with findings reporting that competition between receptive fields leads to sampling. The findings also demonstrates that sampling in light of competition does not rely on aware processes.

The Brain Attention and Time Lab (ANL) is grateful for the support of the James McDonnell Scholar Award in Understanding Human Cognition, ISF Grants 958/16 and 1899/21, TIMECODE ERC Starting Grant No. 852387 as well as Joy Ventures Research Grant and the Product Academy Award.

Talk Session: Saturday, May 18, 2024, 2:30 – 4:15 pm, Talk Room 2

Face and Body Perception: Development, disorders, models, neural mechanisms

Moderator: Benjamin van Buren, The New School

Talk 1, 2:30 pm, 24.21

PERCEIVED AGE IS DISTORTED IN VISUAL MEMORY: A PHENOMENON OF "FORWARD" AND "BACKWARD" AGING FOR FACES

Didi Dunin[1] (dunid901@newschool.edu), Joan Danielle K Ongchoco[2], Benjamin van Buren[1]; [1]The New School, [2]University of British Columbia

When we meet someone, we quickly make judgments about them based on how old they look (e.g. about their physical abilities, cognitive abilities and personality traits). But how is a person's age represented in the mind in the first place? Do we remember certain people as younger, or as older, than they actually were? One possibility is that representations of facial age exhibit 'representational momentum', such that observers remember a face as older than it actually was. Another possibility is that our memory for facial age is biased towards the average of the faces that we have seen previously, in which case observers might misremember faces as closer to middle age. To explore these possibilities, we ran three experiments which tested participants' memory for the age of a briefly presented face. Participants saw a target face which was either young (30 years old) or old (60 years old). Subsequently, they saw two new faces – one 10 years younger and another 10 years older than the target. Participants selected the face that matched the target. Contrary to our initial predictions, we did not find a bias to remember faces as older, or as closer to middle age. Instead, a distinct pattern emerged — observers were biased to remember young targets as younger (i.e. 'backward aging'), and old targets as older (i.e. 'forward aging'). Remarkably, these biases held across sexes (male, female) and races (asian, black, white) of the target face, across artificially-aged and real faces, and regardless of the observers' own age. Further, the results persisted even when the decoys' identities differed from that of the target face — suggesting that this bias operates over abstract representations of age. Thus, social categories of 'young' and 'old' shape and distort our visual memories of faces.

Talk 2, 2:45 pm, 24.22

FACE-SPECIFIC IDENTIFICATION IMPAIRMENTS FOLLOWING SIGHT-PROVIDING TREATMENT MAY BE ALLEVIATED BY AN INITIAL PERIOD OF LOW VISUAL ACUITY

Sharon Gilad-Gutnick[1] (sharongu@mit.edu), Fengping Hu[2], Kirsten A. Dalrymple[3], Priti Gupta[4], Pragya Shah[5], Chetan Ralekar[1], Dhun Verma[6], Kashish Tiwari,[7], Piyush Swami[8], Suma Ganesh[6], Umang Mathur[6], Pawan Sinha[1]; [1]Massachusetts Institute of Technology, [2]New York University, [3]HealthPartners Institute, [4]Institute of Technology, New Delhi, [5]Institute of Human Behavior and Allied Sciences, New Delhi, [6]Dr. Shroff's Charity Eye Hospital, New Delhi, [7]Dr. Rajendra Prasad Center for Ophthalmology, All India Institute of Medical Sciences, New Delhi, [8]Technical University of Denmark

The ability to identify individual faces is critical for social and cognitive functioning, and as such, the human brain has evolved to perform this task quickly and accurately. However, many questions remain about how this skill emerges in early development, and specifically about how early visual experience impacts skill acquisition later in life. In our previously published work, we proposed that the poor visual acuity that newborns experience in the first year of life may play a facilitatory role in scaffolding the processes needed to develop face-identification later in life. Indeed, our computational simulations supported the potential downsides of 'High Initial Acuity' for the development of facial-identification. Motivated by this, we predicted that children who are treated for congenital cataracts late in life and begin their visual journey with higher than newborn acuity will exhibit persistent impairments in face- but not object- identification. We tested this prediction by assessing the development of facial-identification skill in three subject groups: children treated for congenital cataracts whose pre-treatment visual acuity was worse than that of a newborn, those treated for congenital cataracts whose pre-treatment visual acuity was better than that of a newborn, and age-matched controls. As predicted, we found that children with pre-operative acuity worse than a newborn did not show any improvements on face-identification tasks despite years of visual experience and improving on the object-identity tasks. In contrast, those with pre-treatment acuity comparable to a newborn showed improvements on both the object- and face-identification tasks. Overall, our data is consistent with the idea that beginning one's visual journey with a period of low-resolution visual input followed by high-resolution input can be facilitatory for acquiring face-identification skill later in life, whereas higher resolution input right at the outset of vision can be detrimental to facial- but not object-identification.

NIH R01EY020517

Talk 3, 3:00 pm, 24.23

WHAT'S LEFT IN FACE PROCESSING? EVIDENCE FROM HEMISPHERIC DIFFERENCES IN DEVELOPMENTAL PROSOPAGNOSIA

Alison Campbell[1,2] (alison.candice.campbell@gmail.com), Xian Li[1,4], Michael Esterman[1,3], Joseph DeGutis[1,4]; [1]Boston Attention and Learning Laboratory, VA Boston Healthcare System, Boston, MA, [2]Department of Psychiatry, Boston University Chobanian and Avedisian School of Medicine, Boston MA, [3]National Center for PTSD, VA Boston Healthcare System, Boston, MA, [4]Department of Psychiatry, Harvard Medical School, Boston MA

According to classic models, face processing is right-lateralized with little involvement of the left hemisphere. This is challenged by fMRI findings that developmental prosopagnosics (DP) consistently show reduced face-selective responses in the left OFA/FFA and less consistent differences in the right OFA/FFA. To account for this, we hypothesized that right hemisphere regions primarily subserve processes for face perception and, as this is highly variable in DP, we predicted that right-sided regions would only be implicated in those with greater perceptual impairment. In our sample, DPs with low performance (>1SD below controls) on at least two perceptual tests (Benton Face Recognition Test, Cambridge Face Perception Test, and face matching) were classified as perceptually impaired (N=17) and the remaining were classed as perceptually unimpaired (N=18). No controls were impaired (N=22). Using a face localizer (Faces>Objects), we found that perceptually-impaired DPs had reduced face-selective activation in both the left and right OFA, whereas perceptually-unimpaired DPs had reduced activation only in the left OFA. Both groups had reduced activation in the left but not the right FFA. Furthermore, resting-state functional connectivity between the left and right OFA was significantly reduced in perceptually-impaired but not perceptually-unimpaired DPs, consistent with neural abnormalities spanning both hemispheres in the presence of perceptual deficits. The results support the hypothesis that right hemisphere abnormalities (especially in the OFA) reflects a perceptual processing deficit that is variable in DP and explains why not all studies observe right hemisphere differences at the group level. Critically, our results suggest that left hemisphere abnormalities are common to all DPs. Although future work is needed to clarify the functional roles of the left OFA/FFA, their implication in DP suggests that they are essential for normal face recognition and are required for a complete neural model of face processing.

This work was supported by a grant to JD from the National Eye Institute (R01 EY032510-02).

Talk 4, 3:15 pm, 24.24

COMPARING FACE VIEWPOINT, EXPRESSION AND IDENTITY SELECTIVITY IN FMRI-DEFINED FACE PATCHES OF MACAQUE FRONTAL CORTEX

Eline Mergan[1,2] (eline.mergan@kuleuven.be), Qi Zhu[1,2,3], Wim Vanduffel[1,2,4,5]; [1]KU Leuven, [2]Leuven Brain Institute, [3]CEA DRF/JOLIOT/NEUROSPIN, Univ. Paris-Saclay, [4]Radiology, Harvard Medical School, [5]Martinos Centre for Biomedical Imaging

Perceiving and interpreting facial information such as identity, expression, and head orientation is essential for primates as these features provide important social communication cues. To gain insights into the underlying neural mechanisms processing these facial features in prefrontal cortex, we conducted single- and multi-unit recordings in 4 fMRI-defined face patches in 3 macaques: POa and POp in orbitofrontal cortex, PA in ventrolateral prefrontal cortex, in addition to face patch AM in anterior inferotemporal cortex. In each face patch, we found face-selective neurons tuned to identity, expression, and head orientation. A large fraction of these neurons was sensitive to head orientation irrespective of identity or expression. While face-selective neurons preferentially tuned for expressions were mostly present in POa, most face-selective neurons tuned for identity resided in AM. Surprisingly, not only face-selective but also non-face-selective neurons carried similar information about these facial features. Most neurons within each face patch exhibited visual response latencies that were comparable for different face features. At population level, visual response latencies for faces were similar (~70 ms) for both orbitofrontal (POa and POp) and anterior IT (AM) face patches. In prefrontal face patch PA, however, most cells responded much faster to faces, as quick as 30 ms. Neurons generally exhibited fastest face-selective responses (face contrasted with objects), followed by selectivity to head orientations, and yet later for different expressions and identities. While the range of visual response latencies in each face patch is relatively small (interquartile range (iqr): 30 - 80 ms), the selective responses to various face features showed a considerable variation (iqr: 80 - 190 ms). These findings reveal complex prefrontal face-processing signals potentially involving multiple and parallel feedback loops with different areas, prompting a reconsideration of the role of the face-processing system in representing face viewpoint, expression and identity.

FWO G0C1920N, The European Union's Horizon 2020 Framework Program for Research and Innovation under Grant Agreement No 945539 (Human Brain Project SGA3), KU Leuven C14/21/111, CEA PE bottom up 2020 (20P28), ANR-20-CE37-0005

Talk 5, 3:30 pm, 24.25

FROM DIVERGENCE TO CONVERGENCE: A MODEL-GUIDED SYNTHESIS OF FINDINGS IN THE HUMAN AND MACAQUE FACE PROCESSING NETWORKS

Fernando M. Ramirez[1,2] (fernando.ramirez@nih.gov); [1]Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, [2]Section on Functional Imaging Methods, Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health

Recognizing faces regardless of viewpoint is critical for social interactions. Evidence from single-neuron electrophysiological recordings in macaques suggests a three-step architecture revealing

a sharp transition from a strictly view-tuned representation in the macaque middle-lateral/middle-fundus (ML/MF) face patches to a mirror-symmetric representation in the anterior-lateral (AL) face-patch, before achieving viewpoint invariance in the anterior-medial (AM) face-patch, at the highest level of the hierarchy. However, human studies combining functional magnetic resonance imaging (fMRI) and Representational Similarity Analysis (RSA) have led to divergent conclusions in all core face selective areas, including the Fusiform Face Area (FFA). This makes it hard to relate observations within and across species. We previously proposed a geometric configuration in multivariate space that accounts for divergent observations in human FFA. Here, by considering the impact on RSA of signal imbalances across conditions and measurement scale, we show that this geometric configuration is compatible with observations in macaque area ML/MF, but not AL. Our account shows that key assumptions of RSA sometimes break down. Specifically, we show that inferences about neuronal coding with RSA are influenced by translation and rotation of the data. We also show that abstracting from the measurement process and relying directly on the rank-order of entries of dissimilarity matrices to relate representations across species and techniques leads to error when marked signal-imbalances are observed across conditions. We demonstrate with biologically-motivated network models, forward models, as well as previously published empirical fMRI data and single-cell monkey electrophysiological recordings that it is necessary to consider details of the measurement process to validly relate measurements across species and techniques. These findings suggest limitations in RSA, urging a nuanced approach for cross-species comparisons, and support the idea that human FFA is view-tuned like macaque area ML/MF, rather than mirror-symmetrically tuned like area AL.

This work was supported by the Intramural Research Program at NIMH (ZIAMH002783)

Talk 6, 3:45 pm, 24.26

INTRACEREBRAL RECORDINGS EVIDENCE THAT UNFAMILIAR FACE-IDENTITY RECOGNITION IS SUPPORTED BY FACE-SELECTIVE NEURAL POPULATIONS IN THE HUMAN VENTRAL OCCIPITO-TEMPORAL CORTEX

Simen Hagen[1] (simenhagen1@gmail.com), Corentin Jacques[1], Louis Maillard[1,2], Sophie Colnat-Coulbois[1,3], Jacques Jonas[1,2], Bruno Rossion[1,2]; [1]Université de Lorraine, CNRS, Nancy, France, [2]Université de Lorraine, CHRU-Nancy, Service de Neurologie, Nancy, France, [3]Université de Lorraine, CHRU-Nancy, Service de Neurochirurgie, Nancy, France

In humans, the recognition of a visual stimulus as a face – generic face recognition (GFR) - and of its specific identity – face identity recognition (FIR) - are intricately linked and both functions are supported by specialized neural regions in the human ventral occipito-temporal cortex (VOTC). However, whether they are instantiated by the same or different neural populations remains unclear. On the one hand, FIR could rely largely on "different" neural populations that receives input from neural populations involved in GFR. On the other hand, FIR could rely on "shared" neural populations that support both functions potentially at different time scales. Here, we directly compared the spatio-temporal profile of the two recognition functions in a large group of epileptic patients (N=109) implanted with intracerebral electrodes in the gray matter of the VOTC. Both GFR (i.e., significantly different responses to faces vs. non-face visual objects; Jonas et al., 2016) and FIR (i.e., significantly different responses to different unfamiliar face identities; Jacques et al., 2020) neural activity was isolated with separate frequency-tagging protocols within patients. This approach provides an objective measure of the two recognition functions, parceling out general visual responses, and providing high spatial and temporal resolution. Across all the significant FIR recording contacts, we found that ~85% also showed significant GFR responses (i.e., were face-selective). This high spatial overlap was found along the posterior-anterior axis and within all core face regions. Moreover, in the overlap contacts, the amplitudes for the two functions correlated (r>.8) and the temporal onset of amplitudes for GFR and FIR was strikingly similar regardless of posterio-anterior location, but with a relatively slower build-up of the FIR amplitudes. Overall, this original dataset suggest that unfamiliar FIR is essentially supported by face-selective neural populations in the human VOTC, with GFR signals potentially transmitted faster than FIR signals.

Funded by: ANR IGBDEV ANR-22-CE28-0028; ERC AdG HUMANFACE 101055175

Talk 7, 4:00 pm, 24.27

THE HIDDEN DETAILS: EFFECTS OF PARTIAL OCCLUSION ON RESPONSE DYNAMICS IN THE PRIMATE INFEROTEMPORAL CORTEX

Anna Bognar[1] (anna.bognar@kuleuven.be), Ghazaleh Ghamkhari Nejad[1], Rufin Vogels[1]; [1]KU Leuven

The primate brain can recognize objects even when partially concealed by occluders. To investigate the effect of occlusion on temporal dynamics of neuronal responses, we conducted experiments in two male macaques, recording single units in body-responsive regions in the posterior and anterior inferotemporal cortex (PIT & AIT) during fixation. Seven levels of occlusion were applied to static bodies, ranging from 5 to 60 percent occlusion. In both monkeys and regions, three key findings emerged: 1) average response strength decreased and 2) response onset and peak latency gradually increased by ~70 ms with degree of occlusion, PIT responses consistently preceding AIT. 3) The first response peak was followed by a trough and a stronger second peak under occlusion. To examine the role of visual information loss in the latency shifts, reduced responses, and response peaks, we presented, in addition to the partially occluded bodies, the same stimuli on top of the occluding pattern, and with an invisible occluding pattern, creating bodies with cut-outs. Interestingly, onset latency only shifted ~20 ms for the highest cut-out levels and remained unaffected by the background occluding pattern. Thus, onset latency shifts with occlusion may result from bottom-up occluder-related processing. Despite cut-out-induced response weakening, cut-outs with 60% information loss maintained selectivity similar to that observed during occlusion. However, the trough formation was pronounced when bodies were presented on top of the occluder. Intriguingly, the second peak did not align with response onset shifts but maintained latency differences between regions, occurring earlier in PIT. Thus, the second response peak in PIT is unlikely to arise from recurrent processing within the region or feedback from AIT. If generated by top-down feedback, one would expect it to appear earlier in AIT and may expect better body selectivity. Yet, it never surpassed early response selectivity based on neural decoding.

This research was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 856495).

Talk Session: Saturday, May 18, 2024, 5:15 – 6:45 pm, Talk Room 1

Perception of Relations, Intuitive Physics

Moderator: Joan Ongchoco, Humboldt-Universität zu Berlin

Talk 1, 5:15 pm, 25.11

WHAT NEWTON DID NOT KNOW ABOUT NEWTON'S CRADLE: SEPARATING VISUAL ROUTINES FOR CAUSE AND EFFECT

Sven Ohl[1] (sven.ohl@hu-berlin.de), Martin Rolfs[1]; [1]Humboldt-Universität zu Berlin

In Newton's cradle a moving object collides with a line of touching stationary objects, causing the object at the very end of the line to move. This demonstration of Newton's first law of motion is fascinating to watch because the cause and effect of the motion are spatially separated. Here, in a modified version of Newton's cradle, we exploit this separation in a visual adaptation paradigm to show that there are separate visual routines for detecting cause and effect in a causal interaction. We presented launching events in which a moving disc stopped next to another disc with varying degrees of overlap, and asked observers to indicate whether the first disc caused the second disc to move, or whether the first disc simply passed a stationary one. We fitted psychometric functions to each observer's reports as a function of disc overlap and determined how these functions were affected by the prolonged presentation of a modified version of Newton's cradle (i.e., the adaptor). Critically, we obtained psychometric functions for the perceived causality at the cause location and at the effect location in Newton's cradle and we observed significant negative aftereffects at both the cause and at the effect location (Experiment 1 and 2). Observers reported fewer launches at these locations only when the motion direction of the test event was the same as the adaptor's motion direction (Experiment 1). Critically, the adaptation was spatially specific: Perception of launches at the location in-between the cause and the effect locations was not affected by adaptation (Experiment 2). These results provide compelling evidence that the perception of causality integrates information from both the cause and the effect location. This integration allows the detection of causal interaction even when the cause and effect are spatially separated in the visual environment.

Talk 2, 5:30 pm, 25.12

BREAKING DOWN A GOLF SWING: SPATIO-TEMPORAL DYNAMICS OF VISUAL MOTION UNDERLIE HIGH-LEVEL STRUCTURING OF OBSERVED ACTIONS

Zekun Sun[1] (zekun.sun@yale.edu), Wenyan Bi[1], Ilker Yildirim[1,2], Samuel McDougle[1,2]; [1]Department of Psychology, Yale University, [2]Wu Tsai Institute, Yale University

To acquire or demonstrate a motor skill, we often break it down into a sequence of steps (e.g., a golf swing has "backswing" and "downswing" phases). But do we *see* single, smooth actions as containing discrete events? We compiled 20 animations depicting natural actions, spanning sports (e.g., kicking a ball), exercises (e.g., a jumping jack), and everyday tasks (e.g., picking up an object). In Experiment 1, observers determined a "boundary" to divide each action into two meaningful units. Consensus among observers implied a similar interpretation of the event structure of each action. Next, we explored whether these actions are spontaneously segmented during visual processing. We reasoned that if we visually represent actions as being divided into units by boundaries, then subtle changes occurring at these boundaries – specifically during the transition between the units – should be less noticeable relative to non-boundary moments. Experiments 2-3 tested observers' detection of transient slowdowns and frame shifts at boundary, pre-boundary and post-boundary frames. People were worse at detecting changes at boundaries compared to non-boundaries. What kind of information about observed actions drives this effect? Experiments 4-5 applied novel distortions to the videos, removing high-level semantic information while preserving lower-level spatial-temporal dependencies. The boundary effect was weakened yet persisted, suggesting that spatio-temporal dynamics play a crucial role in mental structuring of actions. To quantify these dynamics, we extracted optical flow fields from every two consecutive frames of each video and computed 16 motion statistics from the flow maps to capture global and local motion characteristics. We found that the boundary judgments in Experiment 1 could be predicted by the changes in the magnitude and direction of motion vectors, especially the smoothness of these variations. Our results suggest that the visual system automatically imposes boundaries when observing natural actions via image-computable, spatio-temporal motion patterns.

NIH - R01NS132926

Talk 3, 5:45 pm, 25.13

FAST AND AUTOMATIC PROCESSING OF RELATIONS: THE CASE OF CONTAINMENT AND SUPPORT

Sofie vettori[1,2] (sofievettori@gmail.com), Jean-Rémy Hochmann[1,2], Liuba Papeo[1,2]; [1]Institut des Sciences Cognitives—Marc Jeannerod, UMR5229, Centre National de la Recherche Scientifique (CNRS), [2]Université Claude Bernard Lyon 1

Achieving a meaningful representation of the visual environment, one that can be useful for navigating, planning and acting, requires representing objects and the relations between them. We know that object recognition is efficient, i.e., reportedly fast and automatic; how fast and automatic is the processing of relations? We studied this, focusing on the fundamental relations containment and support, using frequency-tagging electroencephalography (FT-EEG). FT-EEG allows to pinpoint automatic stimulus-locked responses. First, we tested –and demonstrated– that relations between multiple objects are accessed as fast and automatically as the object themselves. Twenty adults viewed a sequence of images with object pairs at a base-frequency (2.5 Hz), where every four stimuli illustrating one relation (support: book on table, knife on chop-board), one oddball-stimulus illustrating the other relation appeared (containment: spoon in cup) (oddball-frequency: 0.625 Hz). EEG signals indicated responses at both frequencies, meaning that participants processed each image and spontaneously detected changes in the relation carried by oddball-stimuli. A control condition demonstrated that the oddball-response was not due to a regular repetition of the objects (spoon and cup). Since the above effect was found with oddball stimuli that involved (different instances of) the same objects (e.g., always spoon in cup), we tested whether the same effect could be found when only the relation remained identical (e.g., containment), while the objects changed for every oddball-stimulus (spoon in cup, fish in bowl). Here, the oddball-response remained significant, demonstrating that it reflected encoding of the relation itself, regardless of the objects involved in it. Finally, the oddball-response remained unchanged when participants were explicitly instructed to attend to the relation, indicating that the encoding of relations is independent from attention. We conclude that relations between objects are encoded rapidly, automatically upon stimulus presentation and in a manner that generalizes over a broad class of objects.

SV was supported by a postdoctoral fellowship awarded by a Marie Skłodowska-Curie individual fellowship (MSCA-IF 101108756) by the European Commission. LP was supported by a European Research Council Grant (Project THEMPO-758473).

TALK 4, 6:00 PM, 25.14

JOINT COMMITMENT IN HUMAN COOPERATIVE HUNTING THROUGH AN "IMAGINED WE"

Siyi Gong[1], Ning Tang[2], Minglu Zhao[1], Jifan Zhou[2], Mowei Shen[2], Tao Gao[1]; [1]University of California, Los Angeles, [2]Zhejiang University

For human cooperation, jointly selecting a goal out of multiple comparable goals and maintaining the team's joint commitment to that goal poses a great challenge. By combining psychophysics and computational modeling, we demonstrate that visual perception can support spontaneous human joint commitment without any communication. We developed a real-time multi-player hunting task where human hunters could team up with human or machine hunters to pursue prey in a 2D environment with Newtonian physics. Joint commitment is modeled through an "Imagined We" (IW) approach, wherein each agent uses Bayesian inference to reason the intention of "We", an imagined supraindividual agent that controls all agents as its body parts. This model is compared against a Reward Sharing (RS) model, which posits cooperation as sharing reward through multi-agent reinforcement learning (MARL). We found that both humans and IW, but not RS, could maintain high team performance by jointly committing to a single prey and coordinating to catch it, regardless of prey quantity or speed. Human observers also rated all hunters of both human and IW teams as having high contributions to the catch, irrespective of their proximity to the prey, suggesting that their high-quality hunting resulted from sophisticated cooperation rather than individual strategies. IW hunters could not only cooperate with their own kind but also with humans, with human-IW teams mirroring the hunting performance and teaming experience of all-human teams. However, substituting human members with more RS hunters reduced both performance and teaming experience. In conclusion, this study demonstrates that humans achieve cooperation through joint commitment that enforces a single goal on the team, rather than merely motivating team members through reward sharing. By extending the joint commitment theory to visually grounded cooperation, our research sheds light on how to build machines that can cooperate with humans in an intuitive and trustworthy manner.

TALK 5, 6:15 PM, 25.15

UNCONSCIOUS INTUITIVE PHYSICS: PRIORITIZED BREAKTHROUGH INTO VISUAL AWARENESS FOR PHYSICALLY UNSTABLE BLOCK TOWERS

Kimberly W. Wong[1] (kimberly.wong@yale.edu), Aalap Shah[1], Brian Scholl[1]; [1]Yale University

A central goal of perception and cognition is to predict how events in our local environments are likely to unfold: what is about to happen? And of course some of the most reliable ways of answering this question involve considering the regularities of physics. Accordingly, a great deal of recent research throughout cognitive science has explored the nature of 'intuitive physics'. The vast majority of this work, however, has involved higher-level reasoning, rather than seeing itself—as when people are asked to deliberate about how objects might move, in response to explicit questions ("Will it fall?"). Here, in contrast, we ask whether the apprehension of certain physical properties of scenes might also occur *unconsciously*, during simple passive viewing. Moreover, we ask whether certain physical regularities are not just processed, but also visually *prioritized*—as when a tower is about to fall. Observers viewed block towers—some stable, some unstable—defined in terms of whether they would collapse as a result of external physical forces (such as gravity) alone. We used continuous flash suppression (CFS) to render the towers initially invisible: observers viewed them monocularly through a mirror haploscope, while a dynamic Mondrian mask was presented to their other eye. We then measured how long towers took to break through this interocular suppression, as observers indicated when they became visually aware of anything other than the mask. The results were clear and striking: unstable towers broke into visual awareness faster than stable towers. And this held even while controlling for other visual properties—e.g. while contrasting pairs of stable vs. unstable towers sharing the same convex hull, and differing only in the horizontal placement of a single block. This work shows how physical instability is both detected and prioritized, not only during overt deliberation, but also in unconscious visual processing.

TALK 6, 6:30 PM, 25.16

DECODING PREDICTED FUTURE STATES FROM THE BRAIN'S 'PHYSICS ENGINE'

RT Pramod[1,2] (pramodrt@mit.edu), Elizabeth Mieczkowski[3], Cyn Fang[1,2], Josh Tenenbaum[1,2], Nancy Kanwisher[1,2]; [1]Department of Brain and Cognitive Sciences, MIT, [2]McGovern Institute for Brain Research, MIT, [3]Princeton University

Successful engagement with the physical world requires rapid online prediction, from swerving to avoid a collision to returning a ping-pong serve. Here we test the hypothesis that physical prediction is implemented in a set of parietal and frontal regions (aka the "hypothesized Physics Network " or hPN) that model the structure of the relevant scene and run forward simulations to predict future states. For physical scene understanding and prediction, contact relationships between objects such as support, containment, and attachment are critical because they constrain an object's fate: if a container moves, so does its containee. In Experiment 1, participants (N = 14) were scanned with fMRI while viewing short videos (~3s) depicting contact (contain, support, attach) and non-contact events. MVPA revealed scenario-invariant decoding of the presence versus absence of a contact relationship that was significant in the hPN but not in the ventral pathway. Experiment 2 tested whether the hPN also carries information about predicted future contact events, as expected if the hPN is engaged in forward simulation. Indeed, the voxel response patterns in hPN distinguishing between perceived contact and no-contact events were similar even for predicted events where contact was predictable but not shown. This prediction of future contact events, which generalized across objects and scenarios, was found even though participants were performing an unrelated one-back task, and was detected only in the hPN, not the ventral visual pathway. In both experiments, the key results were absent in the primary visual cortex, arguing against low-level visual feature confounds accounting for these findings. Thus, we find that the hPN both (a) encodes physical relationships between objects in a scene, and (b) predicts future states of the world, as expected if this network serves as the brain's 'Physics Engine'.

This project was funded by NSF NCS Project 6945933 (NGK)

Talk Session: Saturday, May 18, 2024, 5:15 – 6:45 pm, Talk Room 2

Visual Search 1

Moderator: Ruth Rosenholtz, NVIDIA Research

Talk 1, 5:15 pm, 25.21

SHOPPING AT EYE-KEA: ASSESSING THE HIERARCHY OF VISUAL SEARCH PRIORITIES IN A VIRTUAL FURNITURE STORE

Shea E. Duarte[1,2] (seduarte@ucdavis.edu), Joy Geng[1,2]; [1]University of California, Davis, [2]University of California, Davis Center for Mind & Brain

Most studies of attention use experimental designs with short trial structures and simple two-dimensional displays. These studies have demonstrated that search can be made more efficient by attending to items that match a target along categorical and feature dimensions. However, real world visual search unfolds across space over extended periods of time, such as when shopping for items from a list. It is therefore unclear whether the principles learned from lab studies extend to naturalistic search. In the present study, participants first explore a large (4,216 sq m) IKEA-like furniture store in virtual reality (VR) prior to a surprise visual search task in which they "shop" for items in the store. In the "shopping" task, participants sequentially searched for 10 pieces of furniture that were located in display rooms that contained objects of the same type (e.g., lamps, couches, beds). Fixations were measured via eye tracking. The primary question of interest is how target-matching category features (e.g., small, round, yellow, lamp) guide attention and looking during navigation towards, and then within, the target's display room. Results (N=42) showed that, prior to entering the target display room, fixations were distributed across objects with target-matching features. On the other hand, once the room with the target's display room was entered, search proceeded hierarchically, such that guidance was dominated first by a highly diagnostic feature (size or shape) followed by a secondary diagnostic feature, and then by color. Feature diagnosticity was dependent on the objects themselves and the organization of the room. Notably, in contrast to lab studies, color was not a prioritized feature for any target. These findings demonstrate how categorical and feature-based attentional guidance occur hierarchically to improve search efficiency across large-scale, naturalistic environments.

Talk 2, 5:30 pm, 25.22

DO ACTION VIDEO GAME PLAYERS SEARCH FASTER THAN NON-PLAYERS?

Zoe (Jing) Xu[1] (jingxu9@illinois.edu), Josef Spjut[2], Ben Boudaoud[2], Simona Buetti[1], Alejandro Lleras[1], Ruth Rosenholtz[2]; [1]University of Illinois, Urbana Champaign, [2]NVIDIA

Studies have shown that action video game players have enhanced visual abilities in various domains, such as multiple object tracking, size of the useful field of view, and visual search speed and accuracy. These improvements have been attributed to either a general advantage in "learning to learn" abilities, or domain-specific enhancement(s) in the "common demands" between specific games and experimental tasks. To investigate these two theories, we conducted six experiments examining whether and how players and non-players differ in various aspects of visual search. First, we used a staircase to determine the minimal display duration (Experiment1a) and target-distractor color difference (Experiment1b) required for participants to successfully identify a target in a color search task. Next, we assessed participants' search speed and the cost of switching target and distractor identities when there is one (Experiment2a) or multiple distractor types (Experiment2b). Finally, we measured search speed in harder T/L search (Experiment3a) and game-style figure search (Experiment3b). This study is the first to use both a staircase procedure and standard response time measures to discern differences between players and non-players in visual search. The results suggest that players search faster than non-players only in Experiment2, where performance degraded with increased distractor variability for non-players but not for players. Players also exhibited a smaller cost to switching the target and distractor identities. These findings imply that while there might be no overall enhancement in players' search abilities, they might benefit from holding variable distractor templates and switching their search target, potentially due to gaming experience which often necessitates memorizing and switching among multiple objects to monitor/avoid (as in first-person shooting games). These results support the "common demands" theory. In addition, our collected data on the specific games participants play allow for a more systematic evaluation of which games might enhance which search-related abilities.

TALK 3, 5:45 PM, 25.23

UH OH: DOES 40 YEARS OF VISUAL SEARCH RESEARCH ACTUALLY TELL US ABOUT VISUAL SEARCH IN THE WORLD?

Jeremy Wolfe[1,4] (jwolfe@bwh.harvard.edu), Eduard Objio, Jr[2], Hula Khalifa[3], Ava Mitra[1]; [1]Brigham and Women's Hospital, [2]Boston Latin School, Boston, MA, [3]UMass, Amherst, MA, [4]Harvard Medical School

We have decades of visual search data from experiments where observers look for targets among distractors. Typically, observers are tested in blocks of several hundred trials, and conclusions about underlying mechanisms are inferred from Reaction Time X Set Size functions and errors. The introductions to the subsequently published papers then declare that we are studying how you find your keys or the toaster in the real world. However, in the real world, you never search for your keys 100 times in a row. You search for keys, then a coat, then the doorknob, etc. Maybe the rules, gleaned from blocks of trials, apply only in the lab, with different rules for realistic mixtures of tasks? We used four feature search tasks (easy color, moderate lighting direction, moderate cube orientation, hard vernier offset). Observers completed 400 trials either in blocks of 100 trials or with all four tasks randomly intermixed. Mixing tasks did NOT destroy the standard patterns of RT or accuracy data. We obtained similar pattern of results when all four tasks had the same green O target but different distractors, ranging from easy (blue O) to harder (color x shape conjunction) to very hard (circle among vertical and horizontal ovals). Performance was similar under mixed and blocked conditions. Again, this is good news. The results suggest that rules, established in the lab, should apply in more realistic, mixed conditions. However, at least one important theoretical puzzle appears. Guided Search and other models have long proposed that target absent "quitting times" are established by an adaptive mechanism operating over multiple trials. Our experiments showed no evidence for adaptive learning in the mixed condition. Nevertheless, target-absent responses were not impaired. Observers did not need to learn when to quit. The implication is that standard accounts of search termination may be incorrect.

NEI EY017001, NSF 2146617, NCI CA207490

TALK 4, 6:00 PM, 25.24

SALIENT DISTRACTORS PRODUCE FEWER EYE MOVEMENTS WHEN TARGETS ARE ABSENT IN VISUAL SEARCH

Jeff Moher[1] (jmoher@conncoll.edu), Idah Mbithi[1], Maiko Okamura[1], Skylar Rice[1], Lily Rosan[1], Anjum Shaikh[2], Fabio Solorzano Quesada[1], Trafton Drew[3], Brennan Payne[4]; [1]Connecticut College, [2]University of Connecticut School of Medicine, [3]Sirona Medical, [4]University of Utah

Decades of research have demonstrated that salient distractors can capture attention and delay target acquisition. What effect does a salient distractor have when targets are sometimes absent? In Moher (2020), we found that in these types of visual searches, salient distractors cause observers to quit their search earlier than they otherwise would. As a result of this distractor-induced early quitting, targets are more likely to be missed. In a new pre-registered study, we examined eye movements in a simple visual search task to determine how distractors produce early quitting. In addition to replicating the behavioral results from Moher (2020), we found that when distractors were present, fewer eye movements occurred before target absent responses, and less of the display area was searched. These results suggest that at least part of the distractor-induced early quitting effect is a result of participants searching the display less thoroughly when distractors are present. Surprisingly, there were very few eye movements to the salient distractor itself across all conditions suggesting that distractor suppression may be involved in distractor-induced early quitting. Finally, salient distractors produced an increase in both search errors, in which the target was never fixated, and decision errors, in which the target was fixated but the participant failed to recognize it as a target. These findings demonstrate that salient distractors cause early quitting in visual search by reducing the amount of information that observers extract from the search image. Increases in decision errors, combined with below baseline fixations on the distractor itself, suggest that cognitive load associated with distractor suppression may also increase misses when distractors are present. These results have implications not only for understanding how distractors impact attention, but also for applied fields where visual search occurs in the presence of sometimes salient distractors, such as medical image screening.

This research was supported by NSF BCS-2218384 to JM

TALK 5, 6:15 PM, 25.25

INFORMATION TRANSFER DURING GOAL-DIRECTED VIEWING OF EVERYDAY SCENES

Katarzyna Jurewicz[1], Buxin Liao[1,2], B. Suresh Krishna[1]; [1]McGill University, [2]University of Electronic Science and Technology of China

One of the fundamental issues in visual perception is how and how much visual information is transferred across fixations. Here, we examine how humans actively scan the visual environment when performing goal-directed visual search on photographs of complex everyday scenes. We analyze data from two open datasets of eye-movements made by participants performing either category-search with 18 target categories (COCO-Search18) or free-viewing (COCO-FreeView) on over 4000 unique naturalistic images from the MS-COCO dataset. We focus specifically on the evidence for information transfer across saccades as revealed by saccades made after short inter-saccadic intervals (< 125 ms, short-latency saccades). When the target is present in the scene, participants (n = 10) fixate it after predominantly one saccade (45% of trials) or two saccades (36% of trials). Short-latency second saccades occur frequently (45% of second saccades on average) in goal-directed visual search. These saccades foveate the search target more often than saccades executed after longer intersaccadic intervals (regular-latency saccades). Short-latency second saccades are not small-amplitude corrective saccades: they are both more common and more likely to foveate the target when they follow first saccades that end further away from the target. Further, they are much more frequent during goal-directed visual search with the search target present than when the search target is absent or during free-viewing: active searching, and the top-down salience of the search-target contribute to increasing the frequency of short-latency saccades. The results show that human searchers use a satisficing strategy when actively searching complex everyday scenes for a categorically defined target. Short-latency saccades and information transfer across saccades work towards ensuring that the cost of making additional saccades to distractor stimuli is minimal; this would not be the case if perception began anew at each fixation. Information integration and transfer across saccades plays a prominent role during naturalistic vision.

Funding: BSK - NSERC Discovery Research program (RGPIN-2022-05399, DGECR-2022-00321), Calcul Quebec/Digital Research Alliance of Canada grant; KJ - IVADO Postdoctoral Research Funding (PostDoc-2021a-8859659558_2), Bourse d'excellence UNIQUE.

Talk 6, 6:30 pm, 25.26

ELECTROPHYSIOLOGICAL INDICES OF 'PROACTIVE' DISTRACTOR PROCESSING: CHARACTERIZING THE PD

Dirk van Moorselaar[1,2] (d.van.moorselaar@vu.nl), Jan Theeuwes[1,2,3]; [1]Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, the Netherlands, [2]Institute of Brain and Behaviour Amsterdam (iBBA), the Netherlands, [3]William James Center for Research, ISPA-Instituto Universitario, Lisbon, Portugal

Recent advancements in understanding attentional suppression through learning have sparked discussions on the underlying mechanisms. A key finding is that interference from salient distractors, like warning signs, significantly diminishes, often falling below baseline levels, when the color of the distractor is consistent and predictable. This finding has been pivotal in shaping the signal suppression framework, which posits that learning reduces specific feature gain, leading to proactive (feature) suppression. Research in this domain often relies on an early lateralized event-related potential known as the PD to investigate whether distractors can be filtered out before capturing attention (i.e., proactively). However, the interpretation of the PD as a marker of proactive suppression has faced challenges, with alternative accounts questioning its accuracy in representing proactive suppression. In a series of EEG experiments, we systematically varied elements of standard designs to explore proactive feature suppression. Across experiments participants performed a feature search variant of the singleton paradigm, wherein the target shape was embedded in a heterogeneous search display, while the distractor color was held constant. In Experiment 1, we observed reliable attenuation of the PD when target features were no longer consistent across trials, a design feature typical for these type of experiments. In Experiment 2 we introduced a sensory imbalance in a neutral, and task-irrelevant display preceding search. This irrelevant display elicited an early PD, even though there was no need for suppression. Experiment 3 demonstrated that color and rotating distractors could be ignored, yet the former elicited a PD, while the latter elicited an N2pc. Together these findings challenge the notion that the distractor PD unequivocally indexes suppression; it reflects the upweighting of target features and can be elicited by displays not requiring suppression. These findings challenge the concept of proactive feature suppression and emphasize the need for a more nuanced PD interpretation.

European Research Council (ERC) advanced grant (833029) to Jan Theeuwes

Talk Session: Sunday, May 19, 2024, 8:15 – 9:45 am, Talk Room 1

Motion

Moderator: Kohitij Kar, York University

Talk 1, 8:15 am, 31.11

OBJECT CORRESPONDENCE ACROSS MOVEMENTS AT SACCADIC SPEED

Melis Ince[1,2] (melis.ince@hu-berlin.de), Carolin Hübner[1,3], Martin Rolfs[1,2]; [1]Department of Psychology, Humboldt-Universität zu Berlin, Germany, [2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Germany, [3]Department of Psychology, Technische Universität Chemnitz, Germany

Saccadic eye movements impose rapid motion on the retinal image, raising the question of how object correspondence is established from one fixation to the next. Here, we investigated if the rapid motion itself — by providing spatiotemporal continuity — plays a role in achieving object correspondence. To isolate the contribution of high-speed motion, we simulated saccadic motion using a high-temporal-resolution projector (updating the display every 0.69 ms) while observers fixated their gaze throughout the experiments. We first investigated the contribution of motion at saccadic speed to object correspondence using a two-frame quartet-motion display. We positioned identical Gabor patches as objects at opposing corners within an imaginary rectangle. One object then moved continuously — along a curved trajectory (inward or outward) — to one of the neighboring corners, while the other jumped to the opposite side, completing the quartet. On each trial, participants first reported quartet rotations (clockwise or counterclockwise), indicating perceived object correspondence, and then traced the perceived continuous motion trajectory using a mouse, indicating motion visibility (location and curvature). We found that motion visibility declined as speed increased, eventually reaching chance levels for location and curvature reports. At the same time, continuous motion biased the quartet rotation perception even at the highest (saccade-like) speeds. These results suggest that high-speed motion informs object correspondence, even if that motion is effectively invisible. We are currently following up on this finding in a second study, in which we combine a version of our quartet motion display with the go/no-go reviewing paradigm (Sasi et al., 2023). We investigate if object files are maintained through motion at saccadic speed. By combining objective measures of stimulus visibility, the perception of object correspondence, and the maintenance of object files over time, we aim to shed light on the fundamental mechanisms behind object continuity at saccadic speeds.

Funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [865715 – VIS-A-VIS]) and the Heisenberg Programme of the Deutsche Forschungsgemeinschaft (grants RO3579/8-1 and RO3579/12-1) granted to MR.

TALK 2, 8:30 AM, 31.12

DETECTING MOVING OBJECTS DURING SELF-MOTION

Hope Lutwak[1], Bas Rokers[2], Eero Simoncelli[1,3]; [1]Center for Neural Science, New York University, [2]Psychology, Center for Brain and Health, Aspire Precision Medicine Research Institute, New York University Abu Dhabi, [3]Center for Computational Neuroscience, Flatiron Institute

As we move through the world, the pattern of light projected on our eyes is complex and dynamic, yet we are still able to distinguish moving and stationary objects. One might hypothesize that this is achieved by detecting discontinuities in the spatial pattern of velocities, however this computation is also sensitive to velocity discontinuities at boundaries of stationary objects. We instead propose that humans make use of the specific constraints that self-motion imposes on retinal velocities. When an eye translates and rotates within a rigid 3D world, the velocity at each location on the retina is constrained to a line segment in the 2D space of retinal velocities (Longuet, Higgins, Prazdny 1980). The slope and intercept of this segment is determined by the eye's translation and rotation, and the position along the segment is determined by depth of the scene. Since all possible velocities arising from a rigid world must lie on this segment, velocities not on the segment must correspond to moving objects. We hypothesize that humans make use of these constraints, by partially inferring self-motion based on the global pattern of retinal velocities, and using deviations of local velocity from the resulting constraint lines to detect moving objects. Using a head-mounted virtual reality device we simulated a translation forward in different virtual environments: one consisting of textured cubes above a textured ground plane, and one of scattered depth-matched dots. Participants had to determine if a cued cube/dot moved relative to the scene. Consistent with the hypothesis, we found that performance depended on the deviation of the object velocity from the constraint segment, not on the difference between retinal velocities of the object and its surround. Our findings contrast with previous inconclusive results, that relied on an impoverished stimulus with a limited field of view.

TALK 3, 8:45 AM, 31.13

OBJECT MOTION REPRESENTATION IN THE MACAQUE INFERIOR TEMPORAL CORTEX – A GATEWAY TO UNDERSTANDING THE BRAIN'S INTUITIVE PHYSICS ENGINE

Hamidreza Ramezanpour[1] (hamidram@yorku.ca), Filip Ilic, Richard Wildes, Kohitij Kar; [1]York University

Over the past decade, there have been significant advances in understanding how primates recognize objects in the presence of identity-preserving variations. However, primates' vision encompasses more than object recognition. In the dynamic world, an effective interaction with moving objects and the ability to infer and predict their motion are essential for survival. In this study, we systematically investigated hierarchically connected brain areas in the ventral visual pathway of rhesus macaques (areas V4 and IT), implicated in object recognition, to first characterize their responses to object motion, speed, and direction. Subsequently, we quantified the correlative links between these responses and two distinct object motion-based behaviors, one reliant on information directly available in videos (e.g., velocity discrimination) and the other predicated on predictive motion estimates from videos (e.g., future frame predictions). Further, by employing causal microsimulation strategies, we tested the critical role of the macaque IT cortex in these behaviors. Interestingly, while current computational models of object and action recognition are accurate on stationary object-based tasks, we observed that their predictions suffer significant deficits in our dynamic tasks compared to primates. These findings call into question the widely accepted demarcation of the primate ventral and dorsal cortices into the "what" and "where" pathways. These explorations highlight the imperative to examine the interplay between these cortical hierarchies for a more profound understanding of visual motion perception, which serves as a gateway to intuitive physics. The data also provide valuable empirical constraints to guide the next generation of dynamic brain models.

CIHR, Canada Research Chair Program, Google Research, CFREF, Brain Canada, SFARI

TALK 4, 9:00 AM, 31.14

ACQUISITION OF SECOND-ORDER MOTION PERCEPTION BY LEARNING TO RECOGNIZE THE MOTION OF OBJECTS MADE BY NON-DIFFUSIVE MATERIALS

Zitang Sun[1] (sun.zitang.73u@st.kyoto-u.ac.jp), Yen-Ju Chen[1], Yung-Hao Yang[1], Shin'ya Nishida[1,2]; [1]Cognitive Informatics Lab, Graduate School of Informatics, Kyoto University, Japan, [2]NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, Japan

Many animals, including flies, macaques, and humans, have an ability to visually recognize image motion not only from shifts of spatial patterns defined by luminance modulations (first-order motion) but also from those defined by high-level image features such as temporal modulations and contrast modulations (second-order motion). In the past, second-order motion perception has been extensively studied using carefully designed artificial stimuli (e.g., drift-balanced motion) to control first-order motion components, but why and how the visual system has acquired this perceptual ability in natural environments remains poorly understood. We hypothesized that the biological system might naturally learn second-order motion perception for the purpose of estimating correct physical object motion amidst internal optical fluctuations produced, for example, by highlights of glossy materials and refractions of transparent materials. As a proof concept, we developed a DNN-based model to process both first- and second-order motions in natural scenes. The model was based on our two-stage model (Sun et al., NeurIPS 2023) consisting of a trainable motion energy sensing and a recurrent self-attention network, each inspired by biological computations in V1 and MT. For preprocessing for complex second-order features, we added a second input pathway with a vanilla multi-layered convolution network. The model was trained on two distinct optical flow datasets generated by rendering random object motion: one with purely diffuse reflection (PD) and the other with non-diffuse (ND) material properties, the latter including ample optical turbulence made by specular reflections and transparent refractions. The ND-trained model demonstrated significantly better recognition of various types of second-order motion, aligning closely with human performance measured in our psychophysical experiments. Also, this performance was unachievable without the

second input pathway. The results suggest that second-order motion perception might have evolved, at least partially, to help robust estimation of object motion while countering optical fluctuations under natural environments.

This work is supported by JST JPMJFS2123, MEXT/JSPS JP20H00603 and JP20H05605.

Talk 5, 9:15 AM, 31.15

DEEP FEATURE MATCHING VS SPATIO-TEMPORAL ENERGY FILTERING FOR ROBUST MOVING OBJECT SEGMENTATION

Matthias Tangemann[1] (matthias.tangemann@bethgelab.org), Matthias Kümmerer[1], Matthias Bethge[1]; [1]University of Tübingen, Tübingen AI Center

Recent methods for optical flow estimation achieve remarkable precision and are successfully applied in downstream tasks such as segmenting moving objects. These methods are based on matching deep neural network features across successive video frames. For humans, in contrast, the dominant motion estimation mechanism is believed to rely on spatio-temporal energy filtering. Here, we compare both motion estimation approaches for segregating a moving object from a moving background. We render synthetic videos based on scanned 3d objects and backgrounds to obtain ground truth motion for realistic scenes. We transform the videos by replacing the textures with random dots that follow the motion of the original video. This way, each individual frame does not contain any other information about the object apart from the motion signal. Humans have been shown to be able to use random dot motion for recognizing objects in these stimuli (Robert et al. 2023). We compare segmentation methods based on the recent RAFT optical flow estimator (Teed and Deng 2020) and the spatio-temporal energy model of Simoncelli & Heeger (1998). Our results show that the spatio-temporal energy approach works almost as well as using RAFT for the original videos when combined with an established segmentation architecture. Furthermore, we quantify the amount of segmentation information that can be decoded from both models when using the optimal non-negative superposition of feature maps for each video. This analysis confirms that both optic flow representations can be used for motion segmentation while RAFT performs slightly better for the original videos. For the random dot stimuli however, hardly any information about the object can be decoded from RAFT while the brain-inspired spatio-temporal energy filtering approach is only mildly affected. Based on these results we explore the use of spatio-temporal filtering for building a more robust model for moving object segmentation.

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): Germany's Excellence Strategy – EXC 2064/1 – 390727645 and SFB 1233, TP4, project number: 276693517.

Talk 6, 9:30 AM, 31.16

ANISOTROPY IN PERCEIVED NONRIGIDITY

Akihito Maruya[1] (user3098@sunyopt.edu), Qasim Zaidi[1]; [1]State University of New York, College of Optometry

Shape-from-motion models generally assume that objects are rigid, which simplifies the computations but cannot handle movements and locomotion of organisms, all of which require nonrigid shape deformations. We have demonstrated that rotated rigid objects can appear strikingly nonrigid, depending on speed and shape. We showed that nonrigid percepts arise from the outputs of direction-selective motion cells and are countered by feature-tracking and shape-based priors. Now we present the surprising finding that perceived nonrigidity changes with the rigid object's orientation, and model it with documented cortical anisotropies. When two solid 3D circular rings attached rigidly at an angle are rotated horizontally around a vertical axis at medium speed, observers see either rigid rotation or non-rigid wobbling. A 90° image rotation markedly enhances the non-rigid percept. We observed that the elliptical projections of the rings in the rotated image appear narrower and longer than in the original image, like the increased perceived height versus width when a square is rotated 45° to form a diamond. We successfully model the perceived changes in shape with optimal Bayesian decoding of V1 outputs, by incorporating anisotropies in the number and tuning-widths of orientation selective cells and the probability distribution of orientations in images of natural scenes. We show quantitatively that elongating the ellipses alone leads to more perceived nonrigidity even for horizontal rotation, but the vertical rotation further enhances nonrigidity. We incorporated the cortical anisotropies into motion flow computations. The estimated motion fields were decomposed into gradients of divergence, curl, and deformation and compared to the gradients for physical rotation and wobbling. The gradients for the vertical rotation were a closer match to physical wobbling, while the gradients for the horizontal rotation were in between physical wobbling and rotating. This asymmetry indicates that hardwired cortical anisotropies can explain changes in perceived non-rigidity with motion axis.

Talk Session: Sunday, May 19, 2024, 8:15 – 9:45 AM, Talk Room 2

Scene Perception: Behaviour, psychophysics

Moderator: Michelle Green, Barnard College

Talk 1, 8:15 AM, 31.21

BEYOND WORDS: RAPID SCENE DETECTION IS FACILITATED BY HIGH SEMANTIC COMPLEXITY

Emily Lo[1], Kaiki Chiu[1], Quinn O'Connor[1], Michelle R. Greene[1]; [1]Barnard College

The adage "a picture is worth a thousand words" underscores the notion that visual information conveys rich meaning. However, not all scenes contain equal semantic depth. This study quantified the semantic complexity of images and assessed its implications for early visual processing. We asked 100 online observers to write image descriptions for a previously-used set of 1000 images (Bainbridge & Baker, 2020; Greene & Trivedi, 2023). A composite semantic complexity score was computed from the median word count, variability among descriptions (entropy in a bag of words model), and average pairwise distance between concepts within a description from a word vector model (Word2Vec). We selected 100 images with the highest semantic complexity scores and 100 images with the lowest semantic complexity for a rapid detection experiment. The two image groups did not differ significantly in several measures of visual complexity. We predicted that images with lower semantic complexity convey less information, thus observers would more quickly and accurately detect such images. Observers (N=38) distinguished between scene images and 1/f noise (SOA: ~60 ms, with a dynamic pattern mask. Contrary to our expectations, observers had higher detection sensitivity for images with greater lexical complexity (d': 3.91 vs 3.58, p<0.005). This finding challenges the common expectation of capacity limitations in the face of stimulus complexity. Instead, it suggests that semantic richness may enhance rapid perception. One interpretation is that a more extensive set of contextual associations increases both semantic complexity and visual detectability. Alternatively, richer semantic content may engage top-down processing more effectively, aiding rapid visual detection. These results challenge typical views of cognitive load and point to highly semantic aspects of scene gist that drive early visual detection.

NSF CAREER 2240815 to MRG

TALK 2, 8:30 AM, 31.22

THE PSYCHOPHYSICS OF COMPOSITIONALITY: RELATIONAL SCENE PERCEPTION OCCURS IN A CANONICAL ORDER

Alon Hafri[1] (alon@udel.edu), Zekun Sun[2], Chaz Firestone[3]; [1]University of Delaware, [2]Yale University, [3]Johns Hopkins University

An intriguing proposal in recent literature is that vision is compositional: Just as individual words combine into larger linguistic structures (as when "vase," "table," and "on" compose into the phrase "the vase on the table"), many visual representations contain discrete constituents that combine in systematic ways (as when we perceive a vase on a table in terms of the vase, the table, and the relation physical-support). This raises a question: What principles guide the compositional process? In particular, how are such representations composed in time? Here we explore the psychophysics of scene composition, using spatial relations as a case study. Inspired by insights from psycholinguistics, we test the intriguing hypothesis that the mind builds relational representations in a canonical order, such that 'reference' objects (those that are large, stable, and/or exert physical 'control'; e.g., tables)—rather than 'figure' objects (e.g., vases resting atop them)—take precedence in forming relational representations. In Experiment 1, participants performed a 'manual construction' task, positioning items to compose scenes from sentences (e.g., "the vase is on the table"). As hypothesized, participants placed reference-objects first (e.g., table, then vase). Next, we explored whether this pattern arises in visual processing itself. In Experiment 2, participants were faster to recognize a target scene specified by a sentence when the reference-object (table) appeared before the figure-object (vase) than vice-versa. Notably, this pattern arose regardless of word order (reference- or figure-first) and generalized to different objects and relations. Follow-ups showed that this effect emerges rapidly (within 100ms; Experiment 3), persists in a purely visual task (Experiment 4), and cannot be explained by size or shape differences between objects (Experiment 5). Our findings reveal psychophysical principles underlying visual compositionality: the mind builds relational representations in a canonical order, respecting each element's role in the relation.

NSF BCS #2021053 awarded to C.F.

TALK 3, 8:45 AM, 31.23

THE ROLE OF OBJECT CO-OCCURRENCE IN ATTENTIONAL GUIDANCE: EVIDENCE FROM EYE-MOVEMENTS

Alexadra Theodorou[1] (atheodorou@ucdavis.edu), John Henderson[1]; [1]University of California, Davis

The visual world is complex, yet visual information processing is effortless. During scene viewing semantically related objects are prioritized for attention (Hayes & Henderson, 2021). Previous work has defined semantic relations relevant for gaze guidance based on models from computational linguistics. Here we aim to extend previous findings by investigating relationships between objects derived from their visual scene contexts. Neuroimaging and behavioral data have shown that objects that tend to co-occur in scenes are closely represented in the aPPA while frequently co-occurring objects receive higher similarity judgements in a behavioral task (Bonner & Epstein, 2020; Magri, Elmoznino & Bonner, 2023). Here, we investigate measures of object-object relations derived from their visual co-occurrence statistics in scenes to predict eye-movement behavior. Eye-movement data was collected from 100 participants who each viewed 100 scenes performing a free-viewing task. Using object label embeddings from the object2vec model (Bonner & Epstein, 2020) we constructed map-level representations that encode similarity between objects based on their likelihood to appear within the same scene. We used generalized mixed effects models to estimate gaze behavior as a function of co-occurrence values. Our results suggest that objects that are more highly related to other objects within a scene as a function of their co-occurrence likelihood are more likely to be fixated. These findings underscore the role of statistical regularities, particularly in the form of co-occurrence statistics within visual contexts, in shaping efficient eye-movement behavior. Consequently, our study suggests that object co-occurrence forms an integral part of the semantic representations guiding eye movements, contributing significantly to our understanding of object representational dimensions in scene exploration.

TALK 4, 9:00 AM, 31.24

EFFICIENT CODING OF ENSEMBLE STIMULI RELATIVE TO A DYNAMIC REFERENCE

Long Ni[1], Alan A. Stocker[1]; [1]The University of Pennsylvania

When discriminating the average of a stimulus ensemble against a reference, observers often overweigh those stimuli in the ensemble that have feature values similar to the reference—a behavior known as 'robust averaging'. We previously proposed that this behavior can be explained by a Bayesian decision model constrained by efficient coding. Assuming our visual system rapidly forms efficient representations of ensemble stimuli relative to a dynamic reference, our model captured multiple existing datasets showing robust averaging of low-level stimulus ensembles. Here, we provide further evidence for two key predictions of the model: robust averaging should 1) become progressively more pronounced the longer the visual system is exposed to the ensemble stimuli statistics and 2) be reduced when the distribution of the ensemble stimuli is uniform. To test the first prediction, we had subjects discriminate the average orientation of 12 gratings displayed on a virtual circle against a central reference grating during three sessions. In every trial, ensemble orientations were drawn from a Gaussian distribution with various means relative to the (variable) reference orientation, overall creating an approximately Gaussian distribution of ensemble orientations around the reference. Across the three sessions, subjects' discrimination accuracy continuously improved and the weighting kernel became increasingly non-uniform, attributed by our model to a reduction in internal noise and a progressively better adaptation to the ensemble statistics. We tested the second prediction by sampling orientations from two oppositely 'skewed' linear distributions, resulting in an overall uniform distribution centered at the reference. Subjects completed three sessions each under both Gaussian and uniform conditions. While accuracy was similar in both, robust averaging was largely absent in the uniform condition. The alignment between our model's predictions and empirical data validates our hypothesis that the visual system can dynamically create efficient sensory representations of ensemble stimuli relative to a trial-by-trial varying reference.

Talk 5, 9:15 am, 31.25

MAPPING A SCENE FROM AFAR: ALLOCENTRIC REPRESENTATION OF LOCATIONS IN SCENE-SPACE

Anna Shafer-Skelton[1] (ashafers@sas.upenn.edu), Russell Epstein[1]; [1]University of Pennsylvania

Spatial neuroscience has discovered a great deal about how animals—primarily rodents—encode allocentric (world-centered) cognitive maps. We hypothesized that humans might be able to form such maps from afar, through visual processing alone. Previous work in vision science has explored how we extract the overall shape of scenes from particular points of view, but little is known about how we form allocentric representations of discrete locations within a scene—a key feature of a cognitive map. We tested for such a representation in two behavioral experiments. In Exp. 1, N=30 participants viewed images of a 3D-rendered courtyard, taken from one of 4 possible viewpoints outside and slightly above the courtyard, spaced 90 degrees apart. On each trial, participants saw two courtyard images separated by a brief (500ms) delay. Within each image was an indicator object (a car), in one of six possible allocentric locations; participants reported whether the indicator object was facing the same or different allocentric direction in the two images. The task was designed to direct attention to the location of the indicator object within the allocentric framework of the courtyard without requiring explicit reporting of that location. We observed a significant performance benefit in across-viewpoint trials when the indicator object was in the same allocentric location in both images compared to when it was in different allocentric locations (BIS p=0.009; we also report d-prime: p=0.023, RT: p=0.062). In Exp. 2 (N=30), we replicated this same-location benefit when participants viewed a continuous stream of courtyard images and performed a 1-back task on the facing direction of the indicator object (BIS p=0.004; secondary measures d-prime: p=0.026, RT: 0.023). These results show evidence for an allocentric representation of within-scene locations—a critical ingredient of allocentric cognitive maps—formed via visual exploration, without traversing the space.

This work was supported by a NIH-NEI grant awarded to RAE (R01-EY022350)

Talk 6, 9:30 am, 31.26

AUTOMATIC LOGICAL INFERENCES IN VISUAL SCENE PROCESSING

Nathaniel Braswell[1] (nathaniel.braswell@yale.edu), Chaz Firestone[2], Nicolò Cesana-Arlotti[1]; [1]Yale University, [2]Johns Hopkins University

The human capacity for logic is responsible for some of our grandest achievements; without it, formal mathematics, economic systems, and architectural marvels would be elusive. Yet logical cognition is not limited to rarefied intellectual challenges—it also arises in everyday contexts, such as inferring that a glass on a table must be yours because your friend is holding theirs. Previous work shows that a primitive logical operation—disjunctive syllogism (p OR q; NOT p; therefore, Q)—is deployed by infants to infer the identities of objects (Cesana-Arlotti et al., 2018). This raises an intriguing question: Do such logical inferences arise automatically in adults, and even impact processing of visual scenes? Experiment 1 showed adults events wherein an ambiguous object was 'scooped' by a cup from a two-item set (snake and ball). Upon seeing one of the objects outside the cup (snake), adults responded slower when the revealed object's identity violated their logical prediction (snake) than when it was consistent (ball). The effect persisted over 40 trials, even though the revealed identity was random—suggesting that adults were executing this inference automatically. Put differently, they 'couldn't help' but infer the hidden object's identity, even when they knew they shouldn't. Experiment 2 tested whether this effect resulted from one item's appearance priming the other. We devised scenes with a third item in the cup, preventing logical inferences about the cup's contents. A Bayes Factor analysis found strong evidence for the null hypothesis of no response time differences, confirming that logical inference drives the Experiment 1 effect. These findings open avenues in both logical cognition and scene processing. First, our results suggest that logical inferences may be spontaneously deployed to resolve visually uncertain events. Additionally, methods from vision science may serve as a previously unexplored tool for uncovering the nature of our mind's fundamental logical capacities.

NSF BCS #2021053 awarded to C.F.

Talk Session: Sunday, May 19, 2024, 10:45 am – 12:30 pm, Talk Room 1

Visual Memory: Working and behavior

Moderator: Sven Ohl, Humboldt-Universität zu Berlin

Talk 1, 10:45 am, 32.11

SIGNAL INTRUSION RECONCILES DIVERGENT EFFECTS OF PERCEPTUAL DISTRACTION ON WORKING MEMORY

Ziyao Zhang[1] (ziyaopsy@gmail.com), Jarrod Lewis-Peacock[1]; [1]The University of Texas at Austin

Perceptual distraction distorts visual working memories. Recent research has shown divergent effects of distraction on memory performance, including attractive or repulsive biases in memory reports, improving or impairing memory precision, and increasing or decreasing guess rates. These effects are sensitive to target-distractor similarity and thus have been attributed to sensory interference according to the sensory recruitment hypothesis of working memory. Here, we propose a novel Distractor Intrusion Model (DIM), an extension of the Target Confusability Competition (TCC) framework, to reconcile the discrepant results of perceptual distraction. We hypothesized that sensory interference, in all instances, is driven by the integration of a target memory signal and an intrusive distractor signal. We tested this model against the classical mixture model and other candidate models. Model comparisons showed that TCC-DIM had a superior fit to memory error distributions across six delay-estimation tasks with distraction (N = 220). Both passive and active distraction tasks were examined and target-distractor similarity was varied between 18° and 153°. According to the model, distractor intrusions decreased along with target-distractor similarity, in accordance with the sensory recruitment hypothesis. Moreover, we found that TCC-DIM successfully replicated divergent effects of distraction on memory bias, precision, and guesses using only this one intrusion mechanism. This model also makes a novel, and somewhat surprising, prediction that low-fidelity memories are likely to benefit from distractor intrusions, whereas high-fidelity memories are likely to become impaired. Our data support this prediction such that participants (N=49) with lower memory precision benefited from distraction and showed a reduction in memory errors relative to no-distraction trials. Those with higher memory precision showed greater errors following distraction. These results collectively suggest that perceptual distractors affect working memories through signal intrusions, thus providing a unified mechanism to explain diverse and divergent effects of distraction on working memory performance.

This work was completed with support from the National Institute of Health Grant R01EY028746 awarded to J.A.L.-P

Talk 2, 11:00 am, 32.12

ACTION PLANNING BIASES INTERACTIONS BETWEEN VISUAL WORKING MEMORY REPRESENTATIONS

Caterina Trentin[1] (c.trentin@vu.nl), Christian N.L. Olivers[1], Heleen A. Slagter[1]; [1]Vrije Universiteit Amsterdam

Recent studies suggest that planning an action on an object in visual working memory (VWM) can modulate its sensory representations. In this study, we investigated how planning an action on objects in VWM influences the way in which VWM representations interact – specifically whether different associated action plans also lead to more differentiated mnemonic representations of sensory input. We hypothesized that associating two visual orientations with different action plans in VWM would make them appear more dissimilar in memory than two orientations linked to the same action plan. Participants (n=32) memorized the orientation of two bars, sequentially presented on a touch screen. Following a delay, they manually reproduced each of the orientations. Each bar was followed by an action cue informing participants which action had to be performed at test to reproduce the memorized orientations. In the different action condition, the bars were associated with different action plans, i.e., a grip and a slide action. In the same action condition, they were linked to the same action plan, namely both grip or both slide actions. Our results show that similarly oriented bars repelled each other in both conditions (the well-known repulsion effect), but more so when associated with different action plans. Preliminary results from a control experiment indicate that the observed repulsion effect cannot be explained by differential motor biases, but is driven by action planning-induced changes in the mnemonic representations themselves (i.e., is perceptual in nature). Thus, not only visual features, but also action attributes modulate the way VWM representations interact: planned actions on objects in VWM can influence the extent to which their VWM representations look more or less similar to our mind's eyes.

Talk 3, 11:15 am, 32.13

LATENT MEMORY TRACES FOR PROSPECTIVE ITEMS IN VISUAL WORKING MEMORY

Luzi Xu[1] (l.xu2@uu.nl), Andre Sahakian[1], Surya Gayet[1], Chris Paffen[1], Stefan Van der Stigchel[1]; [1]Utrecht University

Visual working memory is a capacity-limited cognitive system that allows for keeping task-relevant information available for goal-directed actions. When selecting a subset of items for encoding in working memory (e.g., pears, pasta, and yogurt from a shopping list), observers can be simultaneously exposed to other items (e.g., tomatoes and eggs, on the same list) that are not selected for imminent action (hereafter: 'prospective items'). Here, we asked whether prior exposure to such prospective items facilitates subsequent visual working memory encoding of these items, when they are selected for imminent action later. We used a so-called 'copy task', in which participants reproduced an arrangement of colored polygons (the 'model grid'), in an adjacent empty grid. During placement, prospective items (i.e., hitherto unplaced items) in the model grid either remained at a fixed position or were swapped. The latter condition hampered the buildup of memory traces for prospective items. In three experiments, using different approaches to manipulate the stability of prospective items, we consistently observed that - when prospective items remained stable - participants took less time inspecting the model when encoding these items in a later stage (compared to when they were swapped). This reduced inspection duration was not accompanied by a higher number of inspections or an increase in errors. We conclude that the memory system gradually builds up latent memory traces of items that are not selected for imminent action, thus increasing the efficiency of subsequent visual working memory encoding. The present work reveals one way in which the mnemonic system circumvents its capacity limitations to efficiently operate in a complex visual world.

Talk 4, 11:30 am, 32.14

STORAGE IN WORKING MEMORY RECRUITS A MODALITY-INDEPENDENT POINTER SYSTEM

Henry Jones[1,2] (henryjones@uchicago.edu), Darius Suplica[1], William Thyer[1,2], Edward Awh[1,2]; [1]Department of Psychology, University of Chicago, [2]Institute for Mind and Biology, University of Chicago

Prominent theories of working memory (WM) have proposed that distinct working memory systems may support the storage of different types of information. For example, distinct dorsal and ventral stream brain regions are activated during the storage of spatial and object information in visual WM. Although feature-specific activity is likely critical to WM storage, we hypothesize that a content-independent indexing process may also play a role. Specifically, spatiotemporal pointers may be required for the sustained indexing and tracking of items in space and time, even while features change, within an unfolding event. Past evidence for such a content-independent pointer operation includes the finding that signals tracking the number of individuated representations in WM (load) generalize across colors, orientations and conjunctions of those features. However, overlapping orientation and color codes in early visual cortices may mimic a generalizable signal. Here, we provide a stronger demonstration of content-independence by using pairs of features that are as cortically disparate as possible. Study 1 (n=16) used color and motion coherence stimuli, and showed that load decoding models generalized across these disparate features. In addition, we used representational similarity analysis (RSA) to document "pure" load signals that tracked the number of items stored regardless of attended feature, while simultaneously documenting and controlling for feature-specific neural activity. Extending these observations, in Study 2 (n=24; n=16) we applied similar analytic approaches to demonstrate a common load signature between auditory and visual sensory modalities, while controlling for modality-specific neural activity and the spatial extent of covert attention. Our findings suggest that content-independent pointers may play a fundamental role in the storage of information in working memory, and may contribute to its overall limited capacity.

TALK 5, 11:45 AM, 32.15

THE RELATIVE DOMINANCE OF VISUAL AND SEMANTIC INFORMATION WHEN VISUAL STIMULI ARE RETRIEVED FROM MEMORY BASED ON IMAGES OR WORDS

Adva Shoham[1] (advashoham@mail.tau.ac.il), Itay Yaron[1], Liad Mudrik[1], Galit Yovel[1]; [1]Tel Aviv University

Familiar concepts can be described by their visual and semantic features. These types of information are hard to dissociate in mental representations. In a recent study we used visual and language DNNs to disentangle and quantify the unique contributions of visual and semantic information in human mental representations of familiar stimuli. We revealed a larger contribution of visual than semantic information during stimuli presentation in perception but a reversed pattern when recalled from memory based on their names. Here we adopt the same methodology to ask how long after stimulus offset does visual dominance shifts to semantic dominance. The duration for which visual information is retained following stimulus offset has been debated. To that end, across two studies, we manipulated the delay between stimulus offset and its recall from memory. In Study 1, participants rated the visual similarity of pairs of familiar faces in simultaneous presentation and in a sequential presentation with 2sec, 5sec or 10sec delays. We extracted representations of faces from a face-trained DNN, and of their Wikipedia description from a language model. In Study 2, we used data collected by Bainbridge et al (2019), in which participants were presented with an image of a scene and were asked to copy it while looking at the scene, 1 second or 10 mins after it was removed; or draw the scene based on its name with no prior exposure. We extracted representations of drawings from an object-trained DNN fine-tuned for drawings, and of their Wikipedia description from a language model. Both experiments revealed visual dominance after stimulus offset across all delays, and semantic dominance when retrieved from memory based on names. We conclude that visual information is dominant even 10 minutes after visual stimulus offset. Semantic information dominates the representation when a stimulus is recalled based on verbal information

ISF 917/21

TALK 6, 12:00 PM, 32.16

FURTHER EVIDENCE THAT THE SPEED OF WORKING MEMORY CONSOLIDATION IS A STRUCTURAL LIMIT

Benjamin J. Tamber-Rosenau[1] (bjtrbjtr@gmail.com), Lindsay A. Santacroce[1,2], Brandon J. Carlos[1,3]; [1]University of Houston, [2]Toronto Metropolitan University, [3]Ball State University

It has been proposed that the typically slow consolidation of information from vision to working memory (WM) is under flexible control, and thus can be speeded based on task demands. Recently (Carlos et al., 2023, doi: 10.3758/s13414-023-02757-7), we showed that consolidation is not sped even when it is prioritized over a subsequent competing decision task (T2). However, other research (Nieuwenstein et al., 2015, doi: 10.1167/15.12.739; Woytaszek, 2020) has manipulated the proportion of trials with T2 present and suggested that anticipated interference from competing tasks can lead to speeding of consolidation. Here, we present evidence against speeding of consolidation even when interference can be anticipated, providing an additional line of evidence against flexible control of WM consolidation. Using a within-subjects manipulation, participants completed blocks of a WM task with T2 presented at varying delays from the WM sample, on either 50% or 100% of trials. Retroactive interference from T2 onto WM was similar regardless of block (i.e., T2 probability). In another manipulation, we also varied the delay from T2 response to WM probe and found that this second delay's duration had no effect on WM reports. Importantly, this suggests that changes in WM performance with sample-T2 delay measure only the interruption of WM consolidation and are not contaminated by proactive interference from T2 onto the report of information from WM. In sum, the present results are consistent with the transfer of information from vision to WM being a slow process that is not under flexible control—either from explicit volitional prioritization, or implicit demands to counter anticipated interference.

This material was supported by the United States National Science Foundation under grant number 2127822.

TALK 7, 12:15 PM, 32.17

PROBING BIDIRECTIONAL SERIAL DEPENDENCE IN AN N-BACK ORIENTATION ESTIMATION TASK

Jongmin Moon[1], Hoyeon Yoon[1], Oh-Sang Kwon[1]; [1]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology

Vision is continuously shaped by a phenomenon known as serial dependence, wherein the estimation of stimulus features, such as orientation, is systematically biased by past visual input. This bias is believed to leverage the temporal autocorrelation in visual scenes, enhancing perceptual stability and sensitivity to change. To harness the full potential of temporal continuity, observers should consider not only the preceding stimulus but also the following one, when estimating a remembered stimulus feature embedded in a sequence of stimuli. Here, we used an N-back orientation estimation task to investigate whether serial dependence extends to memorized stimuli, with the preceding and/or following stimuli inducing the effect. Subjects were presented with a sequence of randomly oriented Gabor stimuli. The sequence terminated with a constant hazard rate, prompting subjects to recall the orientation of the 1-back stimulus (i.e., the target). Therefore, subjects had to keep in mind both the target and the following stimuli when prompted to recall the target. A probabilistic mixture model was employed to quantify contributions of different sources of error, excluding trials where subjects mistakenly reported the preceding or following stimulus instead of the target. Results revealed a highly consistent pattern of repulsive bias in the forward direction (preceding stimulus biases target estimation) and a weak trend of repulsive bias in the backward direction (following stimulus biases target estimation). Intriguingly, the strength of the repulsive bias was more pronounced for the preceding stimulus, despite the more recent presentation of the following stimulus, which would intuitively be expected to have a stronger working memory trace. These results underscore that our memory of visual scenes is influenced by both preceding and following stimuli, with the bias in forward direction prevailing in bidirectional serial dependence. Overall, our findings contribute to a deeper understanding of mechanisms underlying serial dependence in visual working memory.

This research was supported by the National Research Foundation of Korea (NRF-2020S1A3A2A02097375).

Talk Session: Sunday, May 19, 2024, 10:45 am – 12:30 pm, Talk Room 2

Spatial Vision

Moderator: Michele Rucci, University of Rochester

Talk 1, 10:45 am, 32.21

GEOMETRY OF ANISOTROPIC CONTEXTUAL INTERACTIONS IN THE VISUAL CORTEX PLACES FUNDAMENTAL LIMITS ON SPATIAL VISION.

Mitchell Morton[1,2] (mitchell.morton@yale.edu), Sachira Denagamage[1,2], Nyomi Hudson[1], Anirvan Nandy[1,2,3,4]; [1]Department of Neuroscience, Yale University, New Haven, CT 06510, [2]Interdepartmental Neuroscience Program, Yale University, New Haven, CT 06510, [3]Wu Tsai Institute, Yale University, New Haven, CT 06511, [4]Kavli Institute for Neuroscience, Yale University, New Haven, CT 06511

Crowding, the impaired ability to accurately recognize a target stimulus among distractors, is a major bottleneck in visual perception. The spatial configuration of distractors in relation to the target profoundly influences perceptual fidelity. Notably, when a distractor is placed at a more eccentric point on the radial axis (termed ‘radial-out crowding’), it exerts the strongest impairment. Despite the pronounced perceptual anisotropy, the prevalent assumption underlying our understanding of contextual interactions in the visual cortex assumes isotropy. We investigated how distractor stimuli in different spatial configurations impacted the representation of a target stimulus in laminar microcircuits in the primary visual cortex (V1). Our study reveals that radial-out crowding more strongly impacts the ability to decode the target orientation from V1 population activity compared to other spatial configurations. This effect was strongest among putative excitatory neurons in the superficial and input layers, which are the primary neural populations involved in feed-forward information propagation. Remarkably, the feedback pathway involving the deep cortical layers does not exhibit anisotropy. Mechanistically, the anisotropy is explained by a tuned suppression and untuned facilitation of orientation responses, leading to an anisotropic broadening of tuning curves in the feedforward pathway, but not in the feedback pathway. These results underscore the non-uniform spatial integration of information by neurons in the visual cortex, establishing the presence of anisotropic contextual interactions in the earliest stages of cortical processing. By elucidating the distinct roles of feed-forward and feedback pathways in the context of crowding, this study advances our understanding of the intricate interplay between spatial arrangement, neural circuitry, and the constraints on perceptual fidelity during early visual processing.

NIH/NEI R01 EY032555, NARSAD Young Investigator Grant, Ziegler Foundation Grant, Yale Orthwein Scholar Funds, NIH/NINDS training grants T32-NS007224, T32-NS041228, and NIH/NEI core grant for vision research P30 EY026878

Talk 2, 11:00 am, 32.22

SPATIAL CONFIGURATION OF CONTEXTUAL STIMULI INFLUENCES INTER-LAMINAR INTERACTIONS IN MACAQUE PRIMARY VISUAL CORTEX

Xize Xu[1,2,4] (xize.xu@yale.edu), Mitchell P. Morton[1,3], Nyomi V. Hudson[1], Anirvan S. Nandy[1,3,4,5], Monika P. Jadi[1,2,3,5]; [1]Department of Neuroscience, Yale University, New Haven, CT 06510, [2]Department of Psychiatry, Yale University, New Haven, CT 06510, [3]Interdepartmental Neuroscience Program, Yale University, New Haven, CT 06510, [4]Kavli Institute for Neuroscience, Yale University, New Haven, CT 06511, [5]Wu Tsai Institute, Yale University, New Haven, CT 06511

Our visual experience is a result of the concerted activity of neuronal ensembles in the sensory hierarchy. Yet how the spatial organization of objects influences neural activity in this hierarchy remains poorly understood. We investigate how the inter-laminar interactions in the primary visual cortex (V1) are affected by visual stimuli in isolation or with flanking stimuli at various spatial configurations that are known to exert a “crowding” effect on perception. Visual crowding is thought to be the primary limitation on object perception in peripheral vision, and the psychophysically identified “crowding zone” of impaired object identification is highly non-uniform. By employing dimensionality reduction approaches to simultaneous layer-specific population recordings, we determined the extent to which trial-to-trial fluctuations of population responses in the superficial layers of V1 are related to those in the input layer. We demonstrate that specific spatial configurations of contextual stimuli differentially modulate inter-laminar interactions by changing their fidelity and the balance between feedforward and feedback signaling, but not their structure.

Remarkably, the modulations mirror the spatially non-uniform aspects of perceptual crowding. Our results suggest a model in which perceptual impairment under crowding is mediated by visual context integration in the superficial layers of V1 and posit that the non-uniformity in contextual inputs is the neural substrate of perceptual experience.

R01 EY032555; Swartz fellowship; Kavli fellowship

TALK 3, 11:15 AM, 32.23

CHANGES IN POPULATION RECEPTIVE FIELDS FOLLOWING ARTIFICIAL SCOTOMA; REMAPPING OR NONLINEAR RESPONSES?

Marcus Daghlian[1,2,3,4] (m.daghlian@spinozacentre.nl), Mayra Bittencourt[1,2,3,4], Remco Renken[2,6], Serge Dumoulin[1,3,4,5], Frans Cornelissen[2,6]; [1]Spinoza Centre for Neuroimaging, [2]Laboratory for Experimental Ophthalmology, UMCG, [3]Netherlands Institute for Neuroscience, Royal Netherlands Academy of Sciences, [4]Vrije Universiteit, [5]Utrecht Universiteit, [6]Rijksuniversiteit Groningen

There is considerable debate regarding the extent of plasticity in the adult visual cortex, following retinal lesions. Early electrophysiological studies in animal models provided evidence that receptive fields (RF) inside the lesion projection zone adaptively change position preference towards spared portions of the visual field. However, subsequent multimodal studies did not find evidence of RF position change. Changes in population RFs (pRF) have also been observed in healthy controls following simulated scotoma. Importantly, changes in pRF locations are not limited to the simulated lesion projection zone, as pRFs around the visual field display change their apparent position. This suggests that changes in pRFs around scotoma, either simulated or real, are not necessarily due to plasticity. Previous studies generally used a linear, single gaussian pRF model. However, the visual cortex responds non-linearly to stimuli, hence it is possible that apparent position changes following scotoma are driven by non-linear responses. To test this hypothesis, five participants were shown a standard retinotopic mapping stimulus, with and without simulated scotoma (an unstimulated mean-luminance patch on the screen). We modelled the pRF properties using both a linear single gaussian and a pRF model based on divisive normalization (DN). The DN pRF model captures non-linear responses, such as surround suppression, compression and oversaturation. In line with previous studies, we find changes in preferred position using the linear single gaussian pRF model. When fitting with the DN pRF model, the results also displayed position changes, however, these were significantly smaller. Our results suggest that at least a portion of pRF preferred position changes can be captured by non-linear responses. Thus, non-linear responses may be misinterpreted as signs of plasticity, and we propose that the studies of cortical plasticity and stability must consider non-linear responses of visual cortex which are part of normal cortical dynamics.

TALK 4, 11:30 AM, 32.24

OCULOMOTOR INFLUENCES ON EXTRAFOVEAL SENSITIVITY

Jie Z. Wang[1] (jwang255@ur.rochester.edu), Michele Rucci[1]; [1]University of Rochester

The human eyes are always in motion, alternating rapid gaze shifts (microsaccades) with slow smooth movements (drifts) even when attending to a single point. In the fovea, these fixational eye movements (FEM) have been shown to enhance sensitivity in complementary spatial frequency ranges, in a way that is consistent with their reformatting of spatial patterns into temporal signals: the luminance modulations from microsaccades and drifts emphasize low and high spatial frequencies, respectively. Outside the fovea, however, the perceptual roles of FEM remain unclear. In the periphery, views range from general functions ('refreshing' percepts) to no function because of the little FEM motion relative to receptive fields size. Here we show that FEM lead to similar perceptual consequences inside and outside the fovea. Human observers (N=6) were asked to report the orientation (±45°) of a full-field grating while maintaining fixation. The grating was either at high (10 cpd) or low (0.2 cpd) spatial frequency. To restrict stimulation to the peripheral visual field, a circular gray patch (diameter 15° or 23°) remained stationary on the retina centered on the line of sight. We compared performance in the presence and absence of the retinal motion caused by FEM. In the latter condition, eye movements were counteracted in real-time by moving the stimulus on the display via a custom apparatus. Our results show that, also outside the fovea, drifts and microsaccades selectively improve sensitivity to high and low spatial frequencies, respectively. On average performance dropped by approximately 10% at high frequency when retinal motion was eliminated and improved by a similar amount at low frequency in the trials with microsaccades. Together, these results indicate that FEM operate uniformly throughout the visual field, reformatting luminance patterns into spatiotemporal signals that enhance contrast sensitivity in complementary ranges of spatial frequencies.

This work was supported by NIH EY018363 and P30 EY001319

TALK 5, 11:45 AM, 32.25

SEEING LESS BUT SEEING BETTER: INFORMATION LOSS AND ACCURACY GAIN IN REDUNDANCY MASKING

Bilge Sayim[1] (bilge.sayim@univ-lille.fr), Dogukan Nami Oztas[2], Li L-Miao[1], Nihan Alp[2]; [1]CNRS, University of Lille, [2]Sabanci University

To cope with excessive visual information in the environment, the visual system selects, discards, and compresses information. One compression mechanism is redundancy masking (RM) where redundant visual information is compressed. RM occurs with as few as three items. For example, when presented with three identical items in the visual periphery, observers often report seeing only two items. Here, we investigated to what extent features of masked items withstand or are lost in RM. We presented 3-5 radially arranged bars with varying widths (0.1°, 0.25°, 0.4°) for 150ms in the left or right hemifield (10° eccentricity). Observers reported the number of bars, and then adjusted probe widths and spacings to match the perceived stimulus. We computed deviation scores as the difference between perceived and actual (1) number of bars, (2) bar widths, (3) spacings between bars, and (4) overall widths of the arrays. There was strong RM: The number of bars reported was lower than the number presented. Overall, the width of thin (thick) bars tended to be overestimated (underestimated). In RM trials, the reported width was slightly larger than in non-RM trials. Importantly, except for the thinnest width condition, the reported width was more accurate in RM than in correct trials. The reported spacing between bars was larger in RM compared to correct trials, showing a lower perceived density in RM, while the reported overall extent of the arrays was smaller in RM trials, replicating previous results of visual space compression in RM. Our results suggest that the erroneous perception of smaller numbers of items in RM may go hand in hand with higher accuracy in reporting their features. We discuss how RM can be beneficial beyond the economical use of limited processing capacities by improving perception of individual items.

ANR-19-FRAL-0004; Tubitak 122N748

Talk 6, 12:00 pm, 32.26

COMPUTATIONAL ASPECTS OF GROUPING EXPLAIN VISUAL CROWDING ACROSS SPACE AND TIME

Martina Morea[1] (martina.morea@epfl.ch), Michael Herzog[1], Gregory Francis[2], Mauro Manassi[3]; [1]Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, [2]Department of Psychological Sciences, Purdue University, West Lafayette, USA, [3]School of Psychology, University of Aberdeen, King's College, Aberdeen, United Kingdom

In crowding, perception of a target deteriorates in the presence of flanking elements. Surprisingly, crowding can sometimes be ameliorated when more flanking elements are presented, a phenomenon called uncrowding. It was previously shown that uncrowding occurs when the target ungroups from the flankers. Here, we show that grouping processes can explain not only spatial interactions in crowding, but also interactions in the time domain. First, we show that grouping requires a minimum stimulus duration to occur: when participants discriminated the offset of a target vernier presented alone or flanked by lines or cuboids, a cuboid duration of at least 160 ms was needed for uncrowding to occur. Second, we show that the grouping process can be initiated by presenting only the cuboids for 20 ms before an ISI and then the display with the cuboids and the target vernier for 20 ms. With the preview, uncrowding occurs for short ISIs of 20 ms up to ISIs of 250 ms, pointing out to recurrent grouping processes taking place. Third, when presenting flanking elements during the ISI, the uncrowding effect occurred only when the elements formed a good Gestalt. We show that this body of results can be well explained by the Laminart model, in which recurrent processing segments the visual scene into different objects, which are represented in separate segmentation layers. When the target and flanking cuboids are processed in different layers, uncrowding occurs; when they are in the same layer, crowding occurs. Importantly, the preview of the cuboids gives the model sufficient time to segment the vernier target away from the cuboids. Taken together, our results highlight the importance of recurrent grouping processes in spatial and temporal interactions in vision.

Talk 7, 12:15 pm, 32.27

A CONTINUOUS TRACKING MEASURE OF ORIENTATION SENSITIVITY AND BIAS IN THE VISUAL PERIPHERY

Zainab Haseeb[1] (zainab.haseeb@mail.utoronto.ca), Anna Kosovicheva[1]; [1]University of Toronto

Visual performance varies significantly across the visual field, revealing variations in sensitivity at different locations within and across observers. These include polar angle asymmetries—variations in performance across angular locations. Conventional methods for measuring these variations are time consuming but can be made more efficient with recent continuous tracking methods, in which observers follow a continuously changing target. This method calculates the peak of the cross-correlation between the tracked and reference stimuli, effectively assessing sensitivity. However, it does not directly quantify perceptual bias, which reflects systematic errors in perception. To address this, we introduce a novel approach to simultaneously map bias and sensitivity in orientation perception across the visual field at 8º eccentricity across four locations (upper, lower, right, and left). Participants fixate a central grating and adjust its orientation to match the orientation of a randomly rotating peripheral grating. We measured perceptual sensitivity by calculating the peak of the cross-correlation between the central and peripheral gratings. In addition, we measured bias by calculating the difference between observed and actual orientation values at each orientation, which are then grouped and averaged. To validate this approach, participants completed a second condition, in which we used the tilt illusion to measure biases in perceived orientation with a 45º annular surround for the peripheral grating. We reveal significant variations in the strength of the tilt illusion among participants and locations. Additionally, participants demonstrated significant variation in location-specific sensitivity in tracking the grating, both with and without the annulus in the periphery. Sensitivity was well correlated between the two tasks ($p < .001$), but lower with a surrounding annulus. Our results highlight individual differences in sensitivity and bias across the visual field with our novel continuous tracking paradigm, and variation in the magnitude of the tilt illusion in different peripheral field locations.

This work was supported by an NSERC Discovery Grant to AK

Talk Session: Sunday, May 19, 2024, 2:30 – 4:15 pm, Talk Room 1

Attention: Tracking, shifting, capture

Moderator: Anna Nobre, Yale University

Talk 1, 2:30 pm, 34.11

SPATIOTEMPORAL PROCESSING DRIVES THE CONTRALATERAL DELAY ACTIVITY IN A DUAL WORKING MEMORY AND ATTENTIONAL TRACKING TASK

Piotr Styrkowiec[1,2] (pstyrkowiec@uchicago.edu), William Ngiam[1], William Epstein[1], Ron Gneezy[1], Edward Awh[1], Edward Vogel[1]; [1]University of Chicago, USA, [2]University of Wroclaw, Poland

Recent work has suggested that storage in visual working memory (VWM) occurs through the assignment of spatiotemporal pointers to the to-be-remembered items (Thyer et al., 2022). Thus, VWM capacity limits may not be set by the stimulus content exactly, but rather by attentional processes that define the spatiotemporal pointers for item-based storage. We examined whether this is the case in the contralateral delay activity (CDA), an event-related potential long known to track VWM load. The CDA has been shown to track the number of targets in multiple-object tracking (Drew and Vogel, 2008), but also the number of to-be-remembered colors (Vogel and Machizawa, 2004). To directly contrast the effects of attentional tracking load and stimulus content load on working memory, we developed a novel dual-task paradigm. Participants track either one or two moving discs (attentional tracking load), with either two or four colors displayed across each of the discs (working memory load). Participants completed a 'tracking only' condition, where they would need to monitor the moving target discs like in a multiple-object tracking task, and a 'tracking plus memory' condition, where they would track the discs and remember all displayed colors like in a multiple-identity tracking task. The key question was whether or not CDA amplitude would be determined by the number of individuated items tracked, or by the number of distinct colors associated with the currently tracked items. Strikingly, CDA amplitude was determined almost entirely by the number of items tracked, with no reliable effect of variations in the number of colors per tracked item. These findings suggest the CDA largely reflects the maintenance of spatiotemporal pointers for moving objects, not the number of feature values associated with those objects.

Talk 2, 2:45 pm, 34.12

ONE AT A TIME IN THE MIND'S EYE: SERIAL VERSUS PARALLEL MENTAL SIMULATION OF MOVING OBJECTS

Halely Balaban[1] (halelyb@openu.ac.il), Tomer Ullman[2]; [1]The Open University of Israel, [2]Harvard University

Our everyday environment is highly dynamic, and this creates great challenges for any cognitive system. Decades of research using the Multiple Object Tracking paradigm has shown that when people view moving objects, they can only track a handful of them at a time. But people's real-world tracking extends beyond direct perception: we can also keep track of occluded bodies, as well as imagine the yet-unseen future paths of objects (e.g., in order to estimate where a thrown ball will land). In this work, we examined the capacity limitations of tracking items in imagination, rather than perception. Across 4 experiments, participants (N=136 total) watched short 2D animations of balls moving under gravity. Animations paused mid-motion, and participants were asked to continue the scene in their mind's eye, and indicate when each ball hits the ground in their imagination. Responses were compared with the true impact time, which was extracted from simulated physics. With a single ball, people's imagination-based responses closely matched the true impact time (Exp. 1). However, once another ball was introduced, performance was significantly altered (Exp. 2), and followed the predictions of a computational model of serial simulation, which only moves a single object forward at a time. The serial pattern was not due to response requirements (Exp. 3), and was reduced, but not fully eliminated, by introducing extremely strong grouping cues that even allowed relying on heuristics (Exp. 4). Together, our results show that tracking objects in imagination is, at least under certain conditions, a serial process that operates on a single-object basis. More broadly, examining 'imagination tracking' highlights previously unexplored capacity limits in mental simulation, calling for an updating of current models of how humans make intuitive predictions of physical outcomes.

Talk 3, 3:00 pm, 34.13

LESS SALIENT, MORE CAPTURE: THE CURIOUS CASE OF ABRUPT ONSETS

Han Zhang[1] (hanzh@umich.edu), Kane York[1], John Jonides[1]; [1]University of Michigan

Abrupt onsets are commonly assumed to capture attention due to their high physical salience. Using a technique recently introduced by Stilwell et al. (2023), we directly compared the salience level of a color singleton, an abrupt onset, and a color singleton + abrupt onset item. We then assessed the magnitude of capture by these items in a visual search task, in which these items served as distractors. In an oddball-detection task, 26 participants reported the presence or absence of the critical item among four heterogeneous shapes. Applying the exact algorithm from Stilwell et al. (2023), we determined the exposure threshold for each item type. Abrupt onsets exhibited a significantly larger exposure threshold (61 ms), indicating lower salience, compared to color singletons (25 ms; $t(25) = 5.41$, $p < .001$) and color singleton + abrupt onset items (22 ms; $t(25) = 5.92$, $p < .001$). Next, another group of 30 participants completed a visual search task with the same display. Surprisingly, only the pure onset distractor elicited a capture effect, as indicated by response time differences (23.5 ms, $t = 5.46$, $p < .001$). In contrast, we found suppression effects for both the color singleton distractor (29.1 ms, $t = 5.28$, $p < .001$) and the color singleton + abrupt onset distractor (26.3 ms, $t = 3.45$, $p < .001$). The interaction between distractor presence and distractor type was highly significant, $F(2, 58) = 30.15$, $p < .001$. Finally, these findings were replicated when the same group of participants (N = 26) completed both tasks. Overall, individuals could suppress an abrupt onset distractor with high salience but not with lower salience. These results suggest that abrupt onsets maintain a privileged status in the visual system and tend to capture attention, despite their relatively low salience.

This work was supported by the National Science Foundation [grant number: 1658268] and the National Institute of Mental Health (Unique Federal Award Identification Number (FAIN): R21MH129909).

Talk 4, 3:15 pm, 34.14

SUSCEPTIBILITY TO ATTENTIONAL CAPTURE BY TARGET-MATCHING DISTRACTORS PREDICTS HIGH VISUAL WORKING MEMORY CAPACITY

Zhe Qu[1] (quzhe@mail.sysu.edu.cn), Yulong Ding[2]; [1]Sun Yat-sen University, [2]South China Normal University

Introduction: It is well known that the ability of distractor suppression has a close relation to visual working memory (VWM). Recent event-related potential (ERP) studies (e.g., Gaspar et al., 2016) showed that individuals with high working memory capacity could call on an early suppression (indexed by Pd component) to salient-but-irrelevant distractors whereas those low-capacity ones could not. However, it remains unclear whether such an early suppression mechanism applies to nonsalient distractors that possess the target-defining feature. The main purpose of the present study is to investigate the relation between the VWM capacity and the attentional process of feature-matched distractors. Methods: Fifty-one healthy young adults participated in this study. Like previous studies (e.g., Gaspar et al., 2016), individual VWM capacity was measured by the K-score of change detection task. In the ERP experiment, we adopted a central rapid serial visual presentation (RSVP) task in which distractors with or without the target-defining feature were presented peripherally. Participants were informed to identify a digit of a specified color in the central RSVP stream. We focus on two distractor-elicited ERP components (i.e., N2pc and Pd) that would reflect two distractor-related attentional processes (i.e., attentional capture and attentional suppression), respectively. Results: Surprisingly, we revealed that high-capacity individuals would be captured more attention (reflected by a larger distractor-N2pc emerging in less than 200 ms) compared to low-capacity ones, meaning that they are less able to ignore such distractors in early visual processing. Although feature-matched distractors captured more attention for high-capacity individuals, they received stronger suppression soon afterwards, indexed by a larger distractor-Pd. Conclusion: Compared to low-capacity individuals, high-capacity ones would first be captured more attention by feature-matched distractors, then enact more inhibition to those distractors. Our findings support that, high-capacity individuals could exhibit more flexible ways of attentional processing when dealing with different kinds of distractors.

TALK 5, 3:30 PM, 34.15

CUED SUPPRESSION AND LEARNED SUPPRESSION RELY ON SEPARATE MECHANISMS

Nancy Carlisle[1] (nancy.carlisle@gmail.com), Matthieu Chidharom[1,2]; [1]Lehigh University, [2]University of Chicago

In recent years, two forms of attentional suppression have been identified. In cued suppression, participants avoid certain distractor features which are cued on each trial. In learned suppression, participants gradually learn to ignore bottom-up features, which would normally capture attention, through repeated exposure. Although both types of suppression are well established, it is not clear if both types of suppression rely on the same underlying mechanisms. To address this question, we examined correlations (N=54) between cued suppression and learned suppression, as well as WM capacity and a real-world inattentive trait questionnaire. If cued and learned suppression rely on similar mechanisms, we expect correlated amounts of suppression across individuals. Instead, we found that cued and learned suppression were not correlated. Cued suppression was correlated with WM capacity, in line with the idea that these tasks relied on cognitive control. Learned suppression was correlated with the real-world inattentive trait questionnaire, suggesting that reductions in learning to suppress may create distractibility both inside and outside the lab. These results suggest that cued and learned suppression are separable forms of attentional guidance, not two ways of activating the same suppression mechanism.

1R15EY030247

TALK 6, 3:45 PM, 34.16

CONTINUOUS PSYCHOPHYSICS REVEALS A TEMPORAL COST OF REAL-TIME DISTRACTOR SUPPRESSION DURING ENSEMBLE PERCEPTION

Kevin Ortego[1], Viola Stoermer[1]; [1]Dartmouth College

Ensemble perception plays a fundamental role in how our visual system represents complex scenes. It is commonly studied by briefly presenting sets of stimuli to participants and having them report the average across the feature dimension of interest (e.g. orientation). In real scenes, however, multiple feature sets are often present at the same time, and their inputs to the visual system change continuously, due to variability in the environment or eye and head movements of the observer. Here, we developed a new task to test how participants track feature summaries continuously and how irrelevant distractor features affect the precision and time course of ensemble estimates. During 45-second long tracking trials, participants viewed a set of oriented lines that continuously changed orientations, and concurrently rotated a joystick to reproduce the average orientation of those lines. Following previous work using continuous psychophysics (Bonnen et al., 2015), we computed the cross-correlation between the mean target orientation and the response time series, with the peak amplitude of the resulting cross-correlogram reflecting perceptual sensitivity and the peak latency reflecting processing time. We first validated that our novel task produces results that parallel findings from traditional ensemble tasks: the precision of orientation estimates increased when more items were present ($p=0.003$), and decreased with higher variability of the orientation set ($p<0.001$). Next, we examined the impact of distractors on tracking performance by adding differently-colored distractor lines. We found lower precision ($p=0.001$) and a temporal cost of ~100ms ($p=0.025$) when distractors were present. Together, these results suggest that the presence of distractors in ensemble processing impairs and delays the extraction of relevant feature summaries, and demonstrate the utility of continuous tasks by revealing a temporal cost that may not be captured by traditional reaction time measures.

TALK 7, 4:00 PM, 34.17

SEPARATING RHYTHMS OF SENSORY AND MOTOR PREPARATION

Sage E.P. Boettcher[1], Anna-Katharina Bauer[2], Berit Hartjen[3], Sabine Kastner[3], Anna C. Nobre[4]; [1]University of Oxford, [2]Royal Holloway University of London, [3]Princeton University, [4]Yale University

Covert attention enhances the processing of relevant stimuli within our environment without the need for overt eye movements. Recent work has demonstrated that our covert attentional system rhythmically samples the environment approximately 3 to 8 times per second. Our sensory system must work in tandem with our motor system to produce fluent natural behaviour. That is, once we perceive a task-relevant stimulus, we must execute the appropriate response. Although past work has shown an important role for actions in resetting perceptual sampling, it remains unclear to what extent motor preparation is itself rhythmically modulated. To investigate whether motor preparation follows a similar rhythmic sampling as found in perception, we designed a task that orthogonalized sensory and motor preparation. Compound cues indicated both the most likely location of an upcoming target stimulus (left or right visual field) and the most likely motor response (left or right button press). Both sensory and motor predictions of cues were valid in a majority of trials. However, in a minority of trials, predictions could be invalid in only the sensory domain, only the motor domain, or both domains with equal likelihood. The interval between the cue and target varied between 300 and 1100 ms. This manipulation allowed us to interrogate the effects of valid sensory vs. motor cues orthogonally time. Consistent with previous work we found that sensory cues modulated behaviour in a range between 3 and 8 Hz. In contrast, interestingly, motor preparation showed little evidence of rhythmic modulation.

Support by Wellcome Trust Award (104571/Z/14/Z) and JSMF Award (220020448) to A.C.N. As well as an EPS Postdoctoral fellowship awarded to S.E.P.B. The WIN centre is supported by the Wellcome Trust (203139/Z/16/Z). The work is supported by the NIHR Oxford Health Biomedical Research Centre.

Talk Session: Sunday, May 19, 2024, 2:30 – 4:15 pm, Talk Room 2

Color, Light and Materials: Materials, integrated perception

Moderator: Katja Doerschner, Justus Liebig University, Giessen

Talk 1, 2:30 pm, 34.21

SENSOR-BASED QUANTIZATION OF A COLOR IMAGE DISTURBS MATERIAL PERCEPTION

Masataka Sawayama[1] (masa.sawayama@gmail.com); [1]Graduate School of Information Science and Technology, The University of Tokyo, Japan

Materials in our daily environments undergo diverse color changes based on environmental contexts. For instance, water is inherently colorless, but wetting a surface changes its colors due to optical interactions. Previous studies have explored the effect of colors on material perception while following the literature on object recognition, i.e., examining the effect of categorical colors on a grayscale image. However, unlike object recognition, categorical colors are not always diagnostic for material changes due to context dependence. To address the issue, this study explores color dimensions diagnostic to material perception. Building on recent studies showing that material perception depends on image color entropy (Sawayama et al., 2017), this study investigated the extent to which modulating the image color entropy, defined by the color quantization in a sensor color space (e.g., RGB or LMS), affects the material estimation. Specifically, the experiment leveraged a zero-shot prediction paradigm using pre-trained vision and language machine-learning models. It used 2AFC text prompts related to material perception, such as wet/dry or glossy/matte. The FMD (Sharan et al., 2014) and THINGS (Hebart et al., 2019) datasets were chosen for visual images. Color quantization was applied through the median cut to each image, reducing the quantized numbers from 128 to 2. Additionally, grayscale images were created from the original images. Results showed that the distribution of prediction probabilities was diversely distributed for original and grayscale images across all dataset images. However, when an original image was modulated by color quantization, the distribution diversity was biased heavily towards specific attributes, particularly dry and matte. Further experiments confirm that color quantization has less impact on zero-shot object recognition performance. These findings suggest that diverse material perception of an object image is available for high color entropy, where the color space is defined while mixing chromatic and luminance components.

Talk 2, 2:45 pm, 34.22

CORTICAL REPRESENTATIONS OF CORE VISUAL MATERIAL DIMENSIONS

Hua-Chun Sun[1,2] (hua-chun.sun@psychol.uni-giessen.de), Filipp Schmidt[1,2], Alexandra C. Schmid[3], Martin N. Hebart[1,2,4], Roland W. Fleming[1,2]; [1]Justus Liebig University Giessen, [2]Center for Mind, Brain and Behavior (CMBB), University of Marburg and Justus Liebig University Giessen, [3]Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, [4]Vision and Computational Cognition Group, Max Planck Institute for Human Cognitive and Brain Sciences

In every waking moment, we perceive numerous visual materials from the objects, surfaces and environment around us. How does the brain represent the great diversity of materials and their properties? We recently addressed the mental representation of materials using a novel dataset consisting of 600 images spanning 200 material categories (the STUFF dataset), by crowdsourcing over 1.8 million material similarity judgments. This revealed 36 core dimensions that capture similarity relationships between materials (Schmidt, Hebart, Schmid & Fleming, 2023). To determine the neural representation of these dimensions in the human brain, here we acquired a densely-sampled functional MRI dataset using these images, which we paired with an encoding model of the 36 material dimensions. Each of the 600 images was presented to six participants 14 times each across multiple scanning sessions. The whole brain activation map of each material dimension was then obtained by modeling the dimension score of each image in each of the 36 dimensions (Schmidt et al, 2023). Comparing the voxel-wise activation intensity across material dimensions revealed superimposed cortical maps associated with each of the dimensions. We found that dimensions related to the fine scale granularity of the material are particularly represented in early visual areas (V1-V3). In contrast, dimensions related to hard shapes preferably activated lateral occipital (LO) cortex, indicating a dichotomy between cortical regions associated with shape and fine texture. Flexible and soft material dimensions exhibited particularly strong responses in area hMT+/V5, suggesting that motion sensitive regions also encode the capacity of materials to deform. Finally, color dimensions, which span diverse material categories, were represented less consistently across participants, suggesting that material properties might actually be a more consistent organizing principle than color. Together, our findings provide a comprehensive mapping of material representations across cortical regions in the human brain.

This research is funded by the DFG (222641018 – SFB/TRR 135 TP C1), the HMWK ("The Adaptive Mind") and European Research Council Grant ERC-2022-AdG "STUFF" (project number 101098225).

Talk 3, 3:00 pm, 34.23

PERCEPTION OF MATERIAL PROPERTIES FROM DYNAMIC LINE DRAWINGS

Amna Malik[1] (amnaahmalik55@gmail.com), Ying Yu[5], Huseyin Boyaci[1,2,3,4], Katja Doerschner[1]; [1]Department of Psychology, Justus Liebig University, Giessen, Germany, [2]Interdisciplinary Neuroscience Program, Bilkent University, Ankara, Turkey, [3]Aysel Sabuncu Brain Research Center & National Magnetic Resonance Research Center (UMRAM), Bilkent University, Ankara, Turkey, [4]Department of Psychology, Bilkent University, Ankara, Turkey, [5]EYSZ Inc., United States

Recent studies have shown that people are able to recognize material qualities based on motion information alone (Schmid et al, 2018, Bi et al 2018). In these experiments, stimuli were deforming nonrigid objects (e.g. liquids, jelly, cloth) with dots 'stuck on' at random places in and on the object. The dots 'inherited' the kinematic material properties and global shape deformations of the object and the overall motion pattern yielded vivid nonrigid material qualities, such as wobbliness. Dynamic dot materials contain both, interior-, and boundary motion. However, how much does each of these types of motion contribute to the percept of a given material quality? To answer this question we contrast ratings of material qualities of dynamic dot stimuli with those for dynamic (out)line drawings of the same object-material deformations, as well as with ratings for corresponding full-texture (i.e. color & reflectance) renderings (for sample stimuli, see https://jlubox.uni-giessen.de/getlink/fi9xJ5W9kN1drXPBodkz1HCY/). Animations of five material categories (fabrics, hard breakables, jelly, liquids, smoke) were rated each six times on eight material attributes (dense, flexible, wobbly, fluid, airy motion, motion coherence, oscillatory motion, and motion dynamics), blocked by rendering style (dots, lines, full). Comparing dissimilarity matrices and cluster analysis of attribute ratings between the three rendering conditions suggest that 1) also animated line drawings vividly convey mechanical material properties, 2) similarity in material judgements between line drawings and fully textured animations was larger than that between dynamic dots and fully textured stimuli. We conclude that boundary motion might play a critical role in the perception of mechanical material qualities.

Talk 4, 3:15 pm, 34.24

PERCEIVING MATERIALS AND OBJECTS FROM SEMI-VISIBLE INTERACTIONS

Vivian C. Paulun[1] (vpaulun@mit.edu), Max Siegel[1], Kristine Zheng[1], Joshua Tenenbaum[1]; [1]Massachusetts Institute of Technology

Visual material perception is computationally complex because physical properties such as rigidity or friction are not directly observable. In many cases, however, viewing a dynamic interaction between different objects reveals their internal properties. If a couch cushion deforms dynamically under the weight of a box, we infer the cushion's stiffness as well as the weight of the box. This indicates that the brain jointly interprets the interplay of multiple objects in a physical scene. Can the brain infer the physical structure when only one of the interacting objects is visible, while all others are artificially rendered invisible? To answer this question, we leveraged computer graphics: First, we simulated short interactions of liquid, granular, and non-rigid materials with rigid objects of various shapes. Then, crucially, we rendered only the target material while the remaining scene was black. We presented the videos to 100 observers and asked them to identify which of two alternative interactions showed the same target material as the test video. Match and distractor varied in their material properties (e.g., cohesion), thus implicitly requiring inference of those parameters. Observers were as accurate in judging these videos, as they were when presented with fully rendered versions. Strikingly, we found that observers did not only perceive the target material in rich detail; in most cases, they were able to select which of two alternative 3D shapes was underlying the observed interaction. This finding suggests that the brain imputes the hidden objects in a physically plausible manner. In comparison, a distance-based classifier based on features from pretrained neural networks showed overall lower performance in both tasks and the pattern of errors was different from human observers. Taken together, our results are consistent with the hypothesis that people use an internal generative physics model in online perception.

This work was supported by the German Research Foundation (grant PA 3723/1-1) and National Science Foundation NSF NCS Project 6945933.

Talk 5, 3:30 pm, 34.25

THE SOUND OF SHININESS: CROSS-MODAL INFLUENCE OF AUDITORY PITCH ON THE PERCEPTION OF GLOSS

Silas Choudhury[1] (chous977@newschool.edu), Hong B. Nguyen[1], Benjamin van Buren[1]; [1]The New School

When shopping for jewelry online, we typically just see pictures of items, and cannot hear or feel them. One might assume that this is not a big loss—at the very least, that removing auditory or tactile information should not influence our perception of intuitively visual properties, such as gloss. However, it is also possible that perceivers irresistibly integrate auditory information, such as pitch, into their perception of gloss. We investigated this in two experiments. In Experiment 1, subjects saw pairs of spheres, which were both rendered in the same material (metal, wood, or leather). The spheres differed slightly in glossiness, and subjects reported which was glossier, while ignoring concomitant sounds. In one condition, the glossier sphere was paired with a high-pitched sound and the less glossy sphere with a low-pitched sound; in the other condition, the pairings were reversed. Surprisingly, subjects could not ignore the sounds when discriminating gloss. Rather, they were more accurate when the glossier sphere was paired with the high-pitched sound, suggesting an automatic association between higher pitch and higher gloss. These objects were computer-generated; does the same association also hold when viewing pictures of real objects? In Experiment 2, subjects viewed pictures of jewelry from the Metropolitan Museum of Art's digital archive. Each item was paired once with a high-pitched sound and once with a low-pitched sound. Subjects were instructed to ignore the sounds, and to rate the gloss of each item from "Not at all shiny" to "Very shiny". They rated jewelry items as much shinier when paired with the high-pitched sound, indicating that the association between high pitch and high gloss also holds for pictures of real objects. We conclude that when displaying objects in digital spaces such as online stores and museum catalogs, auditory pitch can be used to drive impressions of gloss.

Talk 6, 3:45 pm, 34.26

LIGHTNESS CONSTANCY CAN BE VERY WEAK IN AN IMMERSIVE VR ENVIRONMENT

Khushbu Y. Patel[1] (khushbup@my.yorku.ca), Laurie M. Wilcox[1], Laurence T. Maloney[2], Krista A. Ehinger[3], Jaykishan Y. Patel[1], Richard F. Murray[1]; [1]York University, [2]New York University, [3]The University of Melbourne

Previous studies have revealed important differences between how viewers perceive real and virtual scenes. Virtual reality (VR) plays a growing role in performance-critical applications such as medical training and vision research, and so it is crucial to characterize perceptual differences between real and VR environments. We compared lightness constancy in real and VR environments. We used a demanding task that required observers to compensate for the orientation of a reference patch relative to a light source in a complex scene. On each trial the reference patch had reflectance 0.40 or 0.58, and a range of 3D orientations (azimuth -50º to 50º). Ten observers adjusted a grey match patch to match the perceived grey of the reference patch. We used a custom-built physical apparatus, and four VR conditions: All-Cues (replicated the physical apparatus); Reduced-Depth (zero disparity, no parallax); Shadowless (no cast shadows); and Reduced-Context (no surrounding objects). Scenes were rendered in Unity and shown in a Rift S headset. Surprisingly, constancy was weak, and approximately the same in all conditions. The mean Thouless ratio (0= no constancy, 1= perfect constancy) was 0.40, with no significant differences between conditions. The above-zero constancy in the Reduced-Context condition, with no cues to support constancy, suggested that observers learned environmental lighting cues in some conditions and transferred this knowledge to other conditions. Accordingly, we re-tested the All-Cue and Reduced-Context conditions in VR, with 10 new observers per condition, and each observer ran in just one condition. Here we found substantially reduced constancy (average Thouless ratio 0.14). We conclude that lightness constancy can be weak in VR, and that observers may use lighting information from real environments to guide performance in virtual environments. We are currently developing experiments with high-performance VR configurations to test whether constancy improves with more realistic rendering of lights and materials.

NSERC, VISTA

Talk 7, 4:00 pm, 34.27

REPRESENTATIONAL MOMENTUM IS DOMAIN-GENERAL: EVIDENCE FROM BRIGHTNESS

Dillon Plunkett[1] (d.plunkett@northeastern.edu), Jorge Morales[1]; [1]Northeastern University

A classic finding in visual cognition is “representational momentum”. Show people a photo of a wave crashing on the beach and they are prone to confuse it with a photo taken a moment later. Abruptly mask a video of a rotating shape or a rapidly melting ice cube and people will overestimate how far they saw the shape rotate or the ice melt. Anticipated motion affects what we see, or at least what we remember seeing. Prior research has argued that representational momentum is strictly limited to anticipated motion, in part by appealing to experiments that found no representational momentum for changes in brightness. Here, we refute this claim with new evidence. In five experiments, we demonstrate that richer stimuli and a more sensitive task reveal people to experience representational momentum for changes in brightness. Participants watched animations of a stationary, achromatic shape that increased in brightness before being masked. Using a slider, they selected a specific frame from each animation to indicate precisely how bright they thought the shape was when it disappeared. Participants reliably judged the shape to have been brighter than what they had truly been shown. We found analogous representational momentum effects for darkening shapes, the brightness of chromatic stimuli, and—generalizing our results beyond 2D shapes—the brightness of the ambient light illuminating a 3D object in a computer-rendered scene. These findings are unlikely to be an artifact of the task design because representational momentum for brightness replicated in a 2AFC version of the task and, most importantly, no analogous effects were observed for changes in hue, which are less intuitively directional and predictable. These results suggest representational momentum is a domain-general phenomenon related to anticipatable change, not one that is narrowly limited to motion. The mind actively anticipates changes in many perceptual domains.

Talk Session: Sunday, May 19, 2024, 5:15 – 7:15 pm, Talk Room 1

Plasticity and Learning

Moderator: Yuka Sasaki, Brown University

Talk 1, 5:15 pm, 35.11

DIFFERENTIAL FUNCTIONAL REORGANIZATION OF VENTRAL AND DORSAL VISUAL PATHWAYS FOLLOWING CHILDHOOD HEMISPHERECTOMY

Vladislav Ayzenberg[1] (vayzenberg90@gmail.com), Michael Granovetter[2], Sophia Robert[2], Christina Patterson[3], Marlene Behrmann [4]; [1]Department of Psychology, University of Pennsylvania, [2]Department of Psychology, Carnegie Mellon University, [3]School of Medicine, University of Pittsburgh, [4]Department of Ophthalmology, University of Pittsburgh

A key signature of neural development is that the brains of children have a greater capacity for recovery following damage or surgery than adults. Indeed, children who have had an entire hemisphere of their brain removed to treat epilepsy (a procedure known as hemispherectomy) show a high degree of perceptual functioning despite the loss of both ventral and dorsal visual pathways in one hemisphere. Yet, accumulating evidence suggests that the dorsal pathway may mature earlier than the ventral pathway – raising the question of whether the two pathways also have a different capacity for recovery after surgery. In the current study, we sought to understand the extent to which functions of the ventral and dorsal pathways reorganize to the contralateral hemisphere following childhood hemispherectomy. We collected fMRI data from an equal number of left and right hemispherectomy patients (N = 8; age-at-surgery = 1-13 years; age-at-testing = 12-37 years) who completed tasks that typically elicit lateralized responses from the ventral or the dorsal pathway in controls, namely, word (left ventral), face (right ventral), tool (left dorsal), and global form (right dorsal) perception. Overall, there was greater evidence of functional reorganization in the ventral pathway than in the dorsal pathway. The majority of hemispherectomy patients showed normal degrees of word and face selectivity in their intact ventral pathway, despite losing the typically preferred hemisphere for each category. By contrast, only one patient showed evidence of normal selectivity for either tools or global form in their intact dorsal pathway. Importantly, because ventral and dorsal reorganization was tested within the same patients, these results cannot be explained by idiosyncratic factors such as disease etiology or age at the time of surgery. These findings suggest that the dorsal pathway has a shorter developmental window of plasticity than the ventral pathway because it matures earlier.

Talk 2, 5:30 pm, 35.12

PATIENTS WITH V1 DAMAGE EXHIBIT INCREASED ORIENTATION DECODING IN HMT+, BUT ONLY IF PULVINAR IS INTACT

Bing Li[1] (bing.li3@nih.gov), Tina T. Liu[1], Matthew R. Cavanaugh[2,3], Helena P. Bachmann[1], Berkeley K. Fahrenthold[2,3], Shruti Japee[1], Krystel R. Huxlin[2,3], Elisha P. Merriam[1]; [1]Laboratory of Brain and Cognition, National Institute of Mental Health, NIH, Bethesda, MD, USA, [2]Flaum Eye Institute, University of Rochester Medical Center, Rochester, NY, USA, [3]Center for Visual Science, University of Rochester, Rochester, NY, USA

Orientation selectivity is a core property of V1 in mammals. Patients with V1 damage can relearn orientation discrimination at trained, blind-field locations. Here, we investigated a potentially key role for pathways bypassing V1, which directly transmit information to downstream visual cortical areas, generating orientation selectivity in these areas. We studied 3 stroke patients (33-63 y/o, all females): one with a large right V1 lesion, a second with a right V2/V3 lesion that spared V1, and a third with lesions affecting both right V1 and pulvinar. Participants were scanned with BOLD fMRI. They viewed small (2.5 deg radius), oriented (45 or 135 deg) gratings in the periphery (7.1-11.2 deg eccentricity) while performing a demanding task at fixation. Stimuli were presented either deep within the blind field or in a mirror-symmetric location in the intact hemifield. We also performed retinotopic mapping, scanning with an MT localizer, and T1-weighted structural scans. In the patient with extensive V1 damage, the ipsilesional hMT+ was visually responsive and able to decode orientation. However, in the patient with a V2/V3 lesion and the patient with right V1 plus pulvinar damage, the ipsilesional hMT+ was visually responsive, but orientation decoding failed to reach significance. Healthy controls exhibited significant orientation decoding in V1 but not in hMT+. Our findings suggest that after V1 damage, strong orientation selectivity may emerge in hMT+ from circuits bypassing V1, including via the pulvinar. When V1 is spared, even when V2/V3 are damaged, these alternative circuits do not generate robust orientation-selective BOLD signals in hMT+. Ongoing work is assessing the functional implications of these findings for perception and rehabilitation potential.

This work was supported by the Intramural Research Program of the National Institute of Mental Health (ZIAMH002966).

Talk 3, 5:45 pm, 35.13

TRAINING-INDUCED FUNCTIONAL HOMOGENIZATION IN THE OCCIPITOTEMPORAL CORTEX: DIFFERENTIAL CROSS-MODAL MECHANISMS IN BLINDNESS VS. SEVERE LOW VISION

Lora Likova[1] (lora@ski.org), Zhangziyi Zhou[1], Michael Liang[1], Christopher Tyler[1]; [1]Smith-Kettlewell Eye Research Institute

No visual motion is available to totally blind (TB) individuals, raising questions about the utilization of the territory of the visual motion complex (hMT+) in the absence of vision, and its potential reorganization through training. This study explores whether such reorganization differs between TB individuals and those with some residual vision, such as severe low vision (SLV). Methods: TB and SLV subjects underwent five sessions of the Cognitive-Kinesthetic Memory-Drawing Training for spatial navigation. Pre- and post-training, whole-brain scans (Siemens 3T Prisma) were conducted while subjects (i) haptically explored and memorized raised-line tactile maps (30 s); after a 20 s rest, they (ii) drew-from-haptic-memory (30 s) the maps using the opposite hand with a stylus. Results: Despite the absence of visual input, hMT+ was robustly activated bilaterally in TB individuals in the right-hand blind Memory-Drawing task. However, the left-hand Haptic Exploration task activated only right hMT+, an unexpected interhemispheric functional asymmetry. Furthermore, following the training, significant brain reorganization occurred in the lateral occipitotemporal cortex, forming clusters of cortical areas TPOJ 1-3, FST, MTG and LOd expressing the same task-response asymmetry as hMT+. In SLV individuals, on the other hand, although a similar large-scale functional clustering occurred, hMT+ was surprisingly excluded and even unilaterally suppressed. Granger Causal connectivity analysis revealed a complex interplay between hMT+, its surrounding cluster, and motor, somatosensory, and memory areas. Conclusions: The multidimensional findings shed light on non-visual hMT+ functionality, its novel interhemispheric asymmetries, and their implications for functional brain architecture and its reorganization through learning. Furthermore, the results reveal for the first time the emergence of training-induced functional-homogenization of extended clusters of occipitotemporal areas around hMT+ in the visually deprived, which in the sighted are functionally distinct, demonstrating a mechanism for a novel type of cross-modal functional reorganization, offering crucial insights into neuroplasticity and sensory compensation.

NIH/NEI EY024056 & NSF SL-CN1640914 to L. Likova

Talk 4, 6:00 pm, 35.14

INDIVIDUAL DIFFERENCES OF FUNCTIONAL BRAIN PLASTICITY IN CENTRAL VISION LOSS

Pinar Demirayak[1] (pinarde@uab.edu), Rachel Chua[1], Leland Fleming[2], Kristina Visscher[1]; [1]University of Alabama at Birmingham, Heersink School of Medicine, Department of Neurobiology, Birmingham, AL 35294, USA, [2]McLean Hospital/Harvard Medical School, Boston, MA 02115, USA

Functional organization of the visual cortex is largely consistent across individuals. However, it is not clear to what extent the functional connectivity (FC) patterns between visual cortex and other areas vary

across individuals, nor is it clear to what extent these connectivity patterns are shaped by experience. Central vision loss provides an excellent model to investigate visual system plasticity because different features of experience impact the same participants, i.e. sensory deprivation in central vision and increased usage of parts of peripheral vision. We studied whole-brain FC patterns for parts of primary visual cortex (V1) corresponding to parts of retina associated with increased and decreased use. We performed both group-level and individual-specific analyses in 21 participants with central vision loss and 22 participants with healthy or corrected-to-normal vision. Group-level FC results revealed that participants with central vision loss have reduced connectivity between sensory-deprived portions of V1 and temporo-parieto-occipital cortical areas, compared to controls. Group-level comparisons did not show any statistically significant difference in mean connection strength to areas of V1 corresponding to increased usage. However, when we implemented an individual-specific approach, we observed that both increased and decreased usage leads to alterations in FC patterns. Results suggested that increased usage leads to idiosyncratic changes in connections whereas decreased usage leads to more stereotyped connection patterns. Further, FC patterns to parts of V1 that process peripheral vision are more stereotyped than patterns of connections to central vision ($F(1,10877)=8.5697$, $p=0.0034$), whereas they are equally stereotyped in patients with central vision loss ($F(1,1447)=0.34051$, $p=0.56$). Thus, more nuanced changes in connections can be observed when inter-individual variability is taken into account. Overall, our study emphasizes the diversity in patterns of brain plasticity following central vision loss and highlights that FC from V1 maintains the capacity to adapt in adulthood.

We would like to thank NIH/NEI Grants to Visscher 1 U01 EY025858-01A1 & 1R01EY031589-01.

Talk 5, 6:15 pm, 35.15

GEOMETRIC CHANGES IN MONKEY V4 AND IT NEURAL RESPONSES DURING VISUAL CATEGORY LEARNING

Jenelle Feather[1,2] (jfeather@flatironinstitute.org), Long Sha[2], Gouki Okazawa[2,3], Nga Yu Lo[1,2], SueYeon Chung[1,2], Roozbeh Kiani[2]; [1]Flatiron Institute, [2]New York University, [3]Institute of Neuroscience, Shanghai

Behavior changes over the course of learning a task. This behavioral change is due to shifts in neural responses that support improved performance. Here, we investigated how the underlying representational geometries in visual areas V4 and IT of the Macaque visual system change during a categorization learning task. Visual stimuli varied in two independent attributes, and monkeys learned to categorize them based on a category boundary in the stimulus space that was defined by a combination of the attributes. Chronic neural population recordings were obtained from V4 and IT over multiple days of training while a monkey learned the task through receiving correct/incorrect feedback. Additionally, we recorded from the same neural populations while a monkey performed a fixation task viewing the same sets of stimuli. In all eight analyzed tasks, the monkey's performance on the categorization task improved with training. To link this behavioral improvement to the underlying population responses, we investigated how the geometry of neural population activity changed over the course of learning. We treated population responses to all stimuli in each of the two categories as manifold-like representations, and analyzed the geometric properties of these representations using mean-field theoretic manifold capacity analysis. As the monkey learned the task, we observed that the representations in both V4 and IT for the two classes became more separable as measured by an increase in manifold capacity. This increase in capacity was associated with a characteristic geometric change in the neural population response geometry. Our results suggest that both V4 and IT responses actively change during category learning in ways that directly lead to increased separability and improved readouts for downstream neural areas, and point towards future work linking these population-level geometric changes to local changes at the single-neuron level.

Talk 6, 6:30 pm, 35.16

DIFFERENTIAL UNCONSCIOUS CONTROL OF THE MEDIAL PREFRONTAL CORTEX DURING NON-REM AND REM SLEEP TO MITIGATE ANTEROGRADE AND RETROGRADE INTERFERENCES IN VISUAL PERCEPTUAL LEARNING

Takashi Yamada[1] (takashi_yamada@brown.edu), Theodore LaBonte-Clark[1], Shazain Khan[1], Peter Sage[1], Kalyan Pooja[1], Hana Berhe[1], Sophia Hopkins[1], Yu-Ang Chen[1], Yusuke Nakashima[1], Takeo Watanabe[1], Yuka Sasaki[1]; [1]Brown University

While it is established that visual perceptual learning (VPL) is associated with changes in visual areas, the roles of the prefrontal cortex (PFC) in regulating top-down signals for VPL remain elusive. This study specifically investigates how the PFC unconsciously controls top-down signals during sleep to reduce anterograde and retrograde interferences between VPL of the two distinct visual tasks trained before and after sleep. Before and after a 90-minute sleep session inside an MRI scanner with polysomnography, two interfering texture discrimination tasks (TDT) were trained. We tested whether training of pre-sleep and post-sleep TDTs interfered with each other. As our prior research (Nature Neuroscience, 2020) demonstrated a significant correlation between the concentration of excitatory-to-inhibitory neurotransmitters (E/I ratio) in the visual cortex during REM sleep and resilience to interference, we measured E/I ratios during non-REM and REM sleep in the medial prefrontal cortex (mPFC) and dorsolateral prefrontal cortex (DLPFC) regarding interference. Subjects who exhibited both NREM and REM sleep demonstrated greater resilience to retrograde interference (from post-sleep TDT to pre-sleep TDT) compared to those who showed NREM sleep alone. In mPFC, the E/I ratio significantly reduced from baseline during REM sleep in correlation with resilience to retrograde interference. However, resilience to anterograde interference (from pre-sleep TDT to post-sleep TDT) was correlated with the E/I ratio during NREM sleep. DLPFC exhibited no significant correlations between E/I ratios and resilience to interference anterogradely or retrogradely, while increases in the E/I ratio in DLPFC during NREM sleep from baseline were significantly correlated with offline performance gains in pre-sleep TDT. These findings indicate the involvement of the mPFC in both anterograde and retrograde interferences but during distinct sleep stages. Despite the prevailing belief that consciousness is predominantly involved in prefrontal controls, our findings suggest that prefrontal control mechanisms operate even during unconscious states such as sleep.

NIH (R01EY031705, R01EY019466, R01EY027841), KAKENHI (JP20KK0268).

Talk 7, 6:45 pm, 35.17

REVERSAL LEARNING IN THE HUMAN VISUAL CORTEX

Caitlin, M Traiser, Richard, T Ward, Hannah, M Engle, Andreas Keil; [1]University of Florida

Reversal learning paradigms are commonly used to investigate cognitive and affective processes, including in neuropsychiatric conditions. The present study uses a novel aversive reversal learning paradigm to investigate visuocortical responses to threat stimuli, with prior research focusing on limbic and frontocortical regions. Participants (N = 44; 18-23 years) viewed flickering Gabor patches at different orientations, driving steady-state visual evoked responses (ssVEP) recorded with EEG. An aversive loud noise was used as the unconditioned stimulus, consistently paired with one orientation (the CS+) and never with the other (CS-). After the initial acquisition phase, the contingency between the conditioned and unconditioned stimuli was reversed. Test phases following initial acquisition and reversal examined frequency-tagged ssVEPs evoked by CS+ and CS-, as well as a neutral accompanying Gabor, allowing us to quantify competition effects as a function of learning. Participants were asked to rate each stimulus in terms of valence, arousal, and expectancy before, during, and after learning. Continuous EEG was recorded using a saline EEG system with 129 electrodes/sensors and artifact-free trials were analyzed in the frequency domain, using the Discrete Fourier Transform, after averaging trials by condition. Statistical analyses were conducted using Matlab. We compared the ssVEP amplitude at the tagging frequencies during the critical test phases, across the entire topography. As expected, the ssVEP evoked by the conditioned threat cue (CS+) was enhanced over posterior sites, compared to the CS-, after the initial acquisition phase. Importantly, this effect reversed after 60 trials of reversal learning, and increased in effect size: The new CS+ (the former CS-) prompted selectively heightened ssVEP signals compared to the new CS-. Findings support the notion that experience changes the amplitude of neural mass activity in human visual cortex. They also show that these changes are malleable, adapting and even reversing with environmental contingencies.

The research was supported by NIH grant R01MH125615 to Andreas Keil, PhD.

Talk 8, 7:00 pm, 35.18

VISUAL PERCEPTUAL LEARNING COMPLETELY TRANSFERS TO A NEW LOCATION WHEN PHASE, INSTEAD OF CONTRAST, IS VARIED DURING TRAINING

Beyza Akkoyunlu[1,2,3,4] (bakkoyunlu@dpz.eu), Caspar Schwiedrzik[1,2,3]; [1]Perception and Plasticity Group, German Primate Center, Göttingen, Germany, [2]Neural Circuits and Cognition Lab, European Neuroscience Institute Göttingen – A Joint Initiative of the University Medical Center Göttingen and the Max-Planck-Society, [3]Leibniz ScienceCampus Primate Cognition, Göttingen, Germany, [4]Graduate School for Neuroscience, Biophysics and Molecular Biosciences (GGNB), Göttingen

Performance on visual tasks can be improved via training or experience, and this is called visual perceptual learning (VPL). However, this improved performance is limited to the trained task's specifics, i.e. when the spatial position of the stimulus is changed, the improvement disappears. Yet, recent research shows that variability along task-irrelevant stimulus dimensions can alter this characteristic. We argue that variability determines which neurons undergo plasticity in VPL, and depending on these neurons' invariance properties, generalization or specificity is achieved. We trained two groups of participants with almost identical tasks, only changing which task-irrelevant dimension varied between trials. In particular, we created variability by randomizing spatial phase in one training group and contrast in the other. After training, we tested for transfer to a new spatial location in both groups. Phase-invariant neurons emerge later in the visual processing hierarchy, compared to contrast-invariant neurons (e.g., complex, and simple cells), and hence phase-invariant neurons have larger spatial receptive fields. Due to this, we hypothesized that varying the phase of the training stimuli over trials would give rise to generalization in space due to neurons which are phase invariant taking a role in the training. On the other hand, as contrast-invariant neurons appear earlier in the hierarchy, we expected the learning to be more specific when participants were trained with varying contrast. We found that the randomizing phase of the training stimulus resulted in complete generalization of the improvement to a new spatial location, contrary to randomizing contrast. Our results show that which neural populations undergo plasticity with VPL is determined by the training task demands, and in turn, this affects generalization and specificity of behavioral improvements.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 802482, named "VarPL")

Talk Session: Sunday, May 19, 2024, 5:15 – 7:15 pm, Talk Room 2

Eye Movements: Early visual processing

Moderator: Miriam Spering, University of British Columbia

Talk 1, 5:15 pm, 35.21

INVESTIGATING THE RELATIONSHIP BETWEEN HUMAN FOVEAL ANATOMY AND FIXATION BEHAVIOR ACROSS DIFFERENT VISUAL TASKS

Benjamin Moon[1] (b.moon@rochester.edu), Ashley M. Clark[1], Krishnamachari Prahalad[1], Austin Roorda[2], Pavan Tiruveedhula[2], Wolf Harmening[3], Aleksandr Gutnikov[3], Samantha K. Jenks[1], Sanjana Kapisthalam[1], Michele Rucci[1], Jannick P. Rolland[1], Martina Poletti[1]; [1]University of Rochester, [2]University of California, Berkeley, [3]University of Bonn

When looking at fine details, normally sighted observers center stimuli on a specific location in the foveola, the preferred retinal locus. This locus is slightly offset from the point of peak cone density and has been reported to remain consistent between tasks. However, differences in fixational oculomotor behavior are known to occur across tasks, which may lead to fine changes in the average stimulus position on the retina. Here we examined the distribution of the retinal positions of the fixated stimulus in different tasks in relation to cone density across the foveola. Using an adaptive optics scanning laser ophthalmoscope for simultaneous stimulus delivery and retinal imaging, we investigated oculomotor behavior of subjects (N=8) in three different tasks: fixation on a blinking square, a moving Maltese Cross, and a high-acuity Snellen task. Stimulus size was the same across tasks. We then quantified the difference between the average stimulus position on the retina across the tasks with respect to the location where cones are most densely packed. Peak cone densities across subjects varied from 13,847 to 20,897 cones per square degree, and the stimuli remained within the region where cone density was above 70% of the peak density in all conditions. This region spanned on average 0.23 degrees squared. Yet, differences across tasks were present. We found that the stimulus positions in the Snellen task spanned an area that was more than 50% larger on average when compared to either fixation task ($p<0.05$). Further, we found a small consistent shift (1.5+/-1.1 arcmin, $p<0.033$) toward the location of peak cone density in the Snellen task compared to the blinking square fixation. Although the mechanism responsible for the observed shift remains unclear, it raises interesting questions, as it does not yield a significant difference in the Nyquist sampling frequency.

Supported by NIH grants EY029788, EY018363, EY001319, and EY023591 and German Research Foundation grants Ha5323/6-1 and Ha5323/8-1

Talk 2, 5:30 pm, 35.22

MINIMAL RETINAL SLIP IS SUFFICIENT FOR PEAK VISUAL ACUITY IN THE FOVEA

Veronika Lukyanova[1], Julius Ameln[1], Aleksandr Gutnikov[1], Jenny Lorén Witten[1], Bilge Sayim[2], Wolf Harmening[1]; [1]Department of Ophthalmology, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany, [2]Sciences Cognitives et Sciences Affectives (SCALab), CNRS, UMR 9193, University of Lille, Lille, France

Humans can resolve visual targets beyond the Nyquist limit of photoreceptor sampling, likely because fixational eye movements (FEM) induce optimal retinal signals. The retinal space covered by a visual stimulus and hence the number of photoreceptors involved in a visual task grows with time, as FEM continuously move the retina across the image. Here, we determined the minimum amount of retinal slip required for achieving maximum visual performance in two visual acuity tasks. Using adaptive optics scanning laser ophthalmoscopy-based microstimulation, Tumbling-E and Two-dot stimuli were displayed in the foveola of two experienced and two naive participants. Therefore, resolution and positional acuity were measured and the cone photoreceptors used in the task were made visible. The amount of naturally occurring retinal slip was manipulated by varying stimulus duration (2-600 ms) and by additionally restricting it in one of the viewing conditions using real-time retinal stabilization. As expected, slip amplitudes increased linearly with increasing stimulus duration in all participants, at individual rates. Drift velocity and the covered area varied across participants, both linked to the individuals' cone density. Across participants, we found maximum acuity between 0.9 and 1.13 of the cone Nyquist limit for resolution and between 0.26 and 0.57 for positional stimuli. Surprisingly, this was achieved after very short presentation times (80 ms), after which thresholds did not improve significantly. Within 80 ms, drift amplitudes ranged from 0.8 - 1.6 arcmin, equivalent to an absolute retinal slip of 1.5 - 2.4 cone diameters, depending on the participant. On average, resolution thresholds were unaffected by retinal stabilization. During short stimulus durations, positional acuity was better when presented retinally stabilized. These results demonstrate that the human visual system can extract spatial information during time frames that do not allow extended motion paths, and that natural motion is not required to reach maximum performance.

Wolf M. Harmening: German Research Foundation, Ha 5323/8-1.
Bilge Sayim: Agence Nationale de la Recherche, ANR-19-FRAL-000.

Talk 3, 5:45 pm, 35.23

ECCENTRICITY, BUT NOT COLOR CONTRAST, INFLUENCES MICROSACCADIC PREVENTION OF VISUAL FADING

Max Levinson[1], Christopher C. Pack[1], Sylvain Baillet[1]; [1]McGill University

When the eyes remain still for an extended period of time, visual boundaries can appear to fade. This illusory phenomenon is traditionally attributed to slow neuronal adaptation to stable retinal input. Microsaccades – small, mostly involuntary ocular movements during gaze fixation – counteract fading, but it is unclear exactly how. They might simply refresh the retinal image to reverse adaptation or introduce a unique visual resampling signal that interacts with adapting neural populations. To better understand boundary fading both in itself and as a probe of microsaccade function, we investigated why a stronger boundary, determined by isoluminant color contrast and eccentricity (distance from the visual field center), takes longer to fade. A stronger boundary could either create a more robust cortical signal requiring greater adaptation, and/or enhance the preventive effect of microsaccades. To test these two possibilities we recorded microsaccades during a perceptual filling-in task. Human participants fixated centrally until they experienced illusory merging of two isoluminant colored surfaces, separated by a circular ring boundary, due to perceptual boundary fading. We fit linear mixed models of trial-wise fading time to color contrast, boundary eccentricity, and fixational eye movement dynamics. While both color contrast and eccentricity altered fading time, they did so differently. Higher color contrast extended overall fading time but did not influence the efficacy of individual microsaccades. Conversely, lower eccentricity prolonged fading exclusively via eye movements. These findings demonstrate that stimulus properties can differently influence slow adaptation and microsaccadic counteraction. We propose that microsaccades prevent visual fading via transient boundary stimulation that scales with cortical magnification but is invariant to boundary contrast.

Talk 4, 6:00 pm, 35.24

HOW THE UNSTABLE EYE SEES A STABLE AND MOVING WORLD: REDUX

Josephine D'Angelo[1] (josephine_dangelo@berkeley.edu), Pavan Tiruveedhula[1], Raymond J. Weber[2], David W. Arathorn[2], Austin Roorda[1]; [1]University of California, Berkeley, [2]Electrical & Computer Engineering, College of Engineering, Montana State University

Despite incessant retinal image motion, humans reliably perceive stable and moving objects in the world. In 2013, Arathorn et al. found that the direction of stimulus motion with respect to fixational eye motion impacts the amount of motion perceived: Subjects perceived motion when stimuli moved in the same direction as eye movements and perceived little to no motion when images moved directly opposite the direction of eye motion, surprisingly even under conditions where that motion was amplified. This suggests that the visual system computes its direction of motion and perceptually renders anything moving in a direction opposite to that eye motion as stable. We asked: How is the direction of eye motion determined? Is world-fixed retinal image background content needed to compute the direction of motion or are non-visual cues (eg. efferent copy) sufficient? We explored this question using an adaptive optics scanning laser ophthalmoscope which performs high resolution eye tracking and delivers retina-contingent stimuli. We quantified perceived motion using a method of adjustment where the subject compared a stimulus moving contingent to the retina and a stimulus moving on a random walk. The subject adjusted the diffusion constant of the random walk stimulus until the perceived motion of both stimuli looked equal. Experiments were done with retinal image background content and with all visual content removed. In the presence of retinal image background content, the perceived motion of stimuli moving with a 2.5X exaggerated retinal slip was suppressed; however, in conditions with no visual content we found that a higher magnitude of motion was perceived. Our results suggest that the presence of retinal image background content provides the primary signal for the visual system to compute its direction of motion, contradicting the intuition that the same content would provide a frame of reference to see that motion.

NIH BRP R01EY023591; NIH T32 EY007043

TALK 5, 6:15 PM, 35.25

SOUND ACTIVATES A DORMANT VISUAL-MOTOR PATHWAY BYPASSING PRIMARY VISUAL CORTEX

Tatiana Malevich[1,2,3] (tatiana.malevich@cin.uni-tuebingen.de), Yue Yu[1,2,3], Matthias P. Baumann[1,2,3], Tong Zhang[1,2,3], Ziad M. Hafed[1,2,3]; [1]Hertie Institute for Clinical Brain Research, Tuebingen, Germany, [2]Werner Reichardt Centre for Integrative Neuroscience, Tuebingen, Germany, [3]University of Tuebingen

Like in other species, the primate visual system contains an anatomical retinal projection bypassing the geniculostriate pathway and innervating the midbrain. However, unlike in some of these species, the functional significance of this alternative visual pathway remains unknown: increasing evidence suggests that it may be completely dormant. We first tested this by performing focal, reversible inactivation of the primary visual cortex (V1) and investigating a short-latency oculomotor reflex believed to rely on subcortical eye-movement control circuits. This reflex, called saccadic inhibition (recently reviewed by Buonocore and Hafed, 2023), is characterized by a short-latency inhibition of saccade generation by visual stimuli, as well as by a concomitant saccade direction biasing, first towards and then away from stimulus location. When we created a localized cortical scotoma, saccadic inhibition was completely abolished for stimuli in the blind field, confirming the geniculostriate pathway's dominance. Superior colliculus visual responses were also eliminated. However, why does the alternative visual pathway, directly targeting oculomotor control circuits, exist at all? We hypothesized that this pathway might still be functional, albeit in a gated manner. During V1 inactivation, we paired a visual onset with a sound pulse (50 ms; 1 KHz; suprathreshold) that was completely uninformative about the visual stimulus' location. Saccadic inhibition was partially restored, and it was different to when the sound pulse occurred alone. Most importantly, there was a re-emergence of saccade direction biasing towards the visual stimulus location, even though the sound was not spatially informative. Guessed visually-guided saccades towards a target presented in the blind field were also mildly more accurate with the uninformative sound. These results demonstrate that multi-sensory information can activate an otherwise dormant visual-motor pathway. These results also highlight the importance of multi-species comparisons of hierarchical sensory-motor processes, and they especially inform models of active sensory-guided behavior invoking parallel processing streams.

TALK 6, 6:30 PM, 35.26

TEMPORAL DYNAMICS OF SERIAL DEPENDENCE IN OCULAR TRACKING

Bao Hong[1,3] (bh2378@nyu.edu), Jing Chen[2,3], Li Li[2,3]; [1]East China Normal University, Shanghai, China, [2]New York University Shanghai, Shanghai, China, [3]NYU-ECNU Institute of Brain and Cognitive Science at New York University Shanghai, Shanghai, China

An outstanding question regarding serial dependence is the level at which it operates—some argue for its direct influence on perception, while others posit that serial dependence only impacts post-perceptual processes. Here we examined when serial dependence appears and disappears to address this issue. We developed an ocular tracking task consisting of distinct temporal phases of tracking (retinal-motion-driven pursuit initiation vs. extraretinal-signal-supported steady-state tracking), coupled with high-resolution eye movement recordings, to provide an ideal testing paradigm for this purpose. Participants (N=16) tracked the step-ramp motion of a target spot (diameter: 0.64°; speed: 16°/s; direction: randomly drawn from the full 360° circle at the step size of 12°). We performed model-dependent analyses to measure the extent to which pursuit directions through the course of the current trial were affected by the previous target moving direction, with positive values indicating attraction and serial dependence and negative values indicating repulsion and adaptation. We observed a strong serial dependence at pursuit initiation that quickly declined over time, followed by a low-amplitude adaptation that remained stable throughout steady-state tracking. This result shows that serial dependence happens before adaptation in ocular tracking, providing evidence that serial dependence influences perception. We also found a strong correlation ($r=0.88$) between the strength of serial dependence and pursuit direction noise. Using a Bayesian observer model constrained by efficient coding, we further found that the temporal dynamics of serial dependence can be predicted by pursuit direction noise over time. This supports our proposal that the visual system may strike a balance between utilizing the temporal continuity of the physical environment (through serial dependence) and optimizing sensitivity to subtle changes in the environment (through adaptation), and this balance is regulated by sensory noise.

Supported by research grants from the National Natural Science Foundation of China (32071041, 32161133009); China Ministry of Education (ECNU 111 Project, Base B1601); the major grant seed fund and the boost fund from NYU Shanghai

Talk 7, 6:45 pm, 35.27

IDENTIFYING THE NEURAL ORIGINS OF PUPIL CONSTRICTIONS IN RESPONSE TO ISOLUMINANT STIMULI USING CONTRAST ADAPTATION

Vasilii Marshev[1] (vasilimarshev@gmail.com), Haley Frey[2], Jan Brascamp[1]; [1]MSU, [2]UC Berkeley

The events causing transient pupil constriction stretch beyond increased illumination. For example, changes in stimulus structure (orientation, color, motion) can cause constriction even if net illumination stays the same. Two explanations for this have been proposed: some argue that altered cortical feature-based responses cause constriction via their input to pupil control nuclei, others argue that local increases in retinal receptor drive a pupil light reflex even if net illuminance does not increase. We tested the two theories using orientation-specific adaptation, which is thought to have a cortical origin. Accordingly, reduced pupil constrictions following orientation-specific adaptation would support a cortical origin of these constrictions. Subjects' pupil size was recorded while oblique test gratings were presented following presentation of a sliding high-contrast (100%) adapter grating. A test grating's orientation could be either parallel or orthogonal to that of the adapter that preceded it. In Experiment 1, low-contrast test gratings, with a net luminance equal to that of the background, caused a constriction, but this response was the same for parallel and orthogonal gratings, even though perceptual judgments confirmed the presence of orientation-specific adaptation. In Experiment 2, test contrast was increased to produce more robust constrictions. Still, no orientation-specific modulation was observed. In Experiment 3 we used a test grating in which no pixel had a higher luminance than the background, thus minimizing the possibility of local increases in retinal receptor drive to light reflex. We again observed a transient pupil constriction but no orientation-specific modulation. While our results confirm that the pupil can constrict in response to stimuli that involve no global (Experiments 1-2) or even local (Experiment 3) luminance increases, they do not support the idea that these constrictions reflect altered cortical feature-based responses.

Talk 8, 7:00 pm, 35.28

EFFECTS OF PUPIL SIZE AS MANIPULATED THROUGH IPRGC ACTIVATION ON VISUAL PROCESSING

Sebastiaan Mathôt[1] (s.mathot@rug.nl), Hermine Berberyan[2], Philipp Büchel[1], Veera Ruuskanen[1], Ana Vilotjević[1], Wouter Kruijne[1]; [1]Department of Psychology, University of Groningen, The Netherlands, [2]PwC, Amsterdam, The Netherlands

The size of the eyes' pupils determines how much light enters the eye and also how well this light is focused. Through this route, pupil size shapes the earliest stages of visual processing. Yet causal effects of pupil size on vision are poorly understood and rarely studied. Here we introduce a new way to manipulate pupil size, which relies on activation of intrinsically photosensitive retinal ganglion cells (ipRGCs) to induce sustained pupil constriction. We report the effects of both experimentally induced and spontaneous changes in pupil size on visual processing as measured through EEG, and compare these to the effects of stimulus intensity and covert visual attention, because previous studies have shown that these factors all have comparable effects on some common measures of early visual processing. Using a mix of neural-network decoding, ERP analyses, and time-frequency analyses, we find that induced pupil size, spontaneous pupil size, stimulus intensity, and covert visual attention all affect EEG responses, mainly over occipital and parietal electrodes, but—crucially—that they do so in qualitatively different ways. Induced and spontaneous pupil-size changes mainly modulate activity patterns (but not overall power or intertrial coherence) in the high-frequency beta range; this may reflect an effect of pupil size on oculomotor activity and/ or visual processing. In addition, spontaneous (but not induced) pupil size tends to correlate positively with intertrial coherence in the alpha band; this may reflect a non-causal relationship, mediated by arousal. Taken together, our findings suggest that pupil size has qualitatively different effects on visual processing from stimulus intensity and covert visual attention. This suggests that pupil size strongly affects visual processing, and provides concrete starting points for further study of this important yet understudied earliest stage of visual processing.

Talk Session: Monday, May 20, 2024, 8:15 – 9:45 am, Talk Room 1

Eye Movements: Perception and timing

Moderator: Markus Lappe, University of Muenster

Talk 1, 8:15 am, 41.11

DECODING REMAPPED SPATIAL INFORMATION IN THE PERI-SACCADIC PERIOD

Caoimhe Moran[1,2] (caoimhe.moran@mail.huji.ac.il), Philippa A. Johnson[3], Ayelet N. Landau[2], Hinze Hogendoorn[4]; [1]The University of Melbourne, [2]The Hebrew University of Jerusalem, [3]Leiden University, [4]Queensland University of Technology

It has been suggested that, prior to a saccade, visual neurons predictively respond to stimuli that will fall in their receptive fields after completion of the saccade. This saccadic remapping process is thought to compensate for the shift of the visual world across the retina caused by eye movements. To map the timing of this predictive process in the brain, we recorded neural activity using electroencephalography (EEG) during a saccade task. Participants made saccades between two fixation points while covertly attending to oriented gratings briefly presented at various locations on the screen. Data recorded during trials in which participants maintained fixation were used to train classifiers on stimuli in different positions. Subsequently, data collected during saccade trials were used to test for the presence of remapped stimulus information at the post-saccadic retinotopic location in the peri-saccadic period, providing unique insight into when remapped information becomes available.

We found that the stimulus could be decoded at the remapped location ~180 ms post-stimulus onset, but only when the stimulus was presented 100-200 ms before saccade onset. Within this range, we found that the timing of remapping was dictated by stimulus onset rather than saccade onset. We conclude that presenting the stimulus immediately before the saccade allows for optimal integration of the corollary discharge signal with the incoming peripheral visual information, resulting in a remapping of activation to the relevant post-saccadic retinotopic neurons.

TALK 2, 8:30 AM, 41.12

PERCEPTION OF CONTINUOUS FLICKER: PHANTOM ARRAY VERSUS MOVING STIMULI

Rytis Stanikunas[1] (rytis.stanikunas@ff.vu.lt), Alvydas Soliunas[1], Remigijus Bliumas[1], Karolina Jocbalyte[1], Algirdas Novickovas[1]; [1]Vilnius University

During saccadic eye movements, an image is displaced on the retina, and the visual system must recalculate its position to maintain perceptual space constancy. However, when the flickering light is presented during the saccade, a phenomenon of the phantom array of lights is perceived. A similar array of lights can be perceived when a moving and flickering stimulus is presented during visual fixation. Here, we investigated differences in spatial and temporal aspects of the perception of flickering light during saccade versus flickering light movement during visual fixation. In the first experiment, the subjects made a saccade across a point light source flashing from 50 Hz to 4 kHz. In the second experiment, a moving and flickering stimulus was presented on the screen while the subjects maintained steady visual fixation. The speed of the stimulus was set to the same as the saccade speed for each subject. Subjects were asked to indicate the beginning and end of the array of lights, to evaluate the length of one dash, and to count the number of dashes. We found that the perceived length and localization of the moving lights array approximately corresponded to the physical representation of the stimulus on the retina, but during the saccade, a shorter length of the phantom array was perceived, and localization greatly varied between subjects. The phantom array was always perceived as composed of a smaller number of dashes compared to the moving lights array. However, the size of one dash was perceived to be of similar length as projected on the retina during both conditions. Therefore, we can state that the visual space is not compressed in size, but is compressed in time during saccades. The visual system reduces information flow by quantization mechanism and removes some repeatable representations of the same object from perceptual space.

This work was supported by the Research Council of Lithuania S-MIP-21-56

TALK 3, 8:45 AM, 41.13

TIMING OF EYE AND HAND MOVEMENTS DURING REACHING DEPENDS ON FUNCTIONAL DEMANDS OF GAZE

Jolande Fooken[1] (jolande.fooken@queensu.ca), Nethmi H. Illamperuma[1], J. Randall Flanagan[1]; [1]Queen's University

When reaching to visual targets, people are unable to shift their gaze away from the reach target to a secondary gaze target until after the reach target has been attained—a phenomenon known as gaze anchoring. Here, we compared gaze anchoring when reaching to a visual target versus a visual-haptic target providing force feedback upon contact. We also examined gaze anchoring in a bimanual context in which participants were instructed to shift their gaze to the secondary target as soon as it appeared and, at the same time, move their other hand to the secondary target. In our task, human participants (n=28) used their right hand to move the handle of a robotic manipulandum to a primary visual or visual-haptic reach target. A secondary target was present at the beginning, halfway, or end of the reaching movement and participants were instructed to make either an eye movement (unimanual trials) or a combined eye and left hand movement (bimanual trials) to this target as soon as it appeared. We found that in unimanual trials with visual targets, saccades were initiated ~125 ms after the hand cursor 'visually contacted' the reach target. In contrast, with visual-haptic targets, saccades were initiated around the time of contact. This suggests that when haptic feedback was provided, central vision was not critical for guiding the hand as it approached the target or checking target attainment. However, gaze anchoring was still observed with visual-haptic targets earlier in the reach when gaze was engaged in directing the hand towards the target. In bimanual trials, gaze anchoring was observed but anchoring did not extend to the left hand, the onset of which was decoupled from gaze. Overall, our findings indicate that the timing of eye and hand movements in object manipulation is linked to the function of target fixations.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) Research Fellowships Grant FO 1347/1-1 (JF). We thank Elissa Robichaud and Bethany Piekkola for their help with data collection and Martin York for technical support.

TALK 4, 9:00 AM, 41.14

PREDICTIVE LOOKING AND PREDICTIVE LOOKING ERRORS IN EVERYDAY ACTIVITIES

Sophie Su[1] (s.sophie@wustl.edu), Matthew Bezdek[3], Tan Nguyen[1], Christopher Hall[2], Jeff Zacks[1]; [1]Washington University in Saint Louis, [2]University of Virginia, [3]Elder Research

Where people look in pictures and movies has been shown to be based not only on the most salient point in the current scene, but also on predictions of what is going to happen next. The accuracy of these predictions fluctuates during movie watching. Some theories of event comprehension propose that spikes in prediction error can trigger working memory updating and the segmentation of ongoing experience into meaningful events. One previous study of predictive looking found evidence for this proposal (Eisenberg et al., 2018, CR:PI), but the paradigm used in that study could only obtain predictions intermittently, because it analyzed predictive looking to objects that an actor was about to contact. Here, we developed a continuous measure of prediction error by modeling predictive looking towards the actor's hands, and we operationalized this prediction error as the residuals from the predictive looking model. Viewers' gaze was tracked while they watched movies of everyday activities, and mixed-effects models were used to predict the actor' hand positions from viewers' previous gaze location. Stepwise model comparison indicated that viewers look predictively as current gaze position accounts for hand location as far as 9 seconds in the future. We compared the time course of gaze predictions with that of predictions generated from a computational model of event comprehension and found that gaze predictions showed higher error at moments when the computational model had higher errors. Furthermore, spikes in gaze prediction error were predictive of increases in event segmentation in a separate group of viewers. These results support proposals that event segmentation is driven by spikes in prediction error, and this method promises to give a general approach for measuring ongoing prediction error noninvasively.

Talk 5, 9:15 am, 41.15

SACCADES TO PARTIALLY OCCLUDED OBJECTS: PERCEPTUAL COMPLETION MEDIATES OCULOMOTOR CONTROL

Michael Paavola[1] (michael-paavola@uiowa.edu), Andrew Hollingworth[1], Cathleen Moore[1]; [1]University of Iowa

Oculomotor behavior is ultimately controlled by patterns of activity in retinotopically organized populations of neurons in areas, such as the superior colliculus and frontal eye fields, that have visuomotor receptive fields. In contrast, gaze is guided by non-retinotopic variables including task goals, attentional state, and the perceived 3 dimensional structure of the environment. We investigated how the implied extent of perceptually completed surfaces behind occluding surfaces impacts saccade landing position while searching for small targets. Each trial included four disks and four truncated disks. On half of the trials, rectangles abutted the truncated disks supporting the perception of completed disks behind occluding surfaces. Observers searched among the disks for small red or green dots, which appeared only when a saccade landed within a disk region. This design leveraged the tendency for saccades to land near the center of objects (e.g., Melcher & Kowler, 1999) to ask what constitutes an “object” to the eye-movement control system, the perceptually extended whole disk or the optically explicit truncated disk? Experiment 1 showed that distributions of landing position were biased toward the center of the implied whole disks and away from the optically explicit portion of the disk when occluders were present. Experiment 2 showed the same bias toward the center of the whole disk despite the location of the colored dot being presented at the center of the image region, which would have given a strategic advantage to use the image level representation. Experiment 3 used complementary contrast regions to demonstrate that the landing position bias shown in Experiments 1 and 2 were not due to low-level stimulus interactions caused by the presence of occluders. Taken together, these results indicate that oculomotor control mechanisms operate over object-level representations during the planning and execution of eye movements.

Talk 6, 9:30 am, 41.16

PERCEIVING THE SELF-GENERATED MOTION ON THE RETINA CAUSED BY SMOOTH PURSUIT

Krischan Koerfer[1] (krischan.koerfer@uni-muenster.de), Tamara Watson[2], Markus Lappe[1]; [1]University of Muenster, [2]Western Sydney University

Visual perception in humans is intermingled with eye movements. Despite the self-generated motion on the retina by smooth pursuit and saccades, we perceive a stable world, a marvelous achievement by the visual system. We developed a novel stimulus that leads to loss of visual stability across saccades and that is perceived differently if pursued, highlighting the limitations of the visual systems to compensate for eye movements and providing new insight into the underlying mechanisms. The stimulus consisted of a random dot distribution. Across frames, dots in a circular zone rotated to create a vortex motion. Independent of the first-order motion within it, the vortex then moved across the screen. We formerly reported that the vortex cannot be pursued smoothly and that tracking the vortex with frequent catch-up saccades causes a loss of visual stability. Here, we altered the vortex to make it pursuable by dislocating the dots once they became part of the motion pattern and when leaving the vortex, creating a slim ring of flickering and discontinuity around the vortex. Once participants were able to pursue the altered vortex, this also restored visual stability. Interestingly, successful smooth pursuit also changed the perception of the vortex motion pattern. When asked to identify the formerly observed motion pattern in a discrimination task, participants more often chose the pattern with additional first-order motion congruent with the motion pattern’s movement across the screen than the correct pattern. This contrasted sharply with trials involving the unaltered vortex, where participants mostly identified the correct pattern. Consequently, this indicates that motion patterns are perceived based on the retinal image, rather than their actual presentation on the screen, uncovering a novel interaction between smooth pursuit and perception.

This work was supported by the German Research Foundation (DFG La 952-7) and has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 734227.

Talk Session: Monday, May 20, 2024, 8:15 – 9:45 am, Talk Room 2

Object Recognition: Neural mechanisms

Moderator: Diane Beck, University of Illinois

Talk 1, 8:15 am, 41.21

VERNIER ACUITY IN SINGLE NEURONS OF MONKEY INFERIOR TEMPORAL CORTEX

Sini Simon[1] (sini.simon@gmail.com), SP Arun[1]; [1]Indian Institute of Science

Humans can discriminate small offsets between nearly collinear lines, a phenomenon known as vernier acuity. However, the underlying neural correlates have not been investigated. To investigate this issue, we asked whether monkeys experience visual acuity like humans and which brain regions form the underlying neural basis. We created stimuli containing a square frame with a disk that could be moved along a horizontal or vertical line, in the presence of a horizontal or vertical bar at the center. If vernier acuity is present in behavioural or neural responses, we should observe greater sensitivity to small changes in the horizontal position of the disk when a nearby bar is oriented vertically rather than horizontally. Conversely, there should be greater sensitivity to changes in vertical position when the nearby bar is oriented horizontally but not vertically. We tested these predictions on monkeys performing a same-different task, as well as using neural responses recorded from their inferior temporal cortex, while they passively viewed the same stimuli. In Experiment 1, we tested 3 monkeys trained to perform a same-different task. Here, all three animals showed higher sensitivity to position changes in the vernier conditions compared to the non-vernier conditions. In Experiment 2, we tested these predictions using wireless brain recordings from the inferior temporal cortex, a region critical for object recognition, while monkeys viewed these stimuli in a fixation task. Here too, we observed greater neural dissimilarity between the stimuli in the vernier condition compared to the non-vernier condition. Interestingly, this effect arose late in the neural response, suggesting that this effect arises through computation and is not simply inherited from the early visual areas. Taken together, our results show that monkeys, like humans, experience vernier acuity and this effect is likely driven by single neurons in the inferior temporal cortex.

This work was supported through a Senior Fellowship from the DBT-Wellcome India Alliance to SPA.

TALK 2, 8:30 AM, 41.22

DOES LEVERAGING THE HUMAN VENTRAL VISUAL STREAM IMPROVE NEURAL NETWORK ROBUSTNESS?

Zhenan Shao[1,2] (zhenans2@illinois.edu), Linjian Ma[3], Bo Li[3,4], Diane M. Beck[1,2]; [1]Department of Psychology, University of Illinois Urbana-Champaign, [2]Beckman Institute, University of Illinois Urbana-Champaign, [3]Department of Computer Science, University of Illinois Urbana-Champaign, [4]Department of Computer Science, University of Chicago

Human object recognition is robust to a variety of object transformations, including changes in lighting, rotations, and translations, as well as other image manipulations, including the addition of various forms of noise. Invariance has been shown to emerge gradually along the ventral visual stream with later regions showing higher tolerance to object transformations. In contrast, despite their unprecedented performance on numerous visual tasks, Deep Neural Networks (DNNs) fall short in achieving human-level robustness to image perturbations (adversarial attacks), even those that are visually imperceptible to humans. One potential explanation for this difference is that brains, but not DNNs, build increasingly disentangled and therefore robust object representations with each successive stage of the ventral visual stream. Here, we asked whether training DNNs to emulate human representation can enhance their robustness and, more importantly, whether different stages of the ventral visual stream enable progressively increased robustness, reflecting the potentially evolving representation crucial for human perceptual invariance. We extracted neural activity patterns in five hierarchical regions of interest (ROIs) in the ventral visual stream: V1, V2, V4, LO, and TO from a 7T fMRI dataset (Allen et al., 2022) obtained when human participants viewed natural images. DNN models were trained to perform image classification tasks while aligning their penultimate layer representations with neural activity from each ROI. Our findings reveal not only a significant improvement in DNN robustness but also a hierarchical effect: greater robustness gains were observed when trained with neural representations from later stages of the visual hierarchy. Our results not only show that ventral visual cortex representations improve DNN robustness but also support the gradual emergence of robustness along the ventral visual stream.

This work used NCSA Delta GPU through allocation SOC230011 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296.

TALK 3, 8:45 AM, 41.23

HUMAN EEG AND ARTIFICIAL NEURAL NETWORKS REVEAL DISENTANGLED REPRESENTATIONS OF OBJECT REAL-WORLD SIZE IN NATURAL IMAGES

Zitong Lu[1] (lu.2637@osu.edu), Julie D. Golomb[1]; [1]The Ohio State University

Remarkably, human brains have the ability to accurately perceive and process the real-world size of objects, despite vast differences in distance and perspective. While previous studies have delved into this phenomenon, distinguishing this ability from other visual perceptions, like depth, has been challenging. Using the THINGS EEG2 dataset with high time-resolution human brain recordings and more ecologically valid naturalistic stimuli, our study uses an innovative approach to disentangle neural representations of object real-world size from visual size and perceived real-world depth in a way that was not previously possible. Leveraging this state-of-the-art dataset, our EEG representational similarity results reveal a pure representation of object real-world size in human brains. We report a representational timeline of visual object processing: pixel-wise differences appeared first, then real-world depth and retinal size, and finally, real-world size. Additionally, we input both these naturalistic images and object-only images without natural background into artificial neural networks. Consistent with the human EEG findings, we also successfully disentangled representation of object real-world size from visual size and real-world depth in all three types of artificial neural networks (visual-only ResNet, visual-language CLIP, and language-only Word2Vec). Moreover, our multi-modal representational comparison framework across human EEG and artificial neural networks reveals real-world size as a stable and higher-level dimension in object space incorporating both visual and semantic information. Our research provides a detailed and clear characterization of the object processing process, which offers further advances and insights into our understanding of object space and the construction of more brain-like visual models.

NIH R01-EY025648 (JG), NSF 1848939 (JG)

TALK 4, 9:00 AM, 41.24

VISUAL PROCESSING OF SOFT OBJECTS AUTOMATICALLY ACTIVATES PHYSICS-BASED REPRESENTATIONS IN THE HUMAN BRAIN

Wenyan Bi[1] (wenyan.bi@yale.edu), Qi Lin[2], Kailong Peng[1], Aalap Shah[1], Ilker Yildirim[1]; [1]Yale University, [2]Riken

When encountering soft objects, say a garment draping on a surface or a pillow being pressed, in the wrinkles and folds they make, we don't just see low-level properties such as edges, contours, or colors, but also seemingly higher level ones, such as mass, elasticity and stiffness. What neural and computational mechanisms underlie these percepts? Previously, using psychophysics and modeling, we found that human soft object perception is best explained by a model that incorporates "intuitive physics", as opposed to performance-matched alternatives that only consider pattern recognition (implemented as a CNN). Here, we hypothesize that, in the human brain, these intuitive physics-based representations (i) are computed spontaneously during visual processing, i.e., in the absence of physics-related tasks, (ii) occur in regions common with that of physical reasoning about rigid objects, and (iii) generalize across qualitatively different scene configurations. To address this, we used fMRI to scan participants (N=20) as they passively viewed animations of cloths at two stiffness levels (stiff and soft) undergoing naturalistic deformations in four different scene configurations (e.g., blowing in the wind, draping on an uneven surface). We identified each participant's regions of interest ("physics-ROI") using a previously validated localizer of physical inferences based on rigid objects. Univariate analysis showed that both physics-ROI and V1 were modulated by the soft vs. stiff cloths, but physics-ROI was modulated by this contrast to a significantly greater degree than V1. Moreover, multivariate analysis revealed successful cross-scene decoding of stiffness levels in physics-ROI (ACC=0.61). Notably, fine-grained rankings of cross-decoding accuracy across different scene configurations were well-captured by representations inferred in our physics-based computational model (Kendall's τ=0.64) but not those in the performance-matched CNN (τ=-0.02). These results help reveal the implementation of physics-based representations of soft objects in the brain.

Talk 5, 9:15 am, 41.25

ASYMMETRY OF NEURAL CIRCUITS FOR WORD AND FACE RECOGNITION IN READERS OF ROMAN AND ARABIC SCRIPT

Zahra Hussain[1,2] (zahra.hussain@plymouth.ac.uk), Sarah Abdou[2], Aaliya Sageer[1], Julien Besle[1,2]; [1]University of Plymouth, [2]American University of Beirut

Words and faces preferentially engage regions in opposite hemispheres of the brain, with corresponding differences in recognition ability between the visual fields. These face and word asymmetries have not been compared simultaneously in readers of different scripts. We compared cortical and behavioural laterality between monolingual readers of Roman or Arabic script, and bilingual readers of both scripts, to evaluate whether reading experience and script properties (e.g., reading direction), alter the representation of words and faces. Cortical activation was measured using 3T fMRI in 21 subjects (6-8 per language group; 3 groups) who viewed faces, English and Arabic words, and control stimuli, whilst performing a one-back task. Cortical regions of interest (ROI) were identified for faces (contrast: faces > phase-scrambled faces and faces > houses; ROIs: fusiform face area, occipital face area, superior temporal sulcus), and for English and Arabic words (contrast: words > phase-scrambled words; ROI: visual word form area). BOLD activation and number of voxels were measured in each ROI. Behaviour was measured outside the scanner in four tasks involving stimuli presented in the left or right visual field (lexical decision, same-different word discrimination, 10AFC face identification, chimeric face identity), with eye movements monitored throughout. We found effects of group on cortical and behavioural laterality for words and faces both. All groups showed cortical left-hemispheric dominance for words in the habitually-read script, but this effect was strongest for English readers, and only English readers showed superior word recognition in the right than left visual field. Likewise, cortical right-hemispheric dominance for faces was strongest in English readers, intermediate for bilinguals, and was virtually absent in Arabic readers. These effects were paralleled in chimeric face judgements. Thus, reading experience or the properties of the habitually-read script alter the symmetry of neural representations for words and faces.

This work was funded by a University Research Board grant from the American University of Beirut

Talk 6, 9:30 am, 41.26

OPTOGENETIC STIMULATION IN MACAQUE V4 CORTEX INDUCES ROBUST DETECTABLE VISUAL EVENTS

Rosa Lafer-Sousa[1] (rosa.lafer-sousa@nih.gov), Lilly Kelemen[1], Reza Azadi[1], Elia Shahbazi[1], Arash Afraz[1]; [1]NIMH

Understanding the nature of the perceptual events evoked by neural perturbations is essential for bridging the causal gap between neuronal activity and vision as a behavior. Here we assess behavioral detectability of optogenetic stimulation in monkey V4 cortex. Two macaque monkeys were chronically implanted with LED arrays over a region of V4 cortex transduced with the depolarizing opsin C1V1. The animals were trained to detect stimulation while fixating at different images. In each trial an image was displayed on the screen for 1s. In half of trials, randomly selected, a 200ms optical impulse was delivered halfway through image presentation, and the animal was rewarded for correctly identifying whether the trial contained cortical stimulation. Both animals learned to perform the task significantly above chance within 11 and 7 sessions respectively (Chi-sq, p-values < 0.01) and improved their performance to 90% and 83% after 27 and 13 more training days (Chi-sq, p-values < 0.001). After the training phase, 20 novel images were used to test whether stimulation detection depends on the choice of onscreen image. The choice of image had a significant effect on stimulation detection (permutation test, p-values < 0.001). Further, the effect varied as a function of cortical location. Taken together, the results suggest the effect of stimulation is visual in nature and stimulation of different subregions evoke different perceptual events. Next we asked whether the stimulation-evoked events are additive in nature, by varying the visibility of the onscreen images. In contrast to inferotemporal cortex, reducing the visibility of the onscreen images did not systematically reduce stimulation detection. These results suggest the events evoked by stimulation in V4 are additive. The findings show for the first time that optogenetic stimulation of V4 cortex induces robust detectable visual events, opening the door to systematic causal studies of V4 with optogenetic methods.

NIMH Intramural Research Training Award (IRTA) Fellowship Program; NIMH Grant ZIAMH002958

Talk Session: Monday, May 20, 2024, 10:45 am – 12:15 pm, Talk Room 1

3D Perception

Moderator: Laurie Wilcox, York University

Talk 1, 10:45 am, 42.11

COMPETITION BETWEEN PRIORS FOR CONVEXITY AND RIGIDITY IN STRUCTURE-FROM-MOTION

Ryne Choi[1,2] (ryne.choi@rutgers.edu), Jacob Feldman[1,2], Manish Singh[1,2]; [1]Rutgers University - New Brunswick, [2]Rutgers University, Center for Cognitive Science

We investigated the competition between priors for convexity and rigidity in Structure-From-Motion. We found that a preference for convexity can override the ubiquitous rigidity assumption: a rigidly rotating plane with a convex hill and a concave valley is perceived as a surface with two hills, moving non-rigidly. Our SFM stimuli consisted of two vertically elongated parts (both half-ellipsoids or bivariate Gaussians; each a convex “hill” or concave “valley”), centered in the left and right halves of a square plane. Competition between priors occurs with “one hill–one valley” stimuli: when convexity wins, the surface is seen as moving non-rigidly, with two hills. When rotated about its central vertical axis (Experiments 1,2), the non-rigid percept is of “folding” along that axis; when rotated about its horizontal axis (Experiments 3,4) it is of “twisting.” We manipulated the strength of convexity through the shape of the parts: half-ellipsoids (more convex) vs. bivariate-Gaussians (less convex); and non-rigidity through the angular range of rotation (larger ranges lead to greater non-rigidity). Observers reported whether the surface was “rigid” or “non-rigid,” and the parts “hills” or “valleys”. Under orthographic projection (Experiments 1,3), observers perceived non-rigid motion on a large percentage of trials for both vertical (93%) and horizontal (53%) axis rotation. The proportion of “non-rigid” responses was higher for the ellipsoids than Gaussians, and for smaller ranges of rotation (consistent with the expected effects of convexity and rigidity respectively). While perspective (Experiments 2,4) lowered the overall percentage of non-rigid responses (31% and 10% for vertical and horizontal axis rotation respectively), they remained significantly above zero (the prediction of the rigidity assumption), while maintaining the trends observed under orthographic projection. The results demonstrate even when a rigid interpretation is available, and even when perspective supports that interpretation, a convexity bias can overcome both, leading to a non-rigid percept.

Talk 2, 11:00 am, 42.12

CAN “PRIOR KNOWLEDGE” OF ISOTROPY BE VARIED TRIAL BY TRIAL? YES, IN SLANT FROM TEXTURE

Zihan SHEN[1] (diasss@163.com), Zhongting Chen[1]; [1]School of Psychology and Cognitive Science, East China Normal University

When perceiving slant from texture, observers tend to presume that elements of texture are initially isotropic. Multiple studies confirm this presumption by showing that variation of aspect ratio of texture (i.e., deviation from isotropy) alters slant perception from texture. However, there is yet little literature on how this “prior knowledge” is made. The current study addressed this issue by attempting to convey “knowledge” of isotropy/anisotropy to observers. We introduced a set of two-folded surfaces, of which the upper parts were slanted and the lower parts frontal-parallel. Both parts were planar but connected smoothly with a curved surface. The whole surface was textured with Voronoi textures and the aspect ratio of texture were independently manipulated for the upper and the lower surfaces, in a range between 0.8 (compressed) to 1.0 (isotropic). The aspect ratio of texture on the connection part gradually changed to avoid any abrupt change of texture. In both experiments (N = 30 for Experiment 1; N = 47 for Experiment 2), observers viewed the two-folded surfaces and estimated 3D slants of the upper surfaces by aligning their hand with the orientations of the upper surfaces while they were asked to ignore the lower surfaces. The estimates from both experiments showed that decrease in aspect ratio of the upper surfaces led to the observers’ overestimation of surface slant, as previous studies have shown. Most interestingly, decrease in aspect ratio of the lower surfaces, which had no direct relations to the task, made the observers significantly underestimate slant of the upper surface, even when aspect ratios of the lower surfaces varied trial by trial. These findings indicate that isotropy is not fixed knowledge but more likely to be contextual information. Observers can flexibly choose whether to adhere to the presumption of isotropy, depending on their environmental understanding."

the Shanghai Municipal Natural Science Foundation [grant No. 23ZR1417900]

Talk 3, 11:15 am, 42.13

SHADING AND CONTOUR COOPERATE TO MODULATE THE PERCEIVED 3D SHAPE OF DISPARITY-DEFINED SURFACES

Celine Aubuchon[1] (celine.d.aubuchon@gmail.com), Roland W. Fleming[2,3], Fulvio Domini[1]; [1]Cognitive, Linguistic, and Psychological Sciences, Brown University, [2]Department of Experimental Psychology, Justus Liebig University Giessen, [3]Centre for Mind, Brain and Behaviour, University of Marburg and Justus Liebig University Giessen

Multiple cues are used to estimate 3D shape, which can mutually constrain each other. Notably, contours have a prominent effect on perceived shape from shading1, where covariations (coordinated changes) in luminance and contour create a strong impression of 3D shape. Conversely, a bumpy shading gradient can be flattened by a smooth contour in special cases where this covariation is disrupted. However, we found no such effect of contour on disparity, which is already highly specified at close distances. This is not surprising given that contour does not generally constrain disparity information. Still, disparity is expected to constrain the interpretation of shading. Here, we tested whether the effect of covariation between contour and luminance would be strong enough to override the shape specified by disparity. We started by cropping a periodic luminance pattern with either a covarying corrugated contour or a smooth contour. This

effectively modulated whether the image was perceived to be a shaded 3D corrugated surface or a smooth surface with light and dark blurry stripes. We then combined these images with disparity fields that were either corrugated or smooth. Observers were asked to report the shape of the surface in a 2AFC task. Remarkably, we found that when the surface specified by disparity was smooth, the presence of covarying luminance and corrugated contour information significantly increased the rate that observers responded 'corrugated'. This was the case despite the null effect of contour on the perceived shape from disparity when the luminance pattern was not present. Together, these findings suggest that the covariation between contour and luminance supports a 3D interpretation that is strong enough to override the otherwise powerful disparity cue. 1Todorović, D. (2014). How shape from contours affects shape from shading. Vision Research, 103, 1-10

his material is based upon work supported by the National Science Foundation under Grant No. 2120610

TALK 4, 11:30 AM, 42.14

FAILURES IN DEPTH MAGNITUDE ESTIMATION IN 3D DISPLAYS

Arleen Aksay[1] (aaksay@yorku.ca), Deborah Giaschi[2,3], Laurie M. Wilcox[1]; [1]York University, [2]The University of British Columbia, [3]British Columbia Children's Hospital

Using naturalistic 3D 'thicket' and 'branch' stimuli we have shown that experienced observers generate accurate depth magnitude estimates for fused targets viewed in virtual reality (VR). However, we have observed that individuals with little to no experience with 3D displays exhibit striking errors in estimating depth from disparity. We conducted a set of experiments with inexperienced viewers to quantify and better understand their poor performance. Our first study was a replication of our experiment with experienced observers where novice participants viewed low and high complexity stimuli using a VR headset. In the 'branch' condition, two branches were presented, one on either side of a central reference branch. The more complex 'thickets' were composed of two clusters of overlapping branches centred on a reference branch. We varied the separation between the branches, and within the thickets, from 1.5 to 12 cm by displacing their components equally in front of and behind the fixation point. Sixteen inexperienced observers indicated the overall depth of the structures with a virtual ruler. In another study we evaluated the role of depth averaging by displacing the branches and thickets in a single direction relative to fixation. Unlike our previous results with experienced observers, in both experiments we found that the association between perceived depth and increasing disparity was weak. This, and the fact that all participants could perceive depth from all disparities in this range using a depth-order (near/far) discrimination task argues against an explanation based on depth averaging. We conclude that cue conflicts, particularly related to the contribution of vergence to estimation of viewing distance, interfere with inexperienced participants' ability to compute depth from disparity. Our working hypothesis is that with extended experience observers learn to disregard such conflicts; how they do this is the focus of ongoing research.

Natural Sciences and Engineering Research Council (NSERC) Grant # RGPIN-2019-06694; CF-REF program Vision Sciences to Applications (VISTA)

TALK 5, 11:45 AM, 42.15

CANONICAL PERSPECTIVES OF RENDERED 3D OBJECTS ARE RELATED TO AFFORDANCE

Athanasios Bourganos[1] (athan.bourganos@mail.utoronto.ca), Dirk B. Walther[1]; [1]University of Toronto

Humans prefer to view objects from some but not other perspectives. Palmer, Rosch, and Chase (1981) were first to use the term "canonical perspectives" to describe these preferred viewing angles. More recently, this phenomenon has been studied as it relates to perspective invariance, object identification (human & algorithmic), and navigation. Contemporary studies rely on some of the foundational observations of early canonical perspective research. However, those original results are contradictory in several respects. Past literature includes contradictory findings on between-observer agreement on preferred perspectives, reaction time effects, and support for mental rotation theories of 3D object perception. To address those contradictions and improve our understanding of canonical perspectives, we constructed a digital dataset of three-dimensional objects from three categories: graspable familiar objects, non-graspable familiar objects, and graspable unfamiliar objects. We rendered the objects as viewed from 26 different orientations, covering the full range of viewing angles. We collected canonical perspective ratings via a pairwise comparison task, where participants indicated their preference between two displayed views in a two-alternative, forced-choice task. We presented 325 pairs of views of each object. Ratings were highly consistent between observers. Some viewing angles of graspable objects (e.g., coffee mug) were rated differently between left- and right-handed participants, based on experienced handle placement. This result indicates a significant connection between canonical perspective and affordance. We see a similar, although slightly weaker effect when comparing canonical viewing angles between participants of different body height. Taller participants are biased toward views from the top, smaller participants to views from the front. In summary, our results suggest that viewing angle influences people's aesthetic preference for viewing objects, and that the preferred canonical perspective is frequently related to the individually specific affordance of a particular view.

This work was supported by NSERC Discovery Grant (RGPIN-2020-04097) and SSHRC Insight Grant (435-2023-0015) to DBW.

TALK 6, 12:00 PM, 42.16

LATE DEVELOPMENT OF SENSITIVITY TO RELATIVE DISPARITY IN HUMAN VISUAL CORTEX IN THE FACE OF PRECOCIAL DEVELOPMENT OF SENSITIVITY TO ABSOLUTE DISPARITY

Anthony Norcia[1] (amnorcia@stanford.edu), Milena Kaestner, Yulan Chen, Caroline Clement; [1]Stanford University

Introduction: Immaturities exist at multiple levels of the developing human visual pathway, starting with immaturities in photon efficiency and spatial sampling in the retina and on through immaturities in early and later stages of cortical processing. Here we use Steady-State

Visual Evoked Potentials (SSVEPs) and controlled visual stimuli to determine the degree to which sensitivity to horizontal retinal disparity is limited by the visibility of the monocular half-images, the ability to encode absolute disparity or the ability to encode relative disparity. Methods: Responses were recorded from male and female participants at average ages of 5 months, 5 and 25 years. SSVEPs were recorded in response to contrast and blur modulation of dynamic random dot patterns to measure sensitivity to the spatio-temporal content of the monocular half-images. Disparity sensitivity was measured using planar stereograms that modulated absolute disparity and in stereograms portraying disparity gratings that additionally had relative disparity in them. Results: Disparity thresholds derived from SSVEP amplitude vs disparity response functions for planar stimuli modulating absolute disparity changed little over development, but those for grating stimuli modulating relative disparity changed by a factor of ~10. Equating subjective contrasts between infants, children and adults did not equate disparity sensitivity. Disparity sensitivity at age 5 was adult-like, but disparity tuning at supra-threshold levels was not. Conclusion: The protracted developmental sequence for relative disparity coding shown in our measurements is not simply inherited from immaturities in encoding absolute disparity, but rather reflects immaturities in the computations needed to represent relative disparity that likely involve extra-striate cortical areas where relative disparity is first extracted.

Acknowledgements: This research was supported by grant EY018875 from the National Eye Institute, National Institutes of Health. The authors would like to thank Vladimir Vildavski and Alexandra Yakovleva for the development of instrumentation used in the experiments.

Talk Session: Monday, May 20, 2024, 10:45 am – 12:15 pm, Talk Room 2

Development

Moderator: Leyla Isik, Johns Hopkins University

Talk 1, 10:45 am, 42.21

INCREASED USE OF PRIORS IN THE OBLIQUE EFFECT IN CHILDREN

Sarit Szpiro[1,2], Eman Mhajne[1]; [1]University of Haifa, [2]The Edmond J. Safra Brain Research Center, University of Haifa

Background: According to Bayesian inference models of vision, perception is affected by incoming input (i.e., likelihood) and by past experiences (i.e., priors). Although perceptual inference has been widely studied in adults, much less is known about how Bayesian inference is modified during development. Here, we study how priors impact perception during development by examining the oblique effect in children. In the oblique effect, adults exhibit lower motion direction thresholds and larger estimation biases for cardinals versus oblique motion directions. Methods: We compared the perception of motion directions using RDK motion stimuli in adults (ages 20-38) and children (ages 7-9). First, we determined individual coherence discrimination thresholds for near oblique and for near horizontal directions using a staircase procedure. Then, participants were presented their at-threshold motion directions (near horizontal/oblique) and moved a mouse to estimate directions. Results: Results indicated that children exhibited higher overall thresholds compared to adults. Both groups showed significantly lower thresholds for horizontal motion compared to oblique motion, with this difference significantly more pronounced in children. Additionally, both groups displayed significantly larger estimation biases for near horizontal directions, which again was significantly more prominent in children. Conclusions: Even though horizontal and oblique stimuli were presented at their corresponding thresholds, there were larger estimation biases for horizontal versus oblique in both groups. This suggests that these biases persist even when equating for difficulty, an important factor to consider in future computational models of bias and discriminability. Importantly, our results reveal that children exhibit larger oblique effects than adults. The overall worse thresholds in children (i.e, likelihood) may drive children to rely more on perceptual priors than adults, impacting their perception of motion. Our study contributes valuable insights into the developmental aspects of perceptual inference and the role of priors in perception during development.

ISF research grant 1198/22 to Sarit Szpiro

Talk 2, 11:00 am, 42.22

THE EMERGENCE OF VWFA LATERALITY: EXAMINING THE ROLE OF WHITE MATTER CONNECTIVITY IN EARLY CHILDHOOD

Jin Li[1,2] (li.9361@buckeyemail.osu.edu), Leah DiRubio[1,2], Zeynep Saygin[1,2]; [1]Department of Psychology, The Ohio State University, [2]Center for Cognitive and Behavioral Brain Imaging (CCBBI), The Ohio State University

The visual word form area is a brain region that selectively responds to visual words as people become literate. Interestingly, while face selectivity is typically dominant on the right hemisphere, word selectivity is dominant on the left. What factors contribute to the development of left laterality for words? Here we tested three potential sources of this word laterality: 1) the development of face laterality, 2) cross-hemispheric structural connectivity, and 3) connectivity with ipsilateral language regions. We scanned children (3-9 years, prereaders and readers) and adults on an fMRI task to extract functional activation to words and faces and diffusion-weighted imaging (DWI) in the same participants to examine white matter connectivity. In children, even though word selectivity became increasingly left-lateralized and face selectivity more right-lateralized over development, they were not directly related: while face laterality increased with age in both readers and prereaders, word laterality increased with age only in readers. Word laterality additionally showed a significant correlation with cross-hemispheric connectivity between left VWFA and its right homotope, especially in readers. Furthermore, we found that in readers, the connectivity of VWFA with ipsilateral frontal language regions was positively related to its laterality. Interestingly, in adults, neither face laterality nor connectivity was correlated with individual differences in word laterality. These results demonstrate how anatomical and developmental factors contribute to changes in laterality of high-level visual cortex, and highlight the role of both cross-hemispheric and ipsilateral white matter connectivity in developing word laterality.

Alfred P. Sloan Research Fellowship (to Z.M.S); CCBBI Gibson Research Award (to J.L.)

Talk 3, 11:15 am, 42.23

EARLY NEURAL DEVELOPMENT OF SOCIAL PERCEPTION: EVIDENCE FROM VOXEL-WISE ENCODING IN YOUNG CHILDREN AND ADULTS

Elizabeth Jiwon Im[1] (imelizabeth@jhu.edu), Angira Shirahatti[1], Leyla Isik[1]; [1]Johns Hopkins University

From a young age, children show advanced social perceptual and reasoning abilities. However, the neural development of these abilities is still poorly understood. To address this gap, we used publicly available fMRI data collected while children and adults watched an engaging and socially rich movie to investigate how the cortical basis of social processing changes throughout development. We annotated segments of the movie with visual and social features, including motion energy, presence of faces, presence of a social interaction, theory of mind (ToM) events, valence and arousal. Using a voxel-wise encoding model trained using these features, we find that visual (motion energy) and social (faces, social interaction, ToM, valence, and arousal) features can both predict brain activity in children as young as three years old across the cortex, with particularly high predictivity in motion selective MT and the superior temporal sulcus (STS). Furthermore, individual social feature models showed that while representations for some social features, like ToM, develop throughout childhood, social interaction representations in the STS appear adult-like in even the youngest children. The current study, for the first time, links neural activity in children to specific social features during naturalistic movie viewing and suggests social interaction perception is supported by early developing neural responses in the STS.

This work was supported with funds from the Clare Boothe Luce Program for Women in STEM and NIMH R01MH132826 awarded to LI. We would like to thank Hillary Richardson and colleagues for data used in this study.

Talk 4, 11:30 am, 42.24

HIGH-RESOLUTION DIFFUSION MRI OF THE CORTICO-CORTICAL CONNECTIONS BETWEEN LOWER VISUAL AREAS REVEALS DIVERGENCE OF CONNECTIONS AND ENHANCED CONNECTIVITY OF THE CENTRAL VISUAL FIELD REPRESENTATION

Ziqi Hao[1] (ziqi.hao@mail.mcgill.ca), Alex Valcourt Caron[2], Janine Mendola[1], Maxime Descoteaux[2], Amir Shmuel[1]; [1]McGill University, Montreal, QC, Canada, [2]Universite de Sherbrooke, Sherbrooke, QC, Canada

Studies of the dense structural connections between human lower visual areas with diffusion MRI (dMRI) are sparse. To investigate this fine connectivity, a dMRI pulse sequence is required to operate at both high-resolution and high signal-to-noise ratio (SNR). gSldier-SMS (Setsompop et al., 2018) makes acquiring diffusion MRI data with high SNR and high resolution possible. Using gSlider-SMS, we aim to image the white matter connections between the densely connected lower visual areas V1, V2, and V3. Data were obtained from 10 healthy subjects. Using gSlider-SMS, we acquired dMRI with isotropic 1 mm voxels. We used Benson's atlas (Benson & Winawer, 2018) to generate the retinotopic polar angle and eccentricity maps of human visual areas V1, V2, and V3. Our diffusion data demonstrated a retinotopically organized connectivity pattern between V1-V2, V2-V3, and V1-V3. Moreover, our results also revealed differential density in connectivity patterns between the central and peripheral visual fields. We found that for each pair of connections between V1, V2, and V3, the central visual field regions were more densely connected than the regions in the periphery. This underscores the notion that the central visual field, responsible for detailed and sharp vision, might have evolved to have stronger connectivity to facilitate visual processing at high acuity. In addition, we evaluated the divergence of connections emerging from an eccentricity range in one visual area and connecting to a second visual area. A Gaussian curve showed a good fit to the divergence distribution. Our study provides novel insight into the visual cortex's intricate connectivity patterns, underlining the importance of central visual field representation, retinotopic organization, and hierarchical processing. Our findings also pave the way for investigating the structural-functional relationship of connections between lower visual areas, potentially leading to a deeper understanding of vision-related disorders.

This work was supported by the Assistant Secretary of Defense for Health Affairs endorsed by the Department of Defense, through the Vision Research Program (VRP) under Award No W81XWH1910853. Opinions, interpretations, conclusions and recommendations of the author are not endorsed by the Department.

Talk 5, 11:45 am, 42.25

REDUCED DISSOCIATION BETWEEN PERCEPTION AND ACTION IN INDIVIDUALS WITH AUTISM

Zoha Ahmad[1] (zohahmad@my.yorku.ca), Orly Aziza[1], Tzvi Ganel[2], Bat-Sheva Hadad[3], Erez Freud[1]; [1]York University, [2]Ben Gurion University of the Negev, [3]University of Haifa

Changes in perceptual behaviors are commonly reported in individuals with Autism. Nevertheless, the mechanisms that mediate these alterations are yet to be determined. One hypothesis is that these changes are the result of reduced cortical specialization. In this study, we scrutinize this hypothesis utilizing the framework of the functional dissociation between perception and action. Past studies have demonstrated that unlike their robust effect on perception, altering the temporal or spatial size context of stimulus presentation have little effect on grasping. Therefore, we manipulated the temporal or spatial context of stimulus presentation, while both autistic and neurotypical participants completed grasping and estimation tasks. In Experiment 1 (spatial context), we used two objects that differ in size, and placed on an illusory Ponzo background such that they were perceived as "close" or "far" from the observer. The neurotypical individuals showed a perception-action dissociation with a robust effect of the illusion only during the estimation task, and no effect of the illusion during grasping. In contrast, autistic individuals were affected by the illusion under both conditions. In Experiment 2 (temporal context), we presented a standard stimulus (40mm) in two temporally distinct blocks. In one block ("wide"), the stimuli used as flanks for the standard stimuli encompassed a larger range (20-60mm). In the other block ("narrow"), the flanks encompassed a smaller range (35-45mm). The neurotypical individuals showed a perception-action dissociation with a robust effect of the experimental temporal context only during the perceptual task. In contrast, autistic individuals showed a consistent effect of the experimental temporal context under the perceptual and the visuomotor task, with larger JNDs for the standard stimulus presented during the "wide" block. Taken together, these results provide converging evidence for reduced functional dissociation between perception and action in autism, such that perceptual representations intrude and modulate visuomotor behaviors.

Talk 6, 12:00 pm, 42.26

VISUAL EXPERIENCE IS REQUIRED TO DEVELOP A FLEXIBLE SENSE OF POSITION.

Carolina Tammurello[1,2] (carolina.tammurello@iit.it), Lara Coelho[1], Maria Bianca Amadeo[1], Claudio Campus[1], Walter Setti[1], Monica Gori[1]; [1]Italian Institute of Technology, [2]University of Genoa

In the somatic Rubber Hand Illusion, a blindfolded participant is guided in brushing a dummy hand, while the experimenter synchronously brushes the participant's hand. This can induce proprioceptive drift (PD), i.e., a shift in perceived position of one's own hand. One population that is immune to PD are blind adults, a fact often attributed to their hypothesized superior proprioceptive precision. To explore the role of visual experience on PD and if a lack of vision causes different developmental trajectories, we recruited Severely Visually Impaired (SVI) children (6-11 y.o.; n=14) and compared their PD to that of sighted-age-matched controls (n=51). As hypothesized, SVI children failed to show PD across the age span, unlike sighted children, whose PD increased as a function of age. We then tested whether I) proprioceptive precision and II) distance between the initially perceived position of their own hand and the placement of the dummy hand predicted PD in sighted children. I) did not predict PD in sighted children, nor were there any differences between groups in this parameter. Instead, II) was significant (i.e., the farther the two, the larger the shift) and not modulated by age. We speculate that sighted children show PD as a reaction to the conflict between inter-hand distance (proprioceptive information) and coherence between the stimulation performed (which is actually on the dummy hand) and the one received (tactile and kinesthetic feedbacks). Thus, an intermediate position, where the stimulated hand would be if participants were stroking their own hand, becomes an attractor. The perceptual system of SVI children, instead, fails to react to this conflict and continues to compute the position of the hand irrespectively of any additional contextual information. Visual experience is therefore necessary to remap the body in external space and achieve a reduction in intersensory conflict.

The research was supported by the MYSpace project (PI Monica Gori), which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 948349)

Talk Session: Tuesday, May 21, 2024, 8:15 – 9:45 am, Talk Room 1

Object Recognition: Categories and features

Moderator: Juan Chen, South China Normal University

Talk 1, 8:15 am, 51.11

THALAMOCORTICAL PATHWAYS UNDERLYING UNCONSCIOUS ACTION-RELATED VISUAL INFORMATION- EVIDENCE FROM THE NEURAL REPRESENTATIONS OF BINOCULARLY SUPPRESSED TOOL IMAGES

Zhiqing Deng[1] (zhiqingdeng@m.scnu.edu.cn), Fuying Zhu[1], Jie Gao[1], Zhiqiang Chen[2,3,4], Peng Zhang[2,3,5], Juan Chen[1,6]; [1]Center for the Study of Applied Psychology, Guangdong Key Laboratory of Mental Health and Cognitive Science, and the School of Psychology, South China Normal University, Guangzhou, Guangdong Province, 510631, China, [2]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China, [3]School of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China, [4]Sino-Danish College, University of Chinese Academy of Sciences, Beijing 100049, China, [5]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 15 230026, China, [6]Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education

In exploring the integral role of human subcortical structures, including the thalamus, brainstem, and basal ganglia, our study delves into their contribution to unconscious high-level perceptual processing, notably focusing on visual categorization. While extensive behavioral and fMRI studies have suggested the significance of subcortical pathways for residual vision within blindsight, the specific involvement of the thalamus, superior colliculus, and basal ganglia in unconscious high-level perceptual processing remains elusive in healthy humans. Here, we employed functional magnetic resonance imaging (fMRI) to investigate the representation of toolness (tools vs. non-tools) and shape (elongated vs. stubby) within subcortical structures while images were made invisible using continuous flash suppression (CFS). Both univariate analysis and multivoxel pattern analysis (MVPA) based on blood-oxygenation level-dependent (BOLD) signals revealed a significant toolness representation in the left thalamus, with the left ventral anterior thalamus (VA, motor-related) as the most important thalamic subregion contributing to the toolness representation under CFS. In the basal ganglia, the left striatum (STR) exhibited robust toolness representation in both univariate analysis and MVPA. Among cortical regions, only area 9a and anterior 10p (part of the dorsolateral prefrontal cortex, DLPFC) demonstrated significant toolness representation. Functional connectivity results indicated that elongated tools increased the connectivity between the bilateral VA in the thalamus and left 9a (DLPFC) and between the right VA and left STR in the basal ganglia, compared to elongated non-tools in CFS. Notably, dynamic causal modeling (DCM) results unveiled a thalamocortical pathway from the left VA in the thalamus to the left 9a

(DLPFC) contributing to toolness representation when tool images are rendered invisible by CFS. These findings shed light on the role of the subcortical structures, particularly the thalamus and basal ganglia, and highlight a thalamocortical pathway in healthy humans engaged in unconscious high-level perceptual processing, especially unconscious visual categorization.

This research was supported by the Supported by two National Natural Science Foundation of China grants (31970981 and 31800908) and by the National Science and Technology Innovation 2030 Major Program (STI2030-Major Projects 2022ZD0204802 to JC).

Talk 2, 8:30 am, 51.12

A GENERAL ABILITY FOR SIMPLE AND COMPLEX ENSEMBLE JUDGMENTS

Isabel Gauthier[1] (isabel.gauthier@vanderbilt.edu), Ting-Yun Chang[1], Oakyoon Cha[2]; [1]Vanderbilt University, [2]Sungshin Women's University

People can report summary statistics for various features about a group of objects. One theory is that different abilities support ensemble judgments about low-level features like color vs. high-level features like identity. Existing research mostly evaluates such claims based on evidence of correlations within and between feature domains. However, correlations between two identical tasks that only differ in the type of feature for ensemble judgments can be inflated by method variance. Another concern is that conclusions about high-level features are mostly based on faces. We used latent variable methods on data from 237 participants to investigate the abilities supporting low-level and high-level feature ensemble judgments. Ensemble judgment was measured with six distinct tests, each requiring judgments for a distinct low-level (orientation, lightness, aspect ratio) or high-level feature (bird species, Ziggerin identity, Transformer identity), using different task requirements in each task (mean estimation, mean matching, diversity comparison). We also controlled for other general visual abilities when examining how low-level and high-level ensemble abilities relate to each other. Confirmatory factor analyses showed a perfect correlation between the two factors, suggesting a single ability. A nested model comparison confirmed that using one ensemble perception (EP) factor rather than two did not impair model fit. There was a strong unique relationship (.9) between these two factors, beyond the influence of object recognition and perceptual speed. Additional results from 117 of the same participants also ruled out an important role for working memory in explaining the EP factor. Our results demonstrate that the ability common to a variety of ensemble judgments with low-level features is the same as that common to a variety of ensemble judgments with high-level features.

This work was supported by the David K. Wilson Chair Research Fund (Vanderbilt University) and the Taiwanese Overseas Pioneers Grants (TOP Grants) for PhD Candidates from Ministry of Science and Technology, Taiwan.

Talk 3, 8:45 am, 51.13

DEVELOPMENTAL CHANGES IN THE PRECISION OF VISUAL CONCEPT KNOWLEDGE

Bria Long[1] (bria@stanford.edu), Wanjing Anya Ma[1], Rebecca Silverman[1], Jason Yeatman[1], Michael C. Frank[1]; [1]Stanford University

How precise is children's visual concept knowledge, and how does this change across development? We created a gamified picture-matching task where children heard a word (e.g., "swordfish") and had to choose the picture "that goes with the word." Critically, we chose distractor items with high, medium, and low similarity to each target word, allowing us to examine the granularity of visual representations. We derived similarity via cosine embedding similarity of the target and distractor words in CLIP, a language-vision pre-training model (Radford et al., 2021). Photographs were taken from the THINGS+ dataset and combined with age-of-acquisition (AoA) ratings, yielding 108 items with unique targets and three distractors with estimated AoA ratings within 3 years of each other; we created 2AFC trials with high similarity distractors, 3AFC with high and medium similarity distractors, and 4AFC trials that included a low similarity distractor. Data were then collected from children in preschools (N=66 3-5 year-olds), 6 elementary schools, and 9 charter schools across multiple states (N=1369, 6-11 year-olds) and adults online (N=205). We modeled changes in the proportion of children who chose a given image for a certain word over development using linear mixed-effect models. We found gradual developmental changes in children's ability to identify the correct category. Error analysis from 3- and 4-AFC trials revealed that children were more likely to choose higher-similarity distractors as they grew older; children's error patterns were increasingly correlated with CLIP target-distractor similarity. Overall, these analyses suggest a transition from coarse to finer-grained visual representations over early and middle childhood. Children's visual concept knowledge gradually becomes more refined as children learn what distinguishes similar visual concepts from one another. Broadly, these findings demonstrate the utility of combining gamified experiments and similarity estimates from computational models to probe the content of children's evolving visual representations.

NIH K99 HD108386

Talk 4, 9:00 am, 51.14

TWO-DIMENSIONAL ATTRIBUTES IN TACTILE DEPICTIONS THAT CONVEY THREE-DIMENSIONALITY OF OBJECTS.

Anchal Sharma[1] (anchalsharma.mail@gmail.com), Srinivasan Venkataraman[1], PV Madhusudhan Rao[1]; [1]Indian Institute of Technology Delhi

Three-dimensional objects are readily identifiable in standardized representations like perspective/orthographic projections in the visual domain. While readily understandable visually, these representations are not easily interpretable by blind individuals, especially those congenitally blind. This leaves a major gap in their education and points to a need for investigating a representation technique to convey the volumetric attributes of an object. In this study, 5 geometric objects and 20 tactile representations were provided to 20 participants comprising born-blind, late-blind, and blindfolded-sighted individuals (aged 18-44). Each object was presented on tactile sheets in four different representation styles in randomized order: Generator-director, surface development, isometric view, and dual-view. Participants were first familiarized with the original objects via haptic exploration. Subsequently, they were presented with the aforementioned tactile representations, and for each stimulus, were required to indicate whether it related to any of the original objects. If identified, detailed descriptions were recorded for what features of a particular representation style made it relatable to the identified 3D object. Analysis of the data revealed that surface development and generator-director representation styles were better associated with the objects. Participants' open-ended responses offer deeper insights into why certain representations were preferred over others. Factors included a better indication of surface details and more discernible spatial arrangements. Interestingly, some objects were well identified irrespective of the representation style they were depicted in, pointing to certain unique features present in all of the styles. We discuss tentative constituents of what makes a 2D representation align closely with its 3D counterpart, including local salient features and specific spatial configurations. These results offer hints regarding the cues that allow for translation between 3D structures and 2D tactile depictions, pointing to interesting questions regarding features that are informative in the visual versus tactile domains, and have relevance for conveying graphical information to blind students.

Talk 5, 9:15 am, 51.15

ERROR CONSISTENCY BETWEEN HUMANS AND MACHINES AS A FUNCTION OF PRESENTATION DURATION

Thomas Klein[1,2] (t.klein@uni-tuebingen.de), Wieland Brendel[2], Felix Wichmann[1]; [1]Universität Tübingen, [2]Max Planck Institute for Intelligent Systems, Tübingen, Germany

Within the last decade, Artificial Neural Networks (ANNs) have emerged as powerful computer vision systems that match or exceed human performance on some benchmark tasks such as image classification. But whether current ANNs are suitable computational models of the human visual system remains an open question: While ANNs have proven to be capable of predicting neural activations in primate visual cortex, psychophysical experiments show behavioral differences between ANNs and human subjects as quantified by error consistency. Error consistency is typically measured by briefly presenting natural or corrupted images to human subjects and asking them to perform an n-way classification task under time pressure. But for how long should stimuli ideally be presented to guarantee a fair comparison with ANNs? Here we investigate the role of presentation time and find that it strongly affects error consistency. We systematically vary presentation times from 8.3ms to >1000ms, followed by a noise mask, and measure human performance and reaction times on natural, lowpass-filtered and noisy images. Our experiment constitutes a fine-grained analysis of human image classification under both image corruption and time pressure, showing that even drastically time-constrained humans who are exposed to the stimuli for only a single frame, i.e. 8.3ms, can still solve our 8-way classification task with success rates above chance. Importantly, the shift and slope of the psychometric function relating recognition accuracy to presentation time also depends on the type of corruption. In addition we find that error consistency also depends systematically on presentation time. Together our findings raise the question of how to properly set presentation time in human-machine comparisons. Second, the differential benefit of longer presentation times depending on image corruption is consistent with the notion that recurrent processing plays a role in human object recognition, at least for images that are difficult to recognise.

Funded by the German Research Foundation (DFG) under Emmy Noether grant BR 6382/1-1. Supported by EXC grant 2064/1, project 390727645 and SFB 1233, project 276693517. TK would like to thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for support.

Talk 6, 9:30 am, 51.16

THE NEURAL REPRESENTATION OF THE FAKE OBJECTS

Xieyi Liu[1], Pinglei Bao[1,2,3]; [1]Peking-Tsinghua Center for Life Sciences, Peking Univ., Beijing, China, [2]IDG/McGovern Institute for Brain Research, Peking Univ., Beijing, China, [3]School of Psychological and Cognitive Sciences, Peking Univ., Beijing, China

Earlier research suggested that the IT cortex's functional structure can be understood through an object space model with DCNN. However, category-specific regions in the IT cortex, such as areas dedicated to faces and bodies, imply that its organization might also be based on semantic categories. To distinguish between these two hypotheses, we used fMRI to measure human subjects' responses to artificial images, referred to as "fake objects", which were generated with GAN and lacked semantic category information. We projected these generated fake objects onto the PC1-PC2 space, built with the fMRI responses to 500 real objects. We chose 100 fake objects based on their projections onto the space, resulting in a ring-like structure. Subjects were instructed to perform three tasks in separate scans: two image categorization tasks based on the images' projection onto the two orthogonal axes in the object space and a fixation color discrimination task. The study's results show that the IT cortex can be effectively modulated by these fake objects, and the modulation of each voxel can be accurately represented by the object space model as the projection on the preferred axis. This holds true even for voxels located in category-selective regions, such as the FFA and EBA. Furthermore, the preferred axis of each voxel in the IT cortex remained consistent across the three tasks, although the absolute selectivity decreased in the fixation task. Additionally, the modulation of the two different image categorization tasks was more noticeable in the frontal and parietal cortex. Our results demonstrate that the functional organization of the IT cortex can be better explained by the object space model than the semantic model, and the representation of object space is relatively stable across different tasks, whose outputs can be read out by the later stages of the brain.

National Science and Technology Innovation 2030-Major Project 2022ZD0204803, Natural Science Foundation of China Grant NSFC 32271081,32230043 to P. B. Natural Science Foundation of China Grant NSFC32200857, China Postdoctoral Science Foundation Grant 2023M740125, 2022T150021, 2021M700004 to J.Y.

Talk Session: Tuesday, May 21, 2024, 8:15 – 9:45 am, Talk Room 2

Attention: Neural mechanisms

Moderator: Li Zhaoping, Max Planck Institute

Talk 1, 8:15 am, 51.21

LONG-RANGE MODULATORY FEEDBACK CONNECTIONS IN DEEP NEURAL NETWORKS SUPPORT TOP-DOWN CATEGORY-BASED ATTENTION

Talia Konkle[1] (talia_konkle@harvard.edu), George A. Alvarez; [1]Harvard University

Many views of the world are cluttered with multiple kinds of objects present, but at any given moment only a subset of this information may be task-relevant. Top-down attention can direct visual encoding based on internal goals, e.g. when looking for keys, attention mechanisms select and amplify the relevant key-like image statistics, aiding detection and modulating the gain across the visual hierarchy. Motivated by visual cognition and visual neuroscience findings, we designed long-range modulatory feedback pathways to outfit deep neural network models, with learnable channel-to-channel influences between source and destination layers that spatially broadcast feature-based gain signals. We trained a series of Alexnets with varying feedback pathways on 1000-way ImageNet classification to be accurate on both their feed-forward and modulated pass. First, we show that models equipped with these feedback pathways naturally show improved image recognition, adversarial robustness, and emergent brain-alignment, relative to baseline models. Critically, the final layer of these models can serve as a flexible communication interface between visual and cognitive systems, where cognitive-level goals (e.g. “key?”) can be specified as a vectors in the output space, and naturally leverage feedback projections to modulate earlier hierarchical processing stages. We compare and identify the effective ways to ‘cognitively steer’ the model based on prototype representations, which dramatically improve recognition of categories in composite images of multiple categories, succeeding where baseline feed-forward models fail. Further, these models recapitulate neural signatures of category-based attention—e.g. showing modulation of face and scene selective units inside the model when attending to either faces or scenes, when presented with a fixed face-scene composite image. Broadly, these models offer a mechanistic account of top-down category-based attention, demonstrating how long-range modulatory feedback pathways can allow different goal states to make flexible use of fixed visual circuity, supporting dynamic goal-based routing of incoming visual information.

CRCNS 8431439-01 to TK and NSF PAC COMP-COG 1946308 to GAA

Talk 2, 8:30 am, 51.22

7T CBV FMRI REVEALS CORTICAL MICROCIRCUITS OF BOTTOM-UP SALIENCY IN THE HUMAN BRAIN

Peng Zhang[1] (zhangpeng@ibp.ac.cn), Chen Liu[1], Chengwen Liu[1], Li Zhaoping[2]; [1]Institute of Biophysics, Chinese Academy of Sciences, [2]Max Planck Institute for Biological Cybernetics

A visual item in sharp contrast with its neighbors automatically captures attention. Whether bottom-up saliency signals arise initially in the primary visual cortex (V1) or in the parietal cortex is still controversial. To distinguish these two hypotheses, we investigated the cortical microcircuits of bottom-up saliency with cortical layer-dependent CBV fMRI at 7 Tesla. Behavioral experiments measured the contrast detection performance to orientation singletons presented either at low (15 degrees) or high (90 degrees) orientation contrast within uniformly oriented background bars. Contrast sensitivity was higher to singletons with high compared to low orientation contrast. CBV-weighted fMRI results showed that the orientation-saliency signal was strongest in the superficial layers of V1, and peaked in the middle layers of V2/V3 and the intraparietal sulcus (IPS). Contrast sensitivities of the orientation singletons also correlated with CBV signals in the superficial layers of V1. These findings support the hypothesis that bottom-up saliency map is initially created by iso-feature suppression through lateral inhibition in V1 superficial layers, and then projects to the parietal cortex through the feedforward connection.

Talk 3, 8:45 am, 51.23

ROLE OF THETA OSCILLATIONS IN TOP-DOWN CONTROL OF FEATURE-BASED ATTENTION

Sreenivasan Meyyappan[1] (smeyyappan@ucdavis.edu), Mingzhou Ding[2], George Ron Mangun[1,3,4]; [1]Center for Mind and Brain, University of California Davis, [2]J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, [3]Department of Psychology, University of California Davis, [4]Department of Neurology, University of California Davis

Applying top-down control to selectively process and distinguish visual stimuli based on their attributes such as color or motion is known as feature-based attention. Attention-control signals from the specialized regions in the frontal and parietal cortex, also known as dorsal attention network (DAN), are reported to bias the activity in the visual cortex in favor of the attended feature. Prior work has been successful in identifying the role of alpha oscillations (8-12 Hz) in modulation of sensory processing in visual cortex. However, it remains unknown whether and which oscillatory neural activity may support network communication and integration within and between the nodes of the attentional control network. We hypothesize that the nodes in the DAN dynamically interact via theta band (3-7 Hz) activity, and this coordination enables the DAN to send top-down control signals to the visual cortex. We investigated this by recording EEG during a cued feature attention experiment where participants were cued on a trial-by-trial basis to attend either the direction of motion or color of the forthcoming stimuli (moving dots). Using multivariate decoding approaches comparing attend-color versus attend-motion in the post-cue/pre-target period, we observe the pattern of theta and alpha activity to be predictive of the attended feature and importantly, the decoding timecourse in theta band to temporally precede the decoding in the alpha band. Further, estimating the spectral coherence between an ensemble of frontal and parietal scalp electrodes as an index of cortical synchronization between attention control networks from different frequency bands (e.g., theta, alpha, beta, and gamma activity), we observed significant decoding only in the theta band compared to decoding on surrogate (temporally shuffled) data. These results highlight the distinct role of theta oscillations in enabling the top-down control of selective sensory processing at the visual cortical level.

Talk 4, 9:00 am, 51.24

OBJECT-BASED ASSOCIATION FIELDS FOR GROUPING AND ATTENTION

Hossein Adeli[1] (hossein.adelijelodar@gmail.com), Seoyoung Ahn[2], Gregory Zelinsky[2], Nikolaus Kriegeskorte[1]; [1]Columbia University, [2]Stony Brook University

What are the neural mechanisms that group visual features into coherent object percepts? Association fields, mediated by long-range horizontal connections, have been shown to dynamically configure the neural response in early visual areas to form objects from collinear line segments. We propose that such association fields also exist in higher visual areas and contribute to object-based grouping and attention. To test this hypothesis, we modeled the connection strengths in the association fields by measuring the similarity between the local image features from a transformer-based vision model. We then tested the effectiveness of these object-based associations using a well-established grouping task—a two-dot paradigm. In this task, the model needs to determine whether a central and a peripheral dot are on the same or different objects in a natural scene. Our model performs this grouping task by gradually spreading attention, mediated by the association field, from the two dot locations to the neighboring areas. We observed remarkable performance in attention staying within the object while spreading, showing for the first time the plausibility of attention spread through horizontal connections as an object grouping mechanism in scenes. The model reaches a 'same-object' decision when two segments show a sufficient level of agreement in their feature representations, according to a predefined threshold. We observed a significant correlation between the time taken by the model to arrive at its decision and the actual human reaction time in the same task (72 participants for 1020 trials; $r = 0.32$, $p < 0.001$), significantly closing the gap between the baseline models and the subject-subject agreement ($r = 0.42$). In this work, we hypothesize and provide evidence for how the existence of object-based association fields can mediate the spread of attention to group objects in natural scenes providing novel hypotheses to be tested in neuroscience.

Talk 5, 9:15 am, 51.25

THE ROLE OF EXPECTATIONS IN VISUAL SPATIAL CODING ACROSS THE VISUAL HIERARCHY

Ningkai Wang[1,2,3] (n.wang@vu.nl), Ralph Wientjens[1,2,3], Jurjen Heij[2], Gilles de Hollander[4], Jan Theeuwes[1,2], Tomas Knapen[1,2,3]; [1]Vrije Universiteit Amsterdam, Amsterdam, Netherlands, [2]Spinoza Centre for Neuroimaging, Royal Netherlands Academy of Sciences, Amsterdam, Netherlands, [3]Institute Brain and Behavior Amsterdam (iBBA), Amsterdam, Netherlands, [4]Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland

Predictive processing theorizes that the brain predicts events based on prior experiences. Mismatches between the predictions and input lead to prediction errors (PEs). Despite the theory's popularity, our understanding of the role of PEs in visual spatial perception remains limited. Here, we investigated predicted and unpredicted coding of visual locations across the visual hierarchy, utilizing the predictability of the standard population receptive field (pRF) mapping paradigm while sampling BOLD responses at ultra-high field fMRI. Our experiment featured different conditions in which unpredictable stimulus omissions and/or violations (different bar location and orientation) were either embedded in the standard stimulus sequence, or presented separately. These conditions serve to produce prediction errors, both of stimulus presence and of stimulus location. For all conditions, we first calculated test-retest reliability of BOLD responses to identical stimulus sequences in different brain regions. We reasoned that if PEs drive BOLD responses, this should increase test-retest reliability across runs relative to a fully predictable stimulus design. We indeed find this pattern of results selectively in higher-level and not lower level visual cortex. Next, we fit a spatial divisive normalisation (DN-pRF) model to the BOLD timecourses in the standard pRF stimulus sequence, and tested whether bold timecourses in conditions with unexpected stimuli follow this model, which is linear in time. This analysis also indicates that PEs drive high-level visual cortex responses more than low-level visual cortex. These findings suggest that prediction error responses in visual cortex follow the evolution of temporal scales of integration, from fast to slow, along the visual hierarchy. This hints at a tight relationship between temporal divisive normalization and predictive processing.

Talk 6, 9:30 am, 51.26

STIMULUS REPRESENTATIONS IN NEURAL PRIORITY MAPS ARE EQUALLY ENHANCED BY ATTENTION INDEPENDENT OF THE NUMBER OF ATTENDED LOCATIONS

Amelia Harrison[1] (aharrison@ucsb.edu), Daniel Thayer[1], Thomas Sprague[1]; [1]UC Santa Barbara

When spatial attention is distributed across multiple visual field locations, performance in visual tasks is often impaired. This bottleneck is evident in behavioral and neural studies, especially when using complex stimuli, and is echoed in behavioral and neural measurements of visual working memory. Some studies have suggested that this is because distributing attention results in lower attentional enhancement in visual cortex compared to focused attention (e.g., McMains & Somers, 2005), while others support a bottleneck at a post-perceptual decision-making stage (e.g., White et al, 2017; Harrison et al., 2022; Chen & Seidemann, 2012). To characterize how stimulus representations in neural priority maps reflect these constraints and discriminate between these models, we scanned participants with fMRI while they performed a selective attention task in which they were cued on each trial to discriminate a target that appeared at the fixation point, 1 cued location, or 2 cued locations. Using a spatial inverted encoding model, we reconstructed images of priority maps from retinotopic brain regions which contained representations of each stimulus. Comparing map activation between focal attention and fixation conditions replicated the canonical finding that attention to one stimulus caused map activation enhancement at the attended stimulus location. Next, we examined map activation when both stimuli were attended. Strikingly, both stimulus representations were enhanced when attended, with an equivalent increase in map activation as observed with focused attention directed to a single stimulus. This pattern was consistent across retinotopic cortex, with no evidence for graded attentional enhancement in any region. Thus, our results are consistent with a model whereby fMRI signals are enhanced when a stimulus is attended, and the degree of enhancement does not wane as the number of attended stimuli increases. Such a 'relevance' marker may be used to identify neural populations for selective readout from relevant locations during decision-making.

Research supported by Cooperative Agreement W911NF-19-2-0026 for the Institute for Collaborative Biotechnologies, a University of California, Santa Barbara Academic Senate Research Grant, and an Alfred P Sloan Research Fellowship.

Talk Session: Tuesday, May 21, 2024, 10:45 am – 12:15 pm, Talk Room 1

Face and Body Perception

Moderator: Chris Baker, National Institutes of Health

Talk 1, 10:45 am, 52.11

THIRD SOCIAL PATHWAY COMPUTES DYNAMIC ACTION UNIT FEATURES FOR EMOTION DECISION BEHAVIOR

Yuening Yan[1] (yuening.yan@glasgow.ac.uk), Jiayu Zhan[2], Oliver Garrod[1], Robin A.A. Ince[1], Rachael Jack[1], Philippe Schyns[1]; [1]University of Glasgow, [2]Peking University

Faces convey stable identity via static 3D shape/complexion features and transient emotions via dynamic movements features (i.e. Action Units, AUs). With a transparent generative Virtual Human (VH), we studied how brain pathways dynamically compute (i.e. represent, communicate, integrate) AUs and 3D identity features for emotion decisions. In a behavioral task, the generative VH presented randomly parametrized AUs applied to 2,400 random 3D identities. This produced a different animation per trial that each participant (N=10) categorized as one emotion (happy, surprise, fear, disgust, anger, sad). Using participant's responses, we modelled the AUs causing their perception of each emotion. In subsequent neuroimaging, each participant categorized their own emotion models applied to 8 new identities while we randomly varied each AU's amplitude and concurrently measured MEG. Using information theoretical analyses, we traced where and when MEG source amplitudes represent each AU and how sources then integrate AUs for decisions. We compared these representations to covarying but decision-irrelevant 3D face identities. Our results replicate across all participants ($p<0.05$, FWER-corrected): (1) Social Pathway (Occipital Cortex to Superior Temporal Gyrus) directly represents AUs with time lags, with no Ventral involvement; (2) AUs represented early are maintained until STG integrates them with later AUs. In contrast, emotion-irrelevant 3D identities are reduced early, within Occipital Cortex. In summary, we show that the third "Social" Brain Pathway (not the dorsal pathway) dynamically represents facial action units with time lags that are resorbed by the time they reach STG, where they are integrated for emotion decision behavior; while the irrelevant 3D face identity is not represented beyond OC.

This work was funded by the Wellcome Trust (Senior Investigator Award, UK; 107802) and the Multidisciplinary University Research Initiative/Engineering and Physical Sciences Research Council (USA, UK; 172046-01), awarded to P.G.S; and the Wellcome Trust [214120/Z/18/Z], awarded to R.I..

Talk 2, 11:00 am, 52.12

LARGE-SCALE DEEP NEURAL NETWORK BENCHMARKING IN DYNAMIC SOCIAL VISION

Kathy Garcia[1], Colin Conwell[1], Emalie McMahon[1], Michael F. Bonner[1], Leyla Isik[1]; [1]Johns Hopkins University

Many Deep Neural Networks (DNNs) with diverse architectures and learning objectives have yielded high brain similarity and hierarchical correspondence to ventral stream responses to static images. However, they have not been evaluated on dynamic social scenes, which are thought to be processed primarily in the recently proposed lateral visual stream. Here, we ask whether DNNs are similarly good models of processing in the lateral stream and the superior temporal sulcus as they are in the ventral stream. To investigate this, we employ large-scale deep neural network benchmarking against fMRI responses to a curated dataset of 200 naturalistic social videos. We examine over 300 DNNs with diverse architectures, objectives, and training sets. Notably, we find a hierarchical correspondence between DNNs and lateral stream responses: earlier DNN layers correlate better with earlier visual areas (including early visual cortex and middle temporal cortex), middle layers match best with mid-level regions (extrastriate body area and lateral occipital cortex), and finally later layers in the most anterior regions (along the superior temporal sulcus). Pairwise permutation tests further confirm significant differences in average depth of the best layer match between each region of interest. Interestingly, we find no systematic differences between diverse network types in terms of either hierarchical correspondence or absolute correlation with neural data, suggesting drastically different network factors (like learning objective and training dataset) play little role in a network's representational match to the lateral stream. Finally, while the best DNNs provided a representational match to ventral stream responses near the level of the noise ceiling, DNN correlations were significantly lower in all lateral stream regions. Together, these results provide evidence for a feedforward visual hierarchy in the lateral stream and underscore the need for further refinement in computational models to adeptly capture the nuances of dynamic, social visual processing.

NIH R01MH132826

Talk 3, 11:15 am, 52.13

HYPER-REALISTIC REVERSE CORRELATION REVEALS A NOVEL GENDER BIAS IN REPRESENTATIONS OF LEADERSHIP ACROSS POLITICAL ORIENTATION

Stefan Uddenberg[1] (stefan.uddenberg@gmail.com), Zen Nguyen[1], Daniel Albohn[1], Alexander Todorov[1]; [1]University of Chicago Booth School of Business

Appearance influences election outcomes via leadership stereotypes -- past work has shown that adults and even children can predict real-world elections solely on the basis of perceived competence judgments via photographs with relatively high accuracy. What are our visual stereotypes of leadership? And how do they differ according to political orientation? Here we explored this question using a novel reverse correlation technique powered by hyper-realistic generative face models (Albohn et al., 2022). Participants (N=300) viewed generated faces one at a time and judged whether they looked like a "good leader", a "bad leader", or "not sure". Applying a simple algorithm to the aggregated choices yielded visually compelling and interpretable mental representations at both individual and group levels. While political group-averaged representations were similar along many subjective attributes (e.g., perceived "trustworthiness", "attractiveness"; Peterson et al., 2022), they revealed a novel gender bias: right-leaning participants' "good leaders" were more masculine than those of left-leaning participants. We directly replicated this result using richer latent face representations (N=300). We then validated individual participant models on new observers (N=150), probing their willingness to vote for different faces generated by past participants in an imaginary election. As predicted, participants were not only more willing to vote for "good" leader faces, but were most willing for faces generated by participants sharing their political orientation. Taken together, our results demonstrate how political orientation is linked to a novel gender bias in leadership representations, showcasing the utility of our reverse correlation technique.

Talk 4, 11:30 am, 52.14

THE MULTIDIMENSIONAL REPRESENTATION OF FACIAL ATTRIBUTES.

Jessica Taubert[1,2], Shruti Japee[2], Amanda Robinson[1], Houqiu Long[1], Tijl Grootswagers[3], Charles Zheng[2], Francisco Pereira[2], Chris Baker[2]; [1]The University of Queensland, QLD Australia, [2]The National Institute of Mental Health, MD United States., [3]Western Sydney University, NSW Australia.

As primates, our social behaviour is shaped by our ability to read the faces of the people around us. Our current understanding of the neural processes governing 'face reading' comes primarily from studies that have focused on the recognition of facial expressions. However, these studies have often used staged facial expressions, potentially disconnecting facial morphology from genuine emotion and circumstance. Therefore, a reliance on staged stimuli might be obscuring our understanding of how faces are perceived and recognised during everyday life. Here our goal was to identify the core dimensions underlying the mental representation of expressive facial stimuli using a data driven approach. In two behavioural experiments (Experiment 1, N = 940; Experiment 2, N = 489), we used an odd-one-out task to measure perceived dissimilarity within two sets of faces; 900 highly-variable, naturalistic, expressive stimuli from the Wild Faces Database (Long, Peluso, et al., 2023 Sci Reports, 13: 5383) and 670 highly-controlled, staged stimuli from the NimStim database (Tottenham, Tanaka, et al., 2009 Psychiatry Res, 168: 3). Using Representational Similarity Analysis, we mapped the representation of the faces in the Wild and NimStim databases, separately, and compared these representations to behavioral and computational models. We also employed the state-of-the-art VICE model (Muttenthaler, Zheng, et al., 2022 Adv Neural Inf Process Syst) to uncover the dimensions that best explained behaviour towards each of the face sets. Collectively, these results indicate that the representation of the Wild Faces was best characterised by perceived social categories, such as gender, and emotional valence. By comparison, facial expression category explained more of the perceived dissimilarity among the NimStim faces than the Wild Faces. These findings underscore the importance of stimulus selection in visual cognition research and suggest that, under naturalistic circumstances, humans spontaneously use information about both social category and expression to evaluate faces.

This research was supported by the Intramural Research Program of the National Institute of Mental Health (ZIAMH002909 to CIB) and the Australian Research Council (FT200100843 to JT)

Talk 5, 11:45 am, 52.15

COMPARING HUMAN EYE-TRACKING HEATMAPS WITH DNN SALIENCY MAPS FOR FACES AT DIFFERENT SPATIAL FREQUENCIES

Michal Fux[1] (fux.michal@gmail.com), Joydeep Monshi[2], Hojin Jang[1], Charlotte H Lahey[3], Suayb S Arslan[1], Walter V Dixon III[2], Matthew Groth[1], Pawan Sinha[1]; [1]The Department of Brain and Cognitive Sciences, MIT, [2]AI, Machine Learning and Computer Vision, GE Research, Niskayuna, NY, USA, [3]Keene State College

Deep neural network (DNN)-based Face recognition (FR) models have improved greatly over the past decades achieving, or even exceeding, human-level accuracies under certain viewing conditions, such as frontal face views. However, as we reported in last year's meeting (XXX et al., 2023), under challenging viewing conditions (e.g. large distances, non-frontal regard) humans outperform DNNs. To shed light on potential explanations for these differences in FR accuracies of humans and DNNs, we turned to eye-tracking paradigms to discern potentially important zones of information uptake for observers, and compare them with DNN-derived saliency maps. Despite the conceptual similarity between human eye tracking-based heat-maps and DNN saliency maps, the literature is sparse in terms of strategic efforts to quantitatively compare the two and translate human gaze and attention strategies to improve machine performance. We obtained gaze-contingent (GC) human eye-tracking heatmaps and DNN saliency maps, for faces, under three stimulus conditions: filtered for low-spatial frequency, high-spatial frequency, and full-resolution images. Human participants saw two sequentially presented faces and were asked to determine whether the individuals depicted were siblings (images from Vieira et. al., 2014) or two images of the same person (Stirling face database). While human eye-tracking heatmaps were collected during each occurrence of face images (sibling/stirling), DNN saliency maps were realized from differences in similarity score between the machine-interpreted face embeddings of pairs of face images using an efficient correlation-based explainable AI approach. We present the characterization and comparison of humans' and DNN's usage of the spatial frequency information in faces, and propose a model-agnostic translation strategy for improved face recognition performance utilizing an efficient training approach to bring DNN saliency maps into closer register with human eye-tracking heatmaps.

This research is supported by ODNI, IARPA. The views are of the authors and shouldn't be interpreted as representing official policies of ODNI, IARPA, or the U.S. Gov., which is authorized to reproduce & distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

TALK 6, 12:00 PM, 52.16

BAYESIAN DECODING REVEALS RETINOTOPIC SELECTIVITY FOR BODY POSITIONS IN BODY-SELECTIVE REGIONS

Yu Zhao[1] (josephzhao@nevada.unr.edu), Arnab Biswas[1], Matthew W. Shinkle[1], Mark D. Lescroart[1]; [1]University of Nevada, Reno

The Extrastriate Body Area (EBA) represents information about human bodies. Though EBA is not usually considered a retinotopic area, past work has demonstrated visual field biases in different parts of EBA. Here, we probe the retinotopic position-specificity of EBA. Past research has used relatively coarse tests of position sensitivity, including contrasts between body parts presented in isolation in a few fixed retinotopic locations. To address this limitation, we modeled BOLD fMRI responses to stimuli consisting of rendered bodies performing actions in different retinotopic positions. To minimize naturalistic confounds of visual field location and motion, we varied the camera trajectory and added moving textures to bodies and backgrounds. We then extracted features describing the presence and retinotopic location of body parts, and applied linear regression to map these features onto fMRI responses and predict responses to withheld stimuli. This model yielded accurate predictions across cortical regions in and around EBA. Variance partitioning against a motion energy model revealed unique variance explained in these voxels by body features. To explore retinotopic position sensitivity in body selective regions, we computed contrasts between weights for body features reflecting different locations of bodies. As expected, these revealed left versus right visual field contralateral selectivity. We used two multivariate analyses to further quantify position selectivity.. Principal component analysis on the model weights revealed a dominant dimension of horizontal selectivity, alongside a less pronounced dimension that subtly suggests vertical selectivity. Consistent with this result, Bayesian decoding of body locations was more reliable than chance in the horizontal direction and in some cases more reliable than chance in the vertical direction as well. Overall, our findings suggest that EBA has more position sensitivity than has previously been appreciated. Even coarse coding of retinotopic body location could reveal socially relevant information about the position of bodies relative to gaze.

TALK SESSION: TUESDAY, MAY 21, 2024, 10:45 AM – 12:15 PM, TALK ROOM 2

Visual Search 2

Moderator: Wilson Geisler, University of Texas at Austin

TALK 1, 10:45 AM, 52.21

TEMPORAL DYNAMICS OF MULTIPLE ATTENTIONAL TEMPLATE ACTIVATION DURING PREPARATION FOR SEARCH

Anna Grubert[1] (anna.k.grubert@durham.ac.uk), Ziyi Wang[1], Mikel Jimenez[1], Ella Wiliams[2], Roger Remington[3], Martin Eimer[4]; [1]Durham University, [2]Oxford University, [3]University of Minnesota, [4]Birkbeck, University of London

Visual search for known objects is guided by attentional templates (target representations held in working memory), which are activated prior to search. We used an RSVP paradigm to track the temporal dynamics of template activation when multiple colours are task relevant. Search displays containing a pre-defined colour target and five differently coloured distractors were shown every 1600ms. Every 200ms between successive searches, a target- or distractor-colour probe was presented. N2pc components (electrophysiological markers indexing attentional capture) were measured at each probe's temporal position prior to search to determine when in time attentional templates were activated. Target-colour probe N2pc amplitudes increased during the preparation period and were largest for probes directly preceding the next search display. This pattern of transient template activation was identical in single- and two-colour search and probe N2pcs were comparable in size when participants searched for one versus two colours, when the two possible target colours were equiprobable or differed in their likelihoods, and when they changed randomly versus predictably. Transient template activation was also observed in three-colour search, but only when target colours appeared randomly. When they alternated predictably between search episodes, only probes that matched the upcoming target colour triggered N2pcs. This suggests that two attentional templates can be activated in parallel without any apparent costs. However, with three templates, participants prefer to make use of strategic opportunities to reduce working memory load. Notably, distractor-colour probes never triggered N2pcs, demonstrating perfect colour-selectivity in one- two-, and three-colour search.

This work was funded by research grants of the Leverhulme Trust (RPG-2020-319) awarded to AG and the Economic and Social Research Council (ES/V002708/1) awarded to ME.

TALK 2, 11:00 AM, 52.22

THE CHARACTERISTICS OF DISTRACTOR TEMPLATES ARISING FROM LEARNED SUPPRESSION

Rory Ferguson[1] (raferguson@csuchico.edu), Bo Yeong Won[1]; [1]California State University, Chico

In visual search, individuals use cognitive representations known as distractor templates to filter out irrelevant distractors and focus on relevant targets. While previous research has predominantly focused on target templates, this study aims to investigate the nature of distractor templates. Specifically, we investigate the type of information—whether perceptual, semantic, or a combination of both—that is derived from the repeated suppression of distractors. During the training phase, participants sought a target object from a specific category (e.g., shoe) among other distractor objects from a different category (e.g., broom). Following the training phase, without explicit transitions, four different types of distractors were introduced: 1) new exemplars of trained distractors (e.g., a new broom), 2) semantically related distractors (e.g., bucket), 3) perceptually related distractors (e.g., spatula), and 4) unrelated distractors (e.g., light bulb). We hypothesized that if the distractor template included perceptual information but not semantic information, perceptually similar distractors (e.g., spatula) would be suppressed more effectively than semantically related (e.g., bucket) or unrelated distractors (e.g., light bulb), resulting in faster search. Conversely, if the distractor template contained semantic information, semantically related distractors (e.g., bucket) would be suppressed more efficiently than perceptually related (e.g., spatula) or unrelated distractors (e.g., bucket), leading to faster search. If the distractor template encompassed both semantic and perceptual information, both semantically related (e.g., bucket) and perceptually related (e.g., spatula) distractors would exhibit faster search than unrelated distractors (e.g., light bulb). We found that the distractor template formed through repeated exposures contains both semantic and perceptual information to some extent. It is noteworthy that our distractor processes extend beyond mere feature-based information, incorporating semantic details of distractors encountered repeatedly. These novel findings highlight how attentional guidance during visual search tasks is influenced not only by feature-based but also semantic-based processes.

Talk 3, 11:15 am, 52.23

LOOKING FOR TAMPA BUCCANEERS: FAMILIAR SPORT LOGOS ARE FOUND MORE EFFICIENTLY IN HYBRID SEARCH

Dyllan Simpson[1] (dysimpson@ucsd.edu), Lauren Williams[1], Viola Stöermer[2], Timothy Brady[1]; [1]University of California, San Diego, [2]Dartmouth College

In hybrid search tasks, observers search the environment for multiple target items they hold in memory (e.g., locating ingredients for a recipe in the supermarket). Previous research showed that search performance deteriorates with the number of targets memorized (Wolfe, 2012). The current studies tested how the familiarity and activation level of memory items affect hybrid search efficiency. In the first two experiments, we contrasted performance for a set of 16 search targets seen once with a set where memory strength was increased by repeating items 8x (Exp. 1) or by repeating and asking questions about the items (e.g., "What is the primary use for this object?") to encourage deeper processing (Exp. 2). In a third experiment we selected participants based on their self-reported expertise in sports to capitalize on their strong memories for certain sport logos (NFL vs. NHL fans). We then compared search for 16 targets in their domain of expertise with 12 targets from the other sport. In all experiments, after the memorization and memory test phases, participants performed a visual search at small or large set sizes (6 vs 12 or 8 vs 16). Search performance was measured using inverse efficiency scores (IES) to account for speed/accuracy tradeoffs. Across all experiments (overall N=143), IES scores were generally lower for the high strength memory condition than the low strength memory condition. This indicates better search performance for target sets with stronger memory representations, even when the stronger memory list was larger. Broadly, these data contrast with the idea of a simple search through lists of items in memory. They show that differences in memory strength — and therefore differences in how accessible items are — can account for key aspects of memory set size effects in hybrid search.

Talk 4, 11:30 am, 52.24

UNDERSTANDING COVERT SEARCH IN NOISE BACKGROUNDS USING HEURISTIC DECISION ANALYSIS

Anqi Zhang[1,2] (anqizhang@utexas.edu), Wilson Geisler[1,3]; [1]Center for Perceptual Systems, University of Texas at Austin, [2]Department of Physics, University of Texas at Austin, [3]Department of Psychology, University of Texas at Austin

A classic covert search paradigm is to measure search accuracy as a function of the number of potential target locations at a fixed retinal eccentricity, which minimizes the differences in sensitivity across the potential locations. For well-separated targets there are many cases where the effect of the number of locations (the set size) is predicted by parallel unlimited processing (a Bayes optimal decision process). Here we measured search accuracy for 19 well-separated potential target locations that tiled the central 16 deg in a triangular array. The search display was presented for 250ms (the duration of a typical fixation in overt search). Each location contained a 3.5 deg patch of white noise. On half the trials there was no target, and on half the trials a small wavelet target was added to the center of one of the 19 locations. The task was to indicate the location of the target or that it was absent. To precisely characterize eccentricity effects, we measured in a separate experiment the detectability of the target at each location. Under the assumption of statistical independence, we found that human search accuracy slightly exceeded that of the Bayes optimum, and that the observers suffered a modest loss of sensitivity in the fovea (foveal neglect). Furthermore, the observers were able to do this even though the Bayes optimal decision process uses precise knowledge of the sensitivity (d') at each potential location, which varied substantially across the search locations. These seemingly impossible results may be explained by two plausible factors. First, we show that a simple heuristic decision rule that assumes a fixed sensitivity at all potential locations is very close to optimal. Second, we show that intrinsic temporal variations in overall sensitivity could explain how search performance can be slightly above the optimal performance predicted assuming statistical independence.

National Institute of Health (EY024662, EY11747).

Talk 5, 11:45 am, 52.25

KNOWING WHAT YOU MISSED IN MIXED HYBRID VISUAL SEARCH

Ava Mitra[1], Jeremy Wolfe[1,2]; [1]Brigham and Women's Hospital, [2]Harvard Medical School

Mixed-hybrid search is a model task for investigating errors in everyday visual search when simultaneously searching for multiple types of targets (e.g., finding a specific freeway exit while also searching for the category of "obstacles" such as barriers, workers, or raccoons). In mixed hybrid search, people are better at finding specific items but often miss the more categorical targets. Using methodologies from the Inattentional Blindness literature, a previous study found that when participants miss items during search, they can identify the correct item significantly above chance in a subsequent 2AFC identification task, even while reporting little or no awareness of missing any items. What type of information is retained about these missed items that later enables one to identify them correctly? Might participants have rough categorical representations about the missed item, even when guessing about the specific item within the category? Our participants searched for two specific items and two categories of items with 0, 1, or 2 targets present. Stimuli were visible until participants responded. Following each trial, participants rated their confidence in their search response (0-100), then performed a 2 or 6AFC task to identify potential missed targets. Finally, they reported their confidence in their forced-choice selections. In the 6AFC task for categorical missed targets, participants identified the exact missed item 46% of the time (chance is 16.6%, $t(12)=5.1$, $p=0.003$). Participants correctly answered the 6AFC when uncertain about their search performance (34 on a 100-point confidence scale). However, when confident that no target was missed (confidence 88), they guessed in the 6AFC task, both about the item and the category (40%, chance=40%). Overall, observers have some awareness of missed information in search even if the search is self-terminated. However, when they are sure they missed nothing, there does not appear to be any subsequently recoverable information.

NEI EY017001

Talk 6, 12:00 pm, 52.26

HOW INFLEXIBLE IS THE ATTENTIONAL BIAS TOWARDS RECENTLY SELECTED LOCATIONS?

Daniel Toledano[1] (danielt@mail.tau.ac.il), Dominique Lamy[1]; [1]Tel Aviv University

Attention is strongly biased towards the location where a previous target was recently found. This priming-of-location (PoL) effect is thought to reflect a primitive mechanism by which selecting an object automatically and proactively enhances the attentional priority at its location. This account predicts that PoL should be unaffected by changes in task context. However, in most previous PoL studies the task context remained constant. Here, we tested this prediction using a probe paradigm. We manipulated task context by interleaving search trials where participants searched for a shape target among nontargets (2/3 of trials), search-probe trials where they reported letters briefly superimposed on the search display after a short delay (1/6), and probe trials where only the letters appeared (1/6). In Experiments 1 and 2, we found that a letter was more likely to be reported when it appeared at the previous target location than elsewhere. Crucially, this bias was similar when task context repeated (search→search-probe sequences) and when it changed (search→probe sequences). However, in these experiments participants expected a search task on most trials. Therefore, when the context changed, the expected context did not. In Experiment 3, we reversed the task probabilities (probe task on 2/3 of the trials) and in Experiment 4, we used an AABB design, such that the upcoming task was known with 100%-certainty. The bias to report the letter from the previous target location was reduced as the task-change expectation increased. Interestingly, in probe->search sequences, RTs in the search task were faster when the target appeared at the location of a previously reported letter than elsewhere, in all experiments - but this effect was not modulated by task-change expectations. Overall, our findings indicate that selecting an object proactively enhances the attentional priority at its location but expectations about the tasks' context reduce this bias.

Talk Session: Tuesday, May 21, 2024, 2:30 – 4:15 pm, Talk Room 1

Visual Memory: Working and neural mechanisms

Moderator: Thomas Sprague, UC Santa Barbara

Talk 1, 2:30 pm, 54.11

THE NEURAL BASIS OF VISUAL WORKING MEMORY OF REAL-WORLD OBJECT

Wanru Li[1] (wanruli@stu.pku.edu.cn), Jia Yang[1], PingLei Bao[1]; [1]Peking University

Sustaining visual information in mind temporally, denoted as visual working memory (VWM), is a core ability of many cognitive functions, such as decision making and problem solving. While existing evidence focusing on low-level features suggests that VWM recruits analogous regions utilized in perceptual tasks, the neural basis of real-world objects within VWM remains elusive. To address this, we investigated the representation of twenty real-world objects across three VWM and three perception tasks using fMRI. In experiment 1, twelve participants performed a retro-cue sequential VWM task, viewing two objects and recalling the cued object after a ten-second delay. We found the identity of the cued objects could be decoded from the lateral occipital complex (LOC) and intraparietal sulcus (IPS), but not in early visual areas. Moreover, representational similarity analysis (RSA) revealed a common representational pattern between VWM and perceptual tasks exclusively in the LOC, indicating that task-relevant high-level visual areas are specifically recruited for VWM maintenance. Experiment 2 leveraged the contralateral bias effect to assess the extent of shared representational properties between VWM and perception. Six participants performed a retro-cue spatial VWM task where they memorized a cued object from two objects presented in separate visual fields. Interestingly, while a strong contralateral bias was confirmed in the perception task, this bias was significantly reduced in the VWM task, evident through an enhanced ipsilateral representation of the cued object. Experiment 3 delayed the retro-cue to the end of the delay, resulting in persistent but significantly reduced involvement of the ipsilateral LOC. Additionally, searchlight RSA revealed larger cortical representational areas in VWM than perception tasks. These results underscore the essential role of both the LOC and IPS in maintaining real-world object representations in VWM, while the sensory-based object representations in high-level visual areas may go beyond the feedforward visual information flow during VWM.

National Science and Technology Innovation 2030-Major Project 2022ZD0204803, Natural Science Foundation of China Grant NSFC 32271081,32230043 to P. B. Natural Science Foundation of China Grant NSFC32200857, China Postdoctoral Science Foundation Grant 2023M740125, 2022T150021, 2021M700004 to J.Y.

Talk 2, 2:45 pm, 54.12

FMRI DECODING REVEALS INDEPENDENCE IN OBJECT REPRESENTATIONS IN VISUAL WORKING MEMORY

Yaoda Xu[1] (xucogneuro@gmail.com); [1]Yale University

Often time in everyday visual perception, we need to retain multiple visual objects together in visual working memory (VWM). Yet recent fMRI decoding studies of VWM have predominately focused on the retention of a single object in human occipitotemporal cortex (OTC) and posterior parietal cortex (PPC). How are multiple objects represented together in VWM in these brain regions? Are they represented in an orthogonal and thus independent manner? Or are they coded interactively? To address this, we asked 12 human participants to retain two target objects in VWM. We trained a linear classifier to decode the fMRI response patterns of a pair of target objects A and B when each was retained with object C (i.e., decoding AC vs BC) and tested the classifier's decoding performance for the same object pair either in the same condition (within-decoding, decoding AC vs BC) or when each was retained with object D (cross-decoding, decoding AD vs BD). Across OTC and PPC, we found no drop in cross-decoding compared to within-decoding during VWM delay, indicating that the two objects in VWM are represented in an orthogonal manner. Such a representational scheme enables independence in VWM representation, effectively preventing interference between the different target objects during VWM retention. Interestingly, during VWM encoding, a cross-decoding drop was observed in OTC (but not in PPC), indicating that an object's representation is modulated by the identity of another object during encoding in this brain region. However, such a modulation appears to dissipate over the course of VWM retention, likely through feedback mechanims from brain regions such as PPC. Together, these results show independence in target object representations in VWM in the human OTC and PPC, and the emergence of such representations from perception to VWM.

This research is supported by NIH grant 1R01EY030854.

TALK 3, 3:00 PM, 54.13

INTROSPECTION OF RELATIVE UNCERTAINTY OF NEURAL WORKING MEMORY REPRESENTATIONS IN HUMAN CORTEX

Yanming (Alison) li[1] (alisonlee1115@gmail.com), Daniel Thayer, Thomas Sprague; [1]University of California, Santa Barbara

Working memory (WM) is the ability to maintain and manipulate a limited amount of information over a short time. Previous research has shown that noisy activation patterns in retinotopic cortex during a WM delay period encode both the memorized feature and the uncertainty regarding how accurately it is represented (Li et al., 2021; Geurts et al, 2022), and that participants can accurately introspect which of several representations they can report most precisely (Fougnie et al, 2012; Suchow et al, 2017; Li & Sprague, 2023). However, how participants read out the relative uncertainty for multiple WM representations from neural activity patterns remains unclear. Here, we acquired fMRI data during a memory guided saccade task in which participants remembered 1 or 2 locations over a 12 s delay period and reported the location of one item with a saccade. Extending our previous study (Li & Sprague, 2023), on each trial, at the end of the delay period participants were either instructed to report a cued item or choose the item they believed they remembered best with a saccade. After the memory report, participants reported their uncertainty about the reported location by adjusting the extent of an arc (as in Li et al, 2021). Results showed that when participants were asked to report their best-remembered item, recall error and uncertainty were both lower compared to randomly-cued trials, consistent with accurate introspection of the relative quality of multiple WM representations. Moreover, delay-period WM representations reconstructed from activation patterns in extrastriate cortex were stronger for memory items reported on "report best" trials as compared to the non-reported item. These findings suggest that participants can simultaneously compare and report the quality of multiple remembered locations, and demonstrate that these reports are based on the quality of neural WM representations in retinotopic cortex.

Research was sponsored in part by the U.S. Army Research Office and accomplished under contract W911NF-19-D-0001 for the Institute for Collaborative Biotechnologies, an Alfred P Sloan Foundation Research Fellowship, and a University of California, Santa Barbara Academic Senate Research Grant

TALK 4, 3:15 PM, 54.14

FLEXIBLE MEMORY INTERPLAYS: SELECTIVE REACTIVATION OF LONG-TERM MEMORIES IN WORKING MEMORY

Eren Günseli[1] (gunseli.eren@gmail.com), Duygu Yücel[1], Nursena Ataseven[1], Yağmur Damla Şentürk[1], Nursima Ünver[1], Şahcan Özdemir[1], Lara Todorova[1], Berna Güler[1], Betül Türk[1], Can Demircan[1], Keisuke Fukuda[2], Christian N L Olivers[3], Tobias Egner[4]; [1]Sabanci University, [2]University of Toronto, [3]Vrije Universiteit Amsterdam, [4]Duke University

Working memory is defined as the online storage space for ongoing tasks. It stores both newly encoded information and retrieved long-term memories. However, there is a growing amount of work to suggest that long-term memories can also guide behavior. This raises the question: Why do humans invest metabolic resources in reactivating long-term memories in working memory instead of guiding behavior directly via long-term memory? We conducted six experiments examining working memory reactivation of long-term memories in anticipation of task demands encompassing protection against interference, behavioral guidance, and adaptation to novel settings. Using behavioral and electrophysiological indices, we measured the extent to which long-term memories are reactivated in working memory in anticipation of these task demands relative to the anticipation of a recognition task, which constituted a baseline. Compared to this baseline, we found equal memory reactivation when anticipating perceptual interference and dual-task interference, and less memory reactivation when anticipating attentional guidance. On the other hand, reactivation was stronger for task switching, contextual changes, and performing mental operations. These results suggest that the reactivation of long-term memories in working memory is not primarily for protection against interference or behavioral guidance. Instead, stronger reactivation occurs when there is a need to update the memories themselves (i.e., perform a mental operation) or the settings in which they are used (i.e., the task rules and the context). This insight implies that the goal of reactivating long-term memories in working memory may be to facilitate adaptation to novel situations. Our research challenges influential memory models and recent empirical work that consider working memory as the default buffer for retrieved long-term memories and instead highlights a flexible and dynamic interplay between long-term memories and working memory.

TALK 5, 3:30 PM, 54.15

ASSOCIATIVE LEARNING CHANGES MULTIVARIATE NEURAL SIGNATURES OF VISUAL WORKING MEMORY

William Ngiam[1,2] (wngiam@uchicago.edu), William Thyer[1,2], Henry Jones[1,2], Darius Suplica[1,2], Will Epstein[1,2], Edward Awh[1,2]; [1]Department of Psychology, University of Chicago, [2]Institute of Mind and Biology, University of Chicago

A hallmark of visual working memory is its sharp capacity limit, though this limit can be circumvented using learned knowledge. For example, when arrays of to-be-remembered items contain statistical regularities, people can learn the associations between items and recall more information overall (Brady et al., 2009; Ngiam et al., 2019). One proposed mechanism for how this recall benefit is achieved is through 'memory compression' – redundancies introduce a reduction of information per item, enabling more items to be stored online. Another proposed mechanism is that pointers are efficiently allocated to each 'chunk' with the benefit coming from long-term memory retrieval rather than changes to working memory itself. In an attempt to distinguish between these possibilities, we turned to an EEG measure that tracks the number of individuated items stored in working memory (mvLoad; Thyer et al., 2022). The memory compression account predicts an overall increase in the number of items stored online, whereas the long-term memory retrieval account predicts a reduction in working memory load. Subjects completed a training session where they learned specific color-color pairs. In a subsequent EEG session, subjects completed a recall task with 2 random colors, 4 random colors, or 2 learned color pairs. mvLoad analysis showed a reduction in working memory load for the 2 learned pairs condition (from 4 towards 2), consistent with the notion that an item-based pointer is assigned to each chunk. Moreover, multidimensional scaling shows an additional independent signal that distinguishes the 2 learned pairs condition from the other conditions. We propose that this additional signal reflects the involvement of long-term memory, consistent with the notion that the learned association is being relied upon to maintain the information.

This research was supported by the National Institute of Health R01-MH087214 grant awarded to Edward Awh and Edward Vogel.

TALK 6, 3:45 PM, 54.16

SUSTAINED STIMULUS-SELECTIVE MULTI-UNIT ACTIVITY IN HUMAN PRIMARY VISUAL CORTEX

Lucas Nadolskis[1] (lgilnadolskis@ucsb.edu), Galen Pogoncheff[1], Jacob Granley[1], Alfonso Rodil[2], Leili Soo[2], Lily Turkstra[1], Thomas Sprague[1], Arantxa Alfaro Saez[2,3], Cristina Soto Sanchez[2], Eduardo Fernandez Jover[2], Michael Beyeler[1]; [1]University of California - Santa Barbara, [2]Universidad Miguel Hernandez de Elche, Spain, [3]Section of Neurology, Hospital Vega Baja, Orihuela, Spain

Introduction: Neural activity related to visual working memory (WM) has been found in various cortical areas of primates and humans. However, the role of multi-unit activity in human V1 during WM tasks is not fully understood. We explored this by examining intracortical recordings from an awake blind human with a visual prosthesis (Utah array in parafoveal V1) during a delayed-response WM task. Methods: In 90 trials, one of three chosen electrodes stimulated a visual percept (phosphene) in the participant (stimulation period), who then had to remember its shape and location for 3 or 5 seconds (delay period), followed by an auditory cue and a recall period, during which the participant was asked to intently visualize the remembered phosphene. Neural activity was recorded and analyzed for multi-unit activity (MUA), entire spiking activity (ESA), and local field potential (LFP). Results: Significant differences in MUA, ESA, and LFP (theta, alpha, and beta bands) were observed across different trial periods (stimulation, delay, recall, and spontaneous; t-tests, $p<.05$). Each electrode's neural signature was distinct during delay and recall (over 90% accuracy in leave-one-trial-out cross-validation), with day-to-day drifts. The directions of maximum variability in the recall period neural activity formed a representative neural basis for each electrode, and enabled classification with a random forest classifier for both delay (97% accuracy) and stimulation period activities (88% accuracy). These shared signatures could be learned from the delay or recall period activity, but not from the electrically evoked activity, suggesting that WM elicits a subset of the full activity present during electrical stimulation. Conclusion: Our findings underscore V1's crucial role in retaining information at the neuronal level over delay periods. The transformation of representations during the recall period suggests that the encoded information is more abstract than the sensory activity evoked during stimulation.

NIH DP2-LM014268, Alfred P. Sloan Foundation Research Fellowship, PDC2022-133952-100 and PID2022-141606OB-I00 from the Spanish "Ministerio de Ciencia, Innovación y Universidades", European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 899287 (NeuraViPeR).

TALK 7, 4:00 PM, 54.17

LINKING BEHAVIORAL AND NEURAL ESTIMATES OF TRIAL-BY-TRIAL WORKING MEMORY INFORMATION CONTENT

Ying Zhou[1,2] (yz5725@nyu.edu), Clayton Curtis[2], Daryl Fougnie[1], Kartik Sreenivasan[1,3]; [1]Division of Science and Mathematics, New York University Abu Dhabi, [2]Department of Psychology, New York University, [3]Center for Brain and Health, New York University Abu Dhabi

How is working memory (WM) information represented in the brain? Neural and computational models have used data aggregated over hundreds of trials to argue for different perspectives on how neural activity encodes individual memories. The two main perspectives are information rich representations such as in probabilistic coding models (a probability distribution over the whole feature space), and information sparse representations, such as in high-threshold (a precise feature value) or drift models (a value with a confidence interval unrelated to the direction of drift). The use of aggregate data represents a key inferential bottleneck that critically limits the ability to adjudicate between different formats of individual memory coding in

WM. This study used a powerful method to link behavioral and neural estimates of WM representation on individual trials. We asked participants (n = 12) to memorize a motion direction over a brief delay. After the delay, instead of making a single report about the memorized direction, they indicated their memory by placing 6 "bets", resulting in a distribution over the 360° direction space that reflected their probabilistic memory representation on individual trials. Additionally, we used a Bayesian decoder to estimate the posterior of the memorized direction given the fMRI signal during memory maintenance on individual trials. Comparing the shape of the behavioral and neural estimates on individual trials, we found a significant correspondence in their width in occipital, parietal and frontal regions (ps < .007; Cohen's ds > .767), and critically, a significant correspondence in their asymmetry in early visual cortex (p < .001; Cohen's d = .779). These results indicate (1) individual WM representations are complex probability distributions that contain more information than that can be deduced from aggregate data; (2) early visual cortex contains richer information about WM than other brain regions, with meaningful asymmetry information influencing behavior.

NYUAD Research Institute grant CG012

Talk Session: Tuesday, May 21, 2024, 2:30 – 4:15 pm, Talk Room 2

Temporal Processing

Moderator: Kate Bonnen, Indiana University

Talk 1, 2:30 pm, 54.21

TEMPORAL INTEGRATION OF VISUAL INFORMATION IS AFFECTED BY FAST SPATIAL GROUPING

Ljubica Jovanovic[1] (lj.m.jovanovic@gmail.com), Pascal Mamassian[1]; [1]Laboratoire des Systèmes Perceptifs, Ecole Normale Supérieure, PSL University, CNRS

Sensory information is integrated over space and time, determining both the content and time of our visual experience. Here we tested how spatial integration of elementary features affects temporal integration across the visual field. An array of 16 equidistant small sinusoidal gratings was presented on a virtual circle centered on the fixation point. The gratings were randomly split in two halves, and each half group was presented sequentially. Test stimuli were created by omitting one grating, and participants had to detect these tests in a sequential 2AFC paradigm. To measure the time course of temporal integration, we varied the duration of the first group of gratings (10–160 ms) and inter-stimulus interval between the two groups (ISI, 0–40 ms). The contrast of the first group was adjusted to match the perceived contrast of the second group. We compared two spatial conditions: the gratings' orientations were either aligned with the virtual circle in both groups (forming a collinear contour), or the gratings' orientations were aligned in one group and orthogonal in the other group (forming an interrupted contour). Increasing the ISI decreased performance, but more surprisingly, increasing the duration of the first group also decreased performance (see also DiLollo, 1977, Nature). The two effects were not additive: performance decreased faster with an increase in ISI than duration. Interestingly, detection of the missing element was better when both parts of the display contained gratings forming collinear, rather than with interrupted contours. The interaction between the duration of the display and ISI on performance is inconsistent with temporal integration being an output of a fixed or sliding temporal window integrating information over time. Furthermore, the effect of relative similarity of gratings suggests an interaction between a fast spatial grouping and temporal integration across the visual field, further informing models of temporal integration.

ANR grant no. ANR-22-CE28-0025

Talk 2, 2:45 pm, 54.22

SENSORY CORRELATION DETECTION BY CHILDREN TREATED FOR CONGENITAL VISUAL DEPRIVATION

Priti Gupta[1,3], Lukas Vogelsang[2,5], Marin Vogelsang[2,5], Manvi Jain[1,3], Naviya Lall[3], Dhun Verma[3], Chetan Ralekar[5], Suma Ganesh[4], Pawan Sinha[5]; [1]Indian Institute of Technology Delhi, [2]École Polytechnique Fédérale de Lausanne, Switzerland, [3]Project Prakash, Dr Shroff's Charity Eye Hospital, New Delhi, [4]Department of Pediatric Opthalmology, Dr Shroff's Charity Eye Hospital, New Delhi, [5]Massachusetts Institute of Technology

The temporal covariance of sensory signals provides critical information for determining the relationship between different entities in the sensorium. How quickly during the developmental timeline does the ability to detect such correlations become evident? Addressing this question is important for assessing whether this ability can help bootstrap the early stages of perceptual learning. Here we report work designed to assess the ability to detect temporal correlations of varying strengths within and across sensory modalities in 15 patients treated for congenital blindness as part of Project Prakash, a humanitarian and scientific effort focused on treating early blind children and through their help, understanding visual development. The performance of Prakash patients was compared with that of 21 normally-sighted blur-matched controls. In the intra-modal condition, participants were asked to determine which of two disks was blinking most in unison with a circumscribing ring. In the inter-modal condition, participants had to identify the disk blinking most congruently with a concurrent audio track of beeps and silences. This experimental design yielded three main results. First, we found that, while not fully reaching the level of normally-sighted controls, Prakash patients were able to detect correlations with markedly above-chance accuracy rapidly after sight onset. Second, for both groups, performance levels in the inter-modal and intra-modal conditions were comparable. Finally, the extent of the time series that participants observed before making a decision was similar between the two groups but markedly longer than would be required when using a pre-defined statistical decision criterion. These results help characterize a foundational process for detecting relationships between environmental entities, point to the resilience of acquiring this ability to early-onset, prolonged visual deprivation, and suggest that it could potentially serve as a bootstrapping mechanism for learning to extract environmental cliques.

This work was funded by Grant R01EY020517 from NEI (NIH) to Pawan Sinha

Talk 3, 3:00 pm, 54.23

PARSING PULSES: TESTING THE LIMITS OF TEMPORAL PHASE PERCEPTION IN HUMAN VISION

Andrew Lisech[1] (alisech@udel.edu), Xinyi Yuan[1,2], Keith A. Schneider[1]; [1]University of Delaware, [2]Beijing Normal University

Introduction: Humans can detect luminance flicker exceeding 60 Hz, but the threshold for perceiving the flicker's phase is much lower (~7-10 Hz). As a precursor to future experiments investigating this temporal bottleneck and the broader dynamics of visual perception, we replicated Aghdaee and Cavanagh (2007) using stimuli devoid of spatial and temporal transients. Methods: Twelve subjects judged whether two monochromatic Gaussians, oscillating sinusoidally between black and white, were in-phase or 180° out-of-phase. A 1440 Hz PROPixx projector (VPixx Technologies) displayed stimulus pairs at 4° eccentricity, spaced 1.8° or 5° apart, either: 1) left and right of the vertical meridian in the opposite-hemifield condition, or 2) above and below the horizontal meridian within the same hemifield. Using the method of constant stimuli, we measured phase detection thresholds across 11 oscillation frequencies (1–31 Hz), conducting 25 repetitions for each of the randomly interleaved conditions. To prevent visual offset artifacts, stimuli oscillated continuously until subjects responded. Thresholds were determined by fitting a cumulative normal function with a lower asymptote parameter. Results: Inter-stimulus spacing distance revealed a significant main effect, indicating subjects discriminated phase at higher frequencies for closely-spaced stimuli (11.13 Hz) than for farther stimuli (8.15 Hz). The main effect of hemifield was not significant, and no significant interaction with distance was observed. Notably, the asymptote parameter differed significantly from zero in the near condition, with subjects retaining a small (~60%) but significant ability to determine phase at even the highest frequencies tested. Conclusion: The advantage of near stimuli suggests the involvement of a low-level primary sensory mechanism, such as local motion detection circuits. In contrast, comparing two far-spaced stimuli requires higher-level (non-local) and slower mechanisms which possess timing consistent with conscious awareness. Future work should consider mechanisms such as discrete perception and onset artifacts similar to the Fröhlich Illusion.

Talk 4, 3:15 pm, 54.24

THE APPEARANCE OF ORIENTATION REPULSION CHANGES WITH DEVELOPING TEMPORAL EXPECTATION

Tomoya Nakamura[1,2,3], Ikuya Murakami[1]; [1]The University of Tokyo, [2]Japan Society for the Promotion of Science, [3]RIKEN Center for Brain Science

Anticipating when future events will happen improves our performance by facilitating visuomotor processing at various stages, from perception to action. We investigated whether such temporal expectation also influences the appearance of orientation repulsion, wherein a vertical target subjectively appears as tilted against a surrounding inducer. As the inducer, eight circularly arranged Gabor patches were continuously presented on both sides of the fixation point. As the target, another Gabor patch was flashed at the center of either one of the inducers. Participants reported whether the target appeared as tilted clockwise or counterclockwise from the vertical. In Experiment 1, prior to the target onset, auditory temporal cues were provided five times at constant intervals of 400 ms. Participants were instructed to attend to the fifth cue moment, as the target most often (with 69% probability) appeared at this moment. The target otherwise appeared 200 ms earlier or later than the anticipated moment. Repulsion significantly decreased when the target appeared earlier than anticipated. In Experiment 2, to isolate the effect of automatic entrainment to the cue rhythm, the cue was repeated every 450 ms throughout a session, and the target was presented either in-phase or out-of-phase with the rhythm with equal probability. Anticipating the target onset was virtually impossible in this setup, and indeed, no change in repulsion was observed. In Experiment 3, to focus on the effect of hazard rates, a single cue was provided, and the target was presented after one of three intervals (200, 400, or 600 ms) with equal probability. Although the cue was uninformative about the target onset, repulsion significantly decreased as the cue-target interval increased. These findings suggest that developing temporal expectation, especially the expectation associated with hazard rates, promotes premature decisions on the perceptual content that has not fully undergone contextual modulation during low-level visual processing.

Supported by KAKENHI 21J20400, 22KJ0555, 18H05523, and 23H01052

Talk 5, 3:30 pm, 54.25

SACCADIC TIME COMPRESSION AND BRAIN DYNAMICS: FROM REGIONS TO WHOLE BRAIN NETWORK

Amirhossein Ghaderi[1] (ghaderi@yorku.ca), Matthias Niemeier[1,2], John Douglas Crawford[1]; [1]York University, [2]University of Toronto

Saccades influence time perception, but the associated neural mechanisms remain elusive. We explored the cortical dynamics of perisaccadic time perception through a combination psychophysics, EEG, source localization, and graph theory analysis (GTA). 21 participants viewed a reference stimuli sequence followed by a test stimulus, either just before saccades or sustained fixation. Following this, participants were asked to judge the duration of the test compared to the reference. In previous studies we found that stimulus repetition and saccades events interacted at the level of sensorimotor brain dynamics (Ghaderi et al. Cerebral Cortex 2023) and perceived stimulus duration (Ghaderi et al. Heliyon 2022). Here, we combined these two approaches to investigate brain dynamics related to perceived stimulus duration. Source localization revealed the dynamics in cortical activation, predominantly starting from early visual and concluding in higher-level 'cognitive' areas (frontal and anterior cingulate cortices (ACC)). The GTA highlighted the pivotal roles of three groups of brain regions: 1) visual, 2) temporal and parahippocampal, and 3) frontal and ACC. The involvement of these regions suggests that the early visual areas may initially influence time perception concerning higher visual processing for underestimated trials. Subsequently, a top-down mechanism could be engaged in processing these visual signals, leading to increased activity in the frontal and decreased activity in the ACC, likely associated with decision-making errors. Additionally, this mechanism involves heightened activity in memory-related regions during the underestimation of time, indicating a potential need for increased activation in these areas concerning errors related to time compression. The whole network analysis revealed significant differences in the network features between underestimated and correct judged trials. These results imply a potential link between time compression and processing in functional networks which suggests network characteristics (integration, segregation, synchronization stability, and complexity) seem to play a role in shaping our perception during brief durations.

Grant Support: an NSERC Discovery Grant and VISTA Fellowship, funded by CFREF.

Talk 6, 3:45 pm, 54.26

OPTOGENETIC STIMULATION OF INFEROTEMPORAL CORTEX IS PERCEIVED EARLIER THAN STIMULATION

Drew Nguyen[1], Elia Shahbazi[1], Timothy Ma[2], Arash Afraz[1]; [1]National Institutes of Health, [2]New York University

Local stimulation in high-level cortical visual areas perturbs the contents of visual perception. We have previously demonstrated that visual events evoked by optogenetic stimulation in IT cortex can be reconstructed using a method dubbed "perceptography". While perceptography informs about the contents of stimulation-evoked perceptual events, we do not know when they are perceived relative to external physical events. In this study, we use high throughput behavioral optogenetics coupled with visual interference to measure when stimulation-evoked perceptual events are perceived relative to concurrent sensory input. An adult macaque monkey was trained to behaviorally detect and report a brief optogenetic excitatory impulse delivered to its central IT cortex. The animal started each 1.6s trial by fixating on a randomly chosen computer-generated image (8 deg.). A ~1x1mm area of the IT cortex was optogenetically stimulated in half of the trials at random for 60ms halfway through the image presentation using an implanted LED array. We hypothesize that interrupting the image presentation at the proper time can mask the stimulation-evoked perceptual event. This was accomplished by presenting high-contrast visual noise (12 deg.) at one of 11 different time points during image presentation. After each trial, the animal reported whether it was stimulated by looking at one of two presented targets, with liquid reward for correct reports. We find that the monkey's performance varies with the onset time of the visual noise. Presentation of visual noise 200ms prior to stimulation elicited a significantly larger miss rate compared to baseline. This is reflected by a significant decrease in the monkey's d' for the same noise onset time against baseline. Furthermore, we find that perceptography with image perturbations presented 200ms prior to stimulation induces a higher false alarm rate compared to image perturbation presented simultaneously to stimulation.

Talk 7, 4:00 pm, 54.27

RELATIONSHIP BETWEEN V1 SPIKING PATTERNS AND SCALP EEG IS FREQUENCY-DEPENDENT

Dixit Sharma[1,2] (ds1663@rutgers.edu), Bart Krekelberg[1]; [1]Rutgers University, [2]BNS Graduate Program

Despite decades of electroencephalography (EEG) research, the relationship between EEG and underlying spiking dynamics remains unclear. This limits our ability to infer intracranial signals from EEG, a critical step to bridge electrophysiological findings across species and to develop non-invasive brain-machine interfaces (BMIs). We recorded spiking activity from a 32-channel floating microarray permanently implanted in parafoveal V1 and scalp-EEG in a male macaque monkey. While the animal fixated, the screen flickered at different temporal frequencies (0, 5, 10, 20, and 40 Hz) to induce steady-state visual evoked potentials (SSVEP). The primary advantage of SSVEPs is that they generate high signal-to-noise ratios. We analyzed the relationship between the SSVEPs in multiunit spiking activity (MUA) and EEG. Both MUA and EEG showed robust SSVEPs, with best response in EEG for 20Hz-stimulus. The MUA also showed strong responses at the harmonics of the stimulus frequencies, which was not evident in EEG. Time-series correlation between trial-averaged EEG and MUA showed strongest relationship for 5Hz- and 10Hz-stimuli. Furthermore, correlating MUA with EEG power at different frequencies (1-200 Hz) showed prominent correlations for 5Hz-stimulus, which was limited to specific EEG bands (5-10, 10-20, and 40-70 Hz). This correlation pattern was consistent across intracranial electrodes placed at different depths in V1, suggesting that the 5Hz stimulus is optimal for estimating spiking activity from EEG. Single-trial EEG-MUA correlations lacked stimulus-specific relationships. However, a 10 ms delay in EEG signal yielded consistent negative correlations with spiking activity across intracranial electrode depths. This suggests that delayed EEG signals may reflect information about the spiking activity and could be used to estimate MUA from EEG. Our study shows robust relationships between V1 spiking activity and EEG under frequency-specific stimulus conditions. These results give direction to better estimate cortical spiking activity using non-invasive scalp EEG.

Talk Session: Tuesday, May 21, 2024, 5:15 – 7:15 pm, Talk Room 1

Multisensory Processing

Moderator: Shinsuke Shimojo, Caltech

Talk 1, 5:15 pm, 55.11

PERCEPTUAL MECHANISMS UNDERLYING HUMAN CLICK-BASED ECHOLOCATION

Haydee Garcia-Lazaro[1] (haydee@ski.org), Pushpita Bhattacharyya[1], Brendyn Chao[2], Santani Teng[1]; [1]Smith-Kettlewell Eye Research Institute, [2]University of Washington

Echolocation is an active sensing strategy used by some blind individuals to navigate their surroundings. Human echolocators emit tongue clicks, leveraging the echoes to detect, discriminate, and localize objects within their environment. Proficient blind echolocators outperform non-expert blind and sighted individuals in most echo-acoustic tasks; while visual experience and expertise play significant roles in echolocation performance, the underlying mechanisms of this advantage remain unclear. Recent research from our lab suggests that the emitted click masks the subsequent fainter echo, and superior performance among experts may be attributed to a more effective release from masking relative to novices. We explore this hypothesis by evaluating the influence of two aspects of masking on echolocation performance: click-echo signal-to-noise ratio (SNR) and click-echo temporal separation. Novice-sighted individuals completed an echo-acoustic localization task. Each trial consisted of 2, 5, 8, or 11 synthesized mouth clicks with spatialized echoes from reflectors 1-meter away and 5°–25° from the midsagittal plane. Participants indicated the reflector's location (left vs. right). In Experiment 1, the click amplitude was variably attenuated relative to its natural amplitude. In Experiment 2, click amplitude was fixed while the click-echo time delay varied from ~6–60 ms, equivalent to 1–10 meters. We hypothesized that performance would improve as click-echo amplitude attenuates (Exp. 1) and click-echo time delay increases (Exp. 2). Our results revealed that novice-sighted individuals, at echo-click-level-difference > -2 dB or click-echo time delay > 52 ms, performed similarly to proficient echolocators presented with naturalistic stimuli (~25 dB; ~6 ms). These findings suggest that a well-tuned click-echo relationship, alongside an optimized click-echo temporal integration window, enhances echolocation performance. Future research will explore their combined or separate roles in auditory filters. Finely tuned click-echo SNRs and a narrower click-echo temporal integration window may underlie echolocation proficiency due to improved click-echo segregation and echo representation.

E. Matilda Ziegler Foundation for the Blind, National Eye Institute 1R21EY032282-01, Smith-Kettlewell Eye Research Institute

Talk 2, 5:30 pm, 55.12

REDUCED CONTEXTUAL EFFECTS AND CROSS-MODAL CALIBRATION DEMONSTRATE ATYPICAL SENSORY PROCESSING IN AUTISM

Avni Ben Zvi Inbar[1], Hagit Hel-Or, Bat-Sheva Hadad; [1]Student

Introduction: Sensory symptoms are part of the core phenotype of autism, but their underlying mechanisms are unknown. We examined whether altered perception of magnitude in autism arises from modulations in biases and contextual effects known to calibrate perceptual sensitivity in neurotypicals. Specifically, we asked whether calibration of duration perception by context is generalized across modalities or rather mediated by within-modality specific mechanisms. Sensitivity of duration perception for visual and auditory stimuli was tested when context was manipulated within- and between the sensory modalities. Method: Individuals with and without autism preformed a two-interval forced choice task to determine the longer of two temporal signals either visual or auditory. Participants performed the task under three conditions: 1) the central standard was presented in two types of modalities: visual or auditory; 2) the contextual standards formed either a wide or a narrow contextual range around the central standard; 3) the contextual standards were presented in two modalities: visual or auditory independent of the central standard, forming "same-modality" or "between-modality" conditions. Thresholds were determined using two staircase methods: constant stimuli and QUEST. Results: For neurotypicals, thresholds determined in the auditory domain were smaller than in vision, suggesting an auditory specialization. Importantly, the narrower context enhanced sensitivity for standards within the same modality but had no effect on standards of different modality, suggesting that perceptual magnitude normally follows a modality specific calibration process. For individuals with autism, thresholds were similar in the auditory and visual domains suggesting no auditory specialization. Context only mildly affected sensitivity and in a similar magnitude for between- and within-modalities context, suggesting an amodal general calibration mechanism. These results suggest that contrary to the specialized, modality specific calibration processes in neurotypicals, overall reduced and less specialized (amodal) calibration process is shown in autism, that may account for sensory dysregulations and symptoms.

Talk 3, 5:45 pm, 55.13

VISION WITHOUT PHOTORECEPTORS: CROSSMODAL PERCEPTION WITHIN THE BLIND SPOTS

Ailene Chan[1] (chanyca@caltech.edu), Noelle R. B. Stiles[1,2], Carmel A. Levitan[3], Armand R. Tanguay, Jr.[1,4], Shinsuke Shimojo[1]; [1]California Institute of Technology, Division of Biology and Biological Engineering, [2]Rutgers University, Departments of Neurology, Ophthalmology and Visual Science, and Biomedical Engineering; Center for Advanced Human Brain Imaging Research in the Brain Health Institute, [3]Occidental College, Cognitive Science, [4]University of Southern California, Departments of Electrical Engineering, Chemical Engineering and Materials Science, Biomedical Engineering, Ophthalmology, and Physics and Astronomy; Neuroscience Graduate Program

Multisensory illusions are a key tool to investigate crossmodal integration and plasticity, given their resilience across manipulations and unique spatial adaptability. We previously tested the classic Double Flash Illusion across retinal locations in low vision participants. These participants reported stronger double flash perception in visual impairment areas relative to neurotypical participants. To examine whether illusions could span regions of no light perception, the present study induced multisensory interactions within the blind spots of neurotypicals using a postdictive illusion. The Audiovisual Rabbit Illusion consists of a sequence of [beep-flash, beep, beep-flash]; an illusory flash is induced by the second beep, located between the first and second beep-flash pairs (all beeps are centrally located). This illusion is postdictive, as a latter sensory stimulus impacts the perception of an already-presented stimulus. We mapped each participant's blind spots (with one eye blindfolded) and placed the beep-flash pairs 0.5° outside the borders. We tested four sequences: left-to-right, right-to-left, top-to-bottom, and bottom-to-top, and three conditions: zero-beep, two-flash (0B2F; control); 2B2F (control); and 3B2F (illusion). Participants reported strong illusory percepts within their blind spots, as well as in visible locations. Performance was comparable between these locations. For 0B2F and 2B2F conditions, participants reported perceiving ~2 flashes, confirming that the illusion only occurs when visual and auditory information are incongruent within each beep-flash sequence. There were no significant differences in the number of flashes perceived between the 0B2F and 2B2F conditions. These results support the hypothesis that filling-in within blind areas can be multisensory. In this case, audition may play a key role in inducing visual propagation across visual space, even into regions without visual input capacity. The blind spot provides an interesting test case for how the brain interprets blind regions in the retina, particularly in comparison to scotomas generated from eye diseases.

Croucher Scholarships for Doctoral Study, Dominic Orr Graduate Fellowship in BBE, The National Institutes of Health, The National Eye Institute

Talk 4, 6:00 pm, 55.14

NEURAL DYNAMICS OF SUPRAMODAL CONSCIOUS PERCEPTION

Andreas Wutz[1] (andreas.wutz@plus.ac.at); [1]University of Salzburg

Is the conscious perception of seeing a flash, hearing a sound or feeling a touch associated with one common core-activity pattern in the brain? Here, I present novel magnetoencephalography (MEG) data that reveal such supramodal neural correlates of conscious perception. On each trial, different visual, auditory or tactile stimuli were shown at individual perceptual thresholds, such that about half of the stimuli were consciously detected, while the other half was missed. Four different stimuli per modality were used (i.e. different Gabor patches, sound-frequencies, stimulated fingers) in order to subsequently leverage representational similarity analysis (RSA) for differentiating modality-specific, sensory processing from supramodal conscious experiences, which are similar across modalities. As expected, there was stronger evoked MEG-activity for detected vs. missed stimuli during sensory processing (<0.5 s) in the respective sensory cortices. Moreover consistent with previous work, there was stronger alpha-frequency band power (8-13 HZ) for missed vs. detected trials in the pre-stimulus period and in a later time window after stimulus onset (>0.5 s) for all three modalities. Critically, the RSA distinguished activity patterns related to modality-specific, sensory processing shortly after stimulus onset (<0.5 s) from later supramodal conscious processing (>0.5 s). Overall, our findings suggest a three-stage model for conscious multisensory experiences, involving pre-stimulus alpha oscillations, modality-specific, sensory processing upon stimulus onset and then later supramodal conscious perception. This temporal processing cascade may serve the integration and updating of pre-stimulus brain states, presumably reflecting top-down predictions about upcoming sensory events, with subsequent conscious experiences irrespective of the specific sensory modality.

This research was supported by project funding from the FWF - the Austrian Science Fund. Grant agreement number: P36214

Talk 5, 6:15 pm, 55.15

CROSS-MODAL TUNING IN EARLY VISUAL AND SOMATOSENSORY CORTICES

Stephanie Badde[1], Ilona Bloem[2,3], Jonathan Winawer[2], Michael S Landy[2]; [1]Tufts University, [2]New York University, [3]Netherlands Institute for Neuroscience

Conflicts between the senses shape our perception. We used functional magnetic resonance imaging to test whether exposure to spatially offset visual and tactile stimuli shifts population-level spatial tuning in early visual and somatosensory cortices. Participants fixated a marker at the center of a sketched outline of a right hand. During visual stimulation, yellow circles expanding and contracting at 4 Hz were superimposed on one fingertip on the displayed outline. Tactile stimuli were amplitude-modulated vibrations at the fingertips of the participants' right hand, also pulsating at 4 Hz. Stimuli swept across the fingers, moving from one finger to the next every 4 s, in ascending or descending order. Within a 4 min run, visual and tactile stimuli were either presented in isolation or synchronously. Visual-tactile stimulus pairs were either always located at the same finger, or always located at adjacent fingers, with the visual stimulus shifted either towards the thumb or little finger. Population receptive field (PRF) mapping confirmed topographically organized neural populations tuned to tactile stimulation of one finger in somatosensory but not visual cortex, and vice versa for visual stimulation. Maps from unisensory stimulation agreed well with those from congruent tactile-visual stimulation. Visual-tactile spatial discrepancy resulted in a PRF shift in all participants. Shift direction was independent of sweep direction ruling out prediction of the upcoming stimulus as the source of the effect. Rather, PRFs in somatosensory cortex were shifted toward the neighboring finger, consistent with tuning for combined visual-haptic locations and vice versa in visual cortex. In sum, our results reveal cross-modal effects on population-level spatial tuning in early visual and somatosensory cortices.

RO1EY08266, R01MH111417

Talk 6, 6:30 pm, 55.16

HIGH-LEVEL SENSORY AND MOTOR REGIONS ENCODE OBJECT MASS AFTER REAL-WORLD OBJECT INTERACTIONS

Shubhankar Saha[1] (shubhankars@iisc.ac.in), Prithu Purkait[1], SP Arun[1]; [1]Indian Institute of Science

We experience real-world objects not just by seeing them but by interacting with them. Such interactions give us information about their physical properties such as mass. Are such physical properties integrated into the underlying object representations? To investigate this fundamental question, we performed wireless brain recordings from two monkeys with electrodes implanted into high-level sensory and motor regions before and after they interacted with real-world objects of varying mass. We created 5 water bottles painted with different colors, and added weights (100-500 grams) chosen to be uncorrelated with their (R,G,B) colors. We then recorded neural responses to images of these bottles on a screen while each animal passively viewed these images, prior to any interaction with these bottles. Each bottle was then loaded with a small juice reward and presented to each monkey in randomized order. Monkeys readily interacted with these bottles, lifting them up to drink the juice, thereby ensuring that they had experience with the varying masses of these bottles. Following these interactions, we again recorded neural responses to images of these bottles on a screen as before. We hypothesized that neural activity would show a greater correspondence with the experienced mass of these objects, following the real-world interaction compared to before the interaction. To this end, we calculated the correlation between the multiunit firing rate from each electrode with the object mass. Our main finding is that neural responses showed an increased correlation with object mass after real-world interactions. This effect was present in the premotor/prefrontal cortex (PMv/vlPFC) as well as in inferior temporal cortex (IT). Taken together, our results show that object mass is rapidly encoded into both high-level sensory and motor regions of the brain following real-world interactions with objects.

This research was funded through a Senior Fellowship from the DBT-Wellcome India Alliance (Grant # IA/S/17/1/503081) and the DBT-

IISc partnership programme to SPA, Prime Minister's Research Fellowship to SS and KVPY Fellowship to PP.

TALK 7, 6:45 PM, 55.17

INFORMATIVE VISION ALTERS TACTILE PERCEPTION

Anupama Nair[1], Jared Medina[2]; [1]University of Delaware, [2]Emory University

Previous studies have shown that individuals are more likely to detect a near-threshold tactile stimulus when seeing touch at the same location, leading to the hypothesis that informative vision enhances tactile perception. However, such results could also be explained by a more liberal response criterion when seeing touch. To examine if viewed touch enhances tactile perception, we presented participants with two tasks. Vibrotactile stimuli were presented at varying intensities to both the index and ring finger while participants watched videos of a hand being touched on one cued finger. In the comparative judgment task, participants indicated which finger received the more intense tactile stimulus on their own hand. Across multiple experiments, participants consistently demonstrated a shifted point of subjective equality, reporting that the tactile stimulus associated with the cued finger was more intense. These results provide evidence that the cue clearly altered performance but are agnostic regarding whether there was a shift in response bias or perceptual enhancement. In the equality judgment task, which is more resistant to response bias, participants indicated whether the stimulus intensities on their fingers were the same or different while watching the videos. For equality judgment performance, we found evidence for noise processes at the tails, as participants were more likely to judge tactile stimuli as equal when they were either near threshold or well-above threshold. For stimuli outside of this noise regime, we found a significant shift in the peak of the equality judgment curve (alpha) such that participants were most likely to respond 'equal' when the cued stimulus was less intense than the uncued stimulus. These findings suggest that viewing informative touch enhances tactile perception.

This material is based upon work supported by the National Science Foundation under grant no. 1632849

TALK 8, 7:00 PM, 55.18

COMMON COMPUTATIONS IN AUTOMATIC CUE COMBINATION AND METACOGNITIVE CONFIDENCE REPORTS

Yi Gao[1] (yi.gao0525@outlook.com), Kai Xue[1], Brian Odegaard[2], Dobrimir Rahnev[1]; [1]Georgia Institute of Technology, [2]University of Florida

Sensory stimuli introduce varying degrees of uncertainty, and it is crucial to accurately estimate and utilize this sensory uncertainty for appropriate perceptual decision making. Previous research has examined the estimation of uncertainty in both low-level multisensory cue combination and metacognitive estimation of confidence. However, it remains unclear whether these two forms of uncertainty estimation share the same computations. To address this question, we used a well-established method to induce a dissociation between confidence and accuracy by manipulating energy levels in a random-dot kinematogram. Subjects (N = 99) completed a direction discrimination task for visual stimuli with low vs. high overall motion energy. We found that the high-energy stimuli led to higher confidence but lower accuracy in a visual-only task. Importantly, we also investigated the impact of these visual stimuli on auditory motion perception in a separate task, where the visual stimuli were irrelevant to the auditory task. The results showed that both the high- and low-energy visual stimuli influenced auditory judgments, presumably through automatic low-level mechanisms. Critically, the high-energy visual stimuli had a stronger influence on auditory judgments compared to the low-energy visual stimuli. This effect was in line with the confidence but contrary to the accuracy differences between the high- and low-energy stimuli in the visual-only task. These effects were captured by a simple computational model that assumes that common computations underly confidence reports and multisensory cue combination. Our results reveal a deep link between automatic sensory processing and metacognitive confidence reports, and suggest that vastly different stages of perceptual decision making rely on common computational principles.

We thank Minzhi Wang for his help with data collection. This work was supported by the National Institute of Health (award: R01MH119189) and the Office of Naval Research (award: N00014-20-1-2622).

TALK SESSION: TUESDAY, MAY 21, 2024, 5:15 – 7:15 PM, TALK ROOM 2

Decision Making

Moderator: Constantin Rothkopf, TU Darmstadt

TALK 1, 5:15 PM, 55.21

INVESTIGATING THE ROLE OF LONG-TERM PERCEPTUAL PRIORS IN CONFIDENCE

Marika Constant[1] (marika.constant@gmail.com), Elisa Filevich[2], Pascal Mamassian[3]; [1]Humboldt-Universität zu Berlin, [2]University of Tübingen, [3]École Normale Supérieure, PSL University, CNRS, Paris, France

According to Bayesian models, both our perceptual decisions and our confidence about those decisions are based on the integration of incoming sensory information with our prior expectations. These models therefore assume that priors influence confidence and decisions in the same way, and to the same extent. While asymmetries have been found in the influence priors have on decisions versus confidence, challenging this assumption, those results were obtained with high level cognitive priors that were induced in the task context. It remains unclear whether this generalises to long-term, perceptual priors. Here, we investigated the influence of a low-level prior, namely the slow-motion prior, on confidence. Stimuli were parallel line segments in motion for which the slow-motion prior biases the perceived direction to be perpendicular to the line orientations. Observers had to decide whether the motion direction was clockwise or counterclockwise relative to a reference, and after two such decisions, judge which decision was more likely to be correct. We contrasted two conditions – one where the percept was dominated by the prior, and another where incoming sensory information was dominant. We then assessed which of these conditions participants were more likely to judge as their more confident decision. We found a confidence bias favouring the prior-dominant condition, even when accounting for differences in perceptual decision performance. This suggests that priors impact confidence more strongly than they do perceptual decisions, even in cases of low-level, perceptual priors. Further computational modelling indicates that this effect may be best explained by confidence using the degree of prior-congruent information as an additional cue, above and beyond the posterior evidence used in perceptual decisions. We propose that participants have a metacognitive bias to incorporate confirmatory evidence in favour of their own prior expectations, even when these priors are low-level and participants are unaware of them.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 337619223 / RTG2386, a Freigeist fellowship from the Volkswagen Foundation, number 9D035-1, and an EC grant HORIZON-MSCA-2022-DN-01 “CODE”.

TALK 2, 5:30 PM, 55.22

BAYESIAN INFERENCE BY VISUOMOTOR NEURONS IN PREFRONTAL CORTEX

Robbe Goris[1] (robbe.goris@utexas.edu), Thomas Langlois[1], Julie Charlton[1]; [1]UT Austin

Perceptual interpretations of the environment emerge from the concerted activity of neural populations in decision-making areas downstream of sensory cortex. When the sensory input is ambiguous, perceptual interpretations can be biased by prior beliefs that reflect knowledge of environmental regularities. These effects are examples of Bayesian reasoning, an inference method in which prior knowledge is leveraged to optimize decisions. However, it is not known how decision-making circuits combine sensory signals and prior beliefs to form a perceptual decision. To address this, we studied neural population activity in the prefrontal cortex of two macaque monkeys trained to report perceptual judgments of ambiguous visual stimuli under different prior statistics. Monkeys judged whether a visual stimulus was oriented clockwise or counterclockwise from vertical and communicated their decision with a saccadic eye movement towards one of two visual targets. The meaning of each response option was signaled by the target's orientation (clockwise vs counterclockwise) and was unrelated to its spatial position. Because the spatial configuration of the choice targets varied randomly from trial to trial, changes in prior stimulus statistics biased the animals' perceptual reports, but not the overt motor responses. We analyzed the component of the neural population response that specifically represents the formation of the perceptual decision (the decision variable, DV), and found that its dynamical evolution reflects the integration of sensory signals and prior beliefs. The DV’s initial value before stimulus onset reflects the prior belief in the future state of the sensory environment, while the dynamic range of the DV's ensuing excursion reflects the relative influence of the incoming sensory signals. These findings reveal how prefrontal circuits integrate prior stimulus expectations and incoming sensory signals at the behaviorally relevant timescale of the single trial, thus exposing a general mechanism by which prefrontal circuits can execute Bayesian inference.

TALK 3, 5:45 PM, 55.23

DIFFERENT STIMULUS MANIPULATIONS PRODUCE DISSOCIABLE CONFIDENCE-ACCURACY RELATIONSHIPS

Herrick Fung[1] (herrickfung@gatech.edu), Dobromir Rahnev[1]; [1]Georgia Institute of Technology

A central goal in visual metacognition is to uncover the underlying computations that give rise to our sense of subjective confidence. Achieving this goal necessitates an understanding of how confidence changes in response to various manipulations. However, existing studies have predominately relied on a single stimulus manipulation under the tacit assumption that different manipulations are likely to have equivalent effects on confidence. Here, we test this assumption by including four distinct stimulus manipulations within a single experiment. Subjects judged the orientation (clockwise vs. counterclockwise from 45°) of Gabor patches. The stimuli varied in (1) size (2.5, 5, and 7.5° visual angle), (2) duration (33, 100, and 500 ms), (3) noise contrast (.1, .75, and .9), and (4) orientation (T/2, T, 2T, where T is the individualized threshold obtained by a staircase procedure). We found that the four manipulations produced vastly different effects on accuracy and confidence. Specifically, the size and noise contrast manipulations had a small effect on accuracy but a substantial effect on confidence. Conversely, the orientation manipulation greatly affected accuracy but had only a modest influence on confidence. The orientation manipulation stood out in yet another aspect: it was the only manipulation for which confidence for incorrect trials was higher for the more difficult compared to the easier conditions. The remaining three manipulations exhibited the opposite pattern. We speculate that these effects were driven by orientation being the only manipulation not immediately obvious to the observers. These results clearly demonstrate that different stimulus manipulations yield extensive differences in the confidence-accuracy relationship. Our findings challenge prominent models of confidence that assume a single, stereotypical relationship between confidence and accuracy.

This work was supported by the National Institute of Health (award: R01MH119189) and the Office of Naval Research (award: N00014-20-1-2622).

TALK 4, 6:00 PM, 55.24

DIRECT PRECISION MANIPULATIONS OF MENTAL REPRESENTATIONS IN VISUAL WORKING MEMORY DRIVE SERIAL DEPENDENCE

Sabrina Hansmann-Roth[1]; [1]University of Iceland

Our behavior is heavily influenced by previous information. Work in the field of serial dependence has investigated how the combination of past and present information affects perception and cognition. These studies revealed strong attractive biases towards previously seen stimuli, especially when stimuli are uncertain. Here, for the first time we directly manipulated the uncertainty of mental presentations instead of the uncertainty in the stimulus, through an intermediate task that observers conducted during the retention interval. Participants were presented with differently oriented Gabors and had to reproduce their orientation. While memorizing the orientation, they judged whether two stimuli were identical. This intermediate task varied in the type of stimuli observers were presented with: They compared the size of circles (that contain no orientation information), the length of differently oriented lines or the orientation of lines, and a control condition was added with no intermediate task. These manipulations allowed for a detailed assessment of the role of load and inter-item competition in memory on the precision of the encoded Gabor and subsequently, on the strength of serial dependence. In line with the variable-precision model the mere presence of an intermediate task decreased the precision of the memorized Gabor orientation, enhancing the attractive bias towards past information. Inter-item similarity between the memorized Gabor and the stimuli from the inter-item task further influenced serial dependence: If the intermediate task also required the memorization of orientation, serial dependence was even stronger than for stimuli that contained no orientation information. These results provide novel evidence of the role of working memory on serial dependence. As the precision of individual representations in memory degrades, a greater weight is placed on previous information to make the correct inferences. Moreover, inter-item similarity also leads to a decrease in precision and as a result to an increase in serial dependence.

TALK 5, 6:15 PM, 55.25

ENHANCED METACOGNITION IN INDIVIDUALS WITH AUTISM SPECTRUM DISORDER WHEN INTEGRATING SENSORY EVIDENCE AND REWARDS, BUT NOT PRIOR KNOWLEDGE

Laurina Fazioli[1] (laurina.fazioli@hotmail.fr), Bat-Sheva Hadad[1], Rachel Denison[2], Amit Yashar[1]; [1]University of Haifa, [2]Boston University

Background: Autism Spectrum Disorder (ASD) is a group of neurodevelopmental disorders with complex and diverse impacts on cognition and behavior. Sensory symptoms are increasingly recognized as a core phenotype of ASD, yet the interrelation of these symptoms with cognitive processes remains poorly understood. At the intersection of perception and cognition is the process of perceptual confidence – the ability to evaluate the accuracy of one's own sensory experiences. However, few studies have explored perceptual confidence in ASD. Objective: This study aims to investigate the differences in perceptual metacognitive abilities between individuals with ASD and neurotypical (NT) controls. Using a Bayesian framework, we quantitatively assess how individuals with ASD integrate prior knowledge, sensory evidence, and reward in tasks requiring judgments of perceptual confidence. Method: Two groups of participants, ASD (n = 52) and NT (n = 93), performed an orientation categorization task, designed to evaluate each Bayesian component independently. We manipulated priors, sensory evidence, and reward by varying base rate, stimulus contrast, and a point system, respectively. Participants simultaneously reported the category orientation distribution of a Gabor stimulus and their perceptual decision confidence (four-level rating) by pressing one of eight keys. Results: Individuals with ASD showed enhanced metacognitive accuracy in experiments manipulating sensory evidence and reward, but not in the prior experiment. Furthermore, the type 2 decision criteria (i.e., probability of giving a high-confidence rating) was influenced by the manipulation of prior knowledge to the same extent between the two groups. Conclusions: Our study uncovers an important difference: enhanced metacognitive judgment abilities in individuals with ASD, specifically when integrating sensory evidence and rewards, but not in the context of prior knowledge. This reveals a key difference in how individuals with ASD reflect on and interpret their own perceptual processes.

TALK 6, 6:30 PM, 55.26

METACOGNITIVE MONITORING OF THE VISUAL SYSTEM IN SUSTAINED ATTENTION

Cheongil Kim[1] (kimcheongil@gmail.com), Sang Chul Chong[1]; [1]Yonsei University

The state of the human visual system undergoes moment-to-moment fluctuations due to various neurocognitive factors, such as mind wandering and vigilance. To deal with this instability in the visual system through timely intervention (e.g., controlling attention and taking a rest), monitoring the state of the visual system might be crucial. In this study, we investigated whether and how people can monitor the state of their own visual system during sustained attention tasks. Participants were required to report the orientation (Experiment 1) or presence (Experiment 2) of a Gabor target every two seconds with a confidence judgment for their response. We presumed that if participants could monitor the state of the visual system, confidence judgments would accurately track task performance fluctuations. In Experiment 1, we observed a positive correlation between orientation discrimination performance and confidence, supporting accurate metacognitive monitoring. Experiment 2 aimed to elucidate the mechanism of metacognitive monitoring: direct monitoring of the state of the visual system (e.g., current states of attention and vigilance) versus indirect monitoring based on the visibility of a target. To address this, we employed a target detection task. Specifically, in detection, confidence judgments can be informed by target visibility in judgments about target presence (e.g., high confidence for high visibility) but not target absence. Therefore, if participants monitor only target visibility, their confidence judgments would correlate with performance fluctuations for target-present responses but not for target-absent responses. We observed a positive correlation between detection performance and confidence for target-present responses, but no correlation for target-absent responses. These results suggest that, in sustained attention, metacognitive monitoring of the visual system relies on the visibility of a target, rather than the state of the visual system itself.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A2C3004133).

TALK 7, 6:45 PM, 55.27

PEOPLE TAKE NEWTONIAN PHYSICS INTO ACCOUNT IN SENSORIMOTOR DECISIONS UNDER RISK

Fabian Tatai[1,2] (fabian.tatai@tu-darmstadt.de), Dominik Straub[1,2], Constantin A. Rothkopf[1,2]; [1]Institute of Psychology, Technical University Darmstadt, [2]Centre for Cognitive Science, Technical University Darmstadt

People skillfully manipulate objects on a daily basis, despite uncertainties in both their perceptual inferences and action outcomes. As actions lead to consequences, every movement subject to uncertainty becomes a decision under risk. Such sensorimotor decisions have been shown to follow the predictions of expected utility theory contrary to economic decisions, which systematically fail to maximize expected gains. However, as object manipulations are inescapably governed by the laws of physics, the question arises, how people act under such circumstances. Here, participants slid pucks to targets for gains and losses within a virtual environment, enabling the subjects to interact with an actual standard hockey puck while viewing its trajectory through a head-mounted display. As this novel setup enables subjects to interact with a real-world object through the use of motion capturing, we ensure that subjects have an immersive, naturalistic experience while playing our puck sliding game. In this task, variability inherent in sensorimotor control interacts with the physical relationships governing objects' kinematics under the influence of friction embedded in an economic decision. Therefore, our task features a unique interaction between three cognitive faculties: 1. Economic decision-making, 2. Sensorimotor control and 3. Intuitive physics. We construct an ideal actor model based on statistical decision theory including the kinematics of sliding and show that subject behavior is in coherence with its predictions. Taken together, this demonstrates that subjects use their sensorimotor uncertainty and its interaction with physical relationships and economic demands of the task in guiding their actions.

This research was supported by 'The Adaptive Mind', funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art. We additionally acknowledge support by the European Research Council (ERC; Consolidator Award 'ACTOR'-project number ERC-CoG-101045783).

Talk 8, 7:00 pm, 55.28

UNRAVELING THE INTRICACIES OF HUMAN VISUOSPATIAL PROBLEM-SOLVING

Markus D. Solbach[1] (solbach@yorku.ca), John K. Tsotsos[1]; [1]York University

Computational learning of visual systems has seen remarkable success, especially during the last decade. A large part of it can be attributed to the availability of large data sets tailored to specific domains. Most training is performed over unordered and assumed independent data samples and more data correlates with better performance. This work considers what we observe from humans as our sample. In hundreds of trials with human subjects, we found that samples are not independent, and ordered sequences are our observation of internal visual functions. We investigate human visuospatial capabilities through a real-world experimental paradigm. Previous literature posits that comparison represents the most rudimentary form of psychophysical tasks. As an exploration into dynamic visual behaviours, we employ the same-different task in 3D: are two physical 3D objects visually identically? Human subjects are presented with the task while afforded freedom of movement to inspect two real objects within a physical 3D space. The experimental protocol is structured to ensure that all eye and head movements are oriented toward the visual task. We show that no training was needed to achieve good accuracy, and we demonstrate that efficiency improves with practice on various levels, contrasting with modern computational learning. Extensive use is made of eye and head movements to acquire visual information from appropriate viewpoints in a purposive manner. Furthermore, we exhibit that fixations and corresponding head movements are well-orchestrated, encompassing visual functions, which are composed dynamically and tailored to task instances. We present a set of triggers that we observed to activate those functions. Furthering the understanding of this intricate interplay plays an essential role in developing human-like computational learning systems. The "why" behind all the functionalities - unravelling their purpose - poses an exciting challenge. While human vision may appear effortless, the intricacy of visuospatial functions is staggering.

This research was supported by grants to the senior author (John K. Tsotsos) from the following sources: Air Force Office of Scientific Research USA, The Canada Research Chairs Program, and the NSERC Canadian Robotics Network.

Talk Session: Wednesday, May 22, 2024, 8:15 – 10:00 am, Talk Room 1

Action

Moderator: Douglas Crawford, York University

Talk 1, 8:15 am, 61.11

PREDICTING ACTION TYPE FROM VISUAL PERCEPTION: A KINEMATIC STUDY.

Annalisa Bosco[1,2] (annalisa.bosco2@unibo.it), Elena Aggius Vella[1], Patrizia Fattori[1,2]; [1]Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy, [2]Alma Mater Research Institute for Human-Centered Artificial Intelligence (Alma Human AI)

The planning of a movement toward an object influences the visual perception of the object properties relevant for the action. This suggests a bidirectional interaction between the motor and the visual systems. In the present study, we investigate whether this interaction can be decoded even during the visual estimation of the object properties before the onset of the movement. To this aim, we tested 15 healthy right-handed participants (males=5, females=10; mean age=21.12) in a task consisting of two subsequent phases: 1) a perceptual phase, in which the participants manually estimated the size and orientation of a visual stimulus by extending the index and thumb and, simultaneously, rotating the grip and 2) an action phase, in which participants performed a grasping or a reaching movement (according to the instruction given at the trial onset) towards the same stimulus. A motion capture system recorded the participant's hand position and movement. In order to test if the action type can be predicted during the estimation phase, i.e. if the type of action requested influences the object estimation, we applied a Random Forest classification model to the perceptual phase. The size and orientation estimations, and the velocity of index and thumb (calculated during the perceptual phase) were used as predictors. We found that the model accuracy in classifying the reaching and grasping was on average 99% for the testing dataset. The corresponding sensitivity (ability in classifying true positives) and specificity (ability in classifying true negatives) of the model were 99,5% and 100%, respectively. The most informative predictor was the orientation estimation that contributed for the 99,94%, followed by the size estimation: 78.02% and the index and thumb velocities: 1.2% and 0.6%, respectively. These results suggest that action-based perceptual information can be optimally used to extract action intentions well before the onset of the movement.

MAIA project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 951910; work supported by Ministry of University and Research, PRIN2020-20208RB4N9 and by National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006).

TALK 2, 8:30 AM, 61.12

NEURAL REPRESENTATIONS OF VISUAL MOTION FOR PERCEPTION AND INTERCEPTION

Deborah A. Barany[1] (dbarany@uga.edu), Casey Delaney[1], Haleh Mahmoudi[1], Michelle Marneweck[2]; [1]University of Georgia, [2]University of Oregon

Eye movements are critical for guiding interactions with moving objects according to a behavioral goal, such as tracking an object to perceive its speed or to intercept it with the hand. There are overlapping brain areas involved in motion perception, eye movements, and visually-guided reaching, yet little is known about how these brain areas govern eye-hand interactions with moving objects in different behavioral contexts. Here, we used functional magnetic resonance imaging (fMRI) to investigate how the task goal (perceive/act) and eye movement (fixation/pursuit) modulate neural representations of visual motion. Participants (N = 20) either passively observed (View) or actively intercepted (Intercept-Go) a target moving at a constant rightward or leftward velocity toward an interception zone while their right-hand position and force were recorded on an MR-compatible tablet. On some interception trials, the target changed color prior to movement initiation, indicating that participants should inhibit their planned interception (Intercept-NoGo). Across trials, participants fixated their eyes on the interception zone (Fixate) or smoothly pursued the moving target (Pursue). In-scanner tablet recordings of hand movements and posthoc decoding of eye movements from the MR signal confirmed adherence to the task conditions. Bayesian variational representational similarity analyses of the fMRI data showed that during the initial target motion phase, neural activity patterns in primary visual cortex and human middle temporal area were specific to the eye movement (Fixate vs. Pursue) and motion direction (Right vs. Left), whereas patterns in motor, premotor, and parietal regions were most sensitive to the task goal (View vs. Intercept-NoGo). Analysis of the execution phase of the Intercept-Go trials showed neural activity patterns in primary visual and motor cortices were strongly sensitive to the direction of the target and hand movement. Together, these results reveal distinct eye- and goal-dependent representations for processing visual motion along the sensorimotor hierarchy.

University of Georgia Mary Frances Early College of Education and University of Georgia Office of Research

TALK 3, 8:45 AM, 61.13

ACTIVITY IN PRIMATE VISUAL AREAS IS MODULATED DURING RUNNING

Declan Rowley[1], Jake Yates[2], Alex Huk[1]; [1]UCLA, [2]UCB

Introduction: Running roughly doubles activity in mouse V1, acting as a ~2x gain change (Niell & Stryker, 2010). We recently tested whether these profound modulations of V1 activity are also present in primate, by recording responses in foveal V1 while marmosets running on a treadmill viewed visual stimuli. We found only modest modulations, with hints of suppression. However, whether running affects peripheral representations and/or later visual areas remains unknown. Methods: Following our recent work (Liska, Rowley, et al., 2023), we presented drifting gratings of various orientations to three marmosets while they were perched on a treadmill. Using Neuropixels probes, we recorded from the foveal and peripheral representations of V1, as well as from V2 and MT. Results: We tested whether baseline activity and visually-driven responses were different during running versus not running. In foveal V1, we replicated our finding of little-to-no running effect. However, in the peripheral representation of V1, activity was higher during running in both stimulus viewing and blank periods. Running:Stationary firing rate ratio was 1.196 during stimulus viewing ([1.126, 1.265], 95% CI, p=1.8e-6, 63 cells) and 1.399 during blanks ([1.312, 1.477], p=2e-10). V2 activity was even more strongly modulated during stimulus viewing (1.252, [1.173, 1.351], p = 0.00016 , 26 cells) and during blanks (1.676, [1.416, 1.757], p = 0.00014). Saccade frequency and amplitudes did not differ strongly between conditions, arguing against these modulations arising from different patterns of eye movements when running. In MT, responses were not significantly affected by running (mean firing rate ratio of 1.032 [0.960, 1.109], 77 cells). Conclusion: Although primate foveal V1 is not much affected by running, peripheral V1 and V2 show clear running-correlated modulations. This forms a connection with the striking results found in mice, and calls for a comprehensive dissection of potential modulatory sources and effects.

TALK 5, 9:00 AM, 61.14

DYNAMIC SEQUENTIAL INTERACTIONS OF SPATIAL UNCERTAINTIES EXPLAIN HUMAN NAVIGATION STRATEGIES, ERRORS, AND VARIABILITY

Fabian Kessler[1] (fabian.kessler@tu-darmstadt.de), Julia Frankenstein[1], Constantin Rothkopf[1]; [1]Centre for Cognitive Science TU Darmstadt

Human spatial navigation involves integrating visual cues about our motion and position relative to landmarks with internal signals from self-motion to form a sense of location and direction. However, navigating in the dark or trying to return to a starting point in an environment reveals the uncertainty of these multisensory inferences. Previous studies have revealed many navigational behaviors, including beaconing and path integration, and puzzling patterns of errors and variability in navigation. Ideal observer accounts of navigation have found evidence for perceptual cue integration, but some studies have reported single cues often dominating homing responses. However, purely perceptual accounts do not explicitly account for internal representations, motor planning, and the sequentiality of perception and action. Here, we present an ideal actor model of goal-directed navigation in terms of path planning in the framework of optimal control under uncertainty. This model explicitly accounts for state estimation and learning (Where am I? Where is my goal?) and planning and control (Where should I go? How do I get there?) while taking uncertainty in perception, action, and representation into account. Through simulation of five different triangle-completion experiments from three different laboratories with a single set of biologically plausible parameters, we demonstrate that the observed patterns of navigation are caused by the continuous and dynamic interaction of these three uncertainties. Contrary to ideal observer models, which attribute human endpoint variability to perceptual cue combination processes only, our ideal actor model provides a unifying account of a wide range of phenomena while considering variability in perception, action, and internal representations jointly. Importantly, these findings highlight how dynamic interactions of spatial uncertainties profoundly shape goal-directed navigation behavior and how active vision results from shaping uncertainties along the navigation trajectory, impacting cognitive maps, route planning, movement execution, and ultimately observed behavioral variability.

Calculations for this research were conducted on the Lichtenberg high performance computer of the TU Darmstadt. This research was supported by the European Research Council (ERC; Consolidator Award 'ACTOR'-project number ERC-CoG-101045783).

Talk 6, 9:15 am, 61.15

UNCONSTRAINED VISUALLY-GUIDED GRASPING IS NOT PRECISION GRIP

Fabrizio Lepori[1] (fabrizio.lep96@gmail.com), Frieder Hartmann[2], Kira Dehn[2], Manuela Chessa[3], Roland W. Fleming[2,4], Guido Maiello[1]; [1]Department of Psychology, University of Southampton, [2]Department of Experimental Psychology, Justus Liebig University Giessen, [3]Department of Informatics, Bioengineering, Robotics, and Systems Engineering, University of Genoa, [4]Centre for Mind, Brain and Behaviour (CMBB), University of Marburg and Justus Liebig University, Giessen

In everyday life our hands serve as versatile tools, interacting with objects in numerous ways. Yet most research on visually-guided grasping focuses on precision grips, constraining participants to use only their thumb and index finger. We asked whether insights gained from precision grip studies extend to situations in which participants are free to grasp objects however they chose. To test this, we used a subset of 3D objects from a recent study investigating how participants select precision grips on multi-material (brass and wood) objects [Klein, Maiello et al 2020]. Twenty participants grasped these objects while we tracked their hand movements using a Qualisys passive marker motion capture system. In a first, unconstrained grasping session, participants were free to grasp the stimulus objects however they wanted. In a second, precision grip session, participants were required to grasp the objects using only thumb and index finger. We find that in unconstrained sessions participants rarely employed two-digit precision grips, which accounted for only 9.5% of unconstrained trials (p<.001), and the average position of the digits on the objects differed significantly between precision and unconstrained sessions (p=.019). Nevertheless, in both precision grip and unconstrained sessions participants shifted their grasps towards the objects' center of mass to minimize grip torque (p=.023). Our data thus confirmed the influence of object visual material appearance—previously observed in precision grip experiments—and extended this result to unconstrained grasping. Additionally, upon closer inspection we found that the position of the thumb and index finger on the stimulus objects did not significantly differ between precision and unconstrained sessions (p=.218), suggesting that the remaining fingers primarily provided a support function. Thus, while participants may rarely spontaneously choose two-digit grasps, previous insights gained from precision gip experiments may still extend to natural, unconstrained grasping behaviours.

Talk 6, 9:30 am, 61.16

MODELING THE INFORMATION-BASED CONTROL OF STEERING THROUGH MULTIPLE WAYPOINTS

Brett Fajen[1] (fajenb@rpi.edu), AJ Jansen[1]; [1]Rensselaer Polytechnic Institute

Modeling the visual control of steering has been an active area of research for decades, but the majority of work up to this point has focused on steering to a single target or along a winding road. There is much less work on the strategies used to steer through multiple waypoints, which is relevant in locomotor tasks such as slalom skiing. A critical open issue for modeling the multiple-waypoint task is how to capture the influence of information from waypoints that lie beyond the most immediate one. Recently, we found that humans do use such information, often altering their approach to the nearest waypoint, affording a smoother trajectory through the subsequent waypoint. The aim of the present study was to develop and test competing models that capture human steering behavior observed in multiple-waypoint tasks. We consider four models: (1) the behavioral dynamics model (Fajen & Warren, 2003) with a single goal (most immediate waypoint), (2) the behavioral dynamics model with two goals (two upcoming waypoints), (3) a pure-pursuit controller with a single goal (fixated waypoint) (Tuhkanen et al., 2023), and (4) a new model that relies on information about the constant-radius path that passes through the two upcoming waypoints. We simulated all four models and compared the model-generated trajectories to those produced by human subjects in a task that involves steering through multiple waypoints. Only Model 4 captures the shape of the human trajectories, initially veering away from the nearest waypoint before turning back, setting up a smoother trajectory through both waypoints. The other three models either do not anticipate (Model 1) or were influenced by the future waypoint but not in the way that was consistent with human behavior (Models 2 and 3). The findings demonstrate that human-like anticipation of multiple waypoints can be captured within an information-based framework.

NSF 2218220

Talk 7, 9:45 am, 61.17

MODULARITY OF BRAIN NETWORKS FOR EGOCENTRIC AND ALLOCENTRIC MEMORY-GUIDED REACHING

Lina Musa[1,3] (lmusa09@my.yorku.ca), Amirhossein Ghaderi[1], Ying Chen[6], J. Douglas Crawford[1-5]; [1]Centre for Vision Research, York University, Toronto, ON, Canada, [2]Vision Science to Applications (VISTA), York University, Toronto, ON, Canada, [3]Department of

Psychology York University, Toronto, ON, Canada, [4]Department of Biology York University, Toronto, ON, Canada, [5]Department of Kinesiology York University, Toronto, ON, Canada, [6]Centre for Neuroscience Studies, Queen's University

The brain encodes targets for reaching in egocentric (EGO) and/or allocentric (ALLO) reference frames (Byrne and Crawford 2010). Differences in the cortical activation, but not functional organization, of these two representations have been described (Chen et al., 2014; Neggers et al., 2006). Based on previous findings, we expected increased integration & hubness in the ventral visual stream in ALLO brain networks. Here, we performed a secondary analysis of an event-related fMRI task (Chen et al., 2014). The paradigm consisted of 3 tasks with identical stimulus display but different instructions: remember absolute target location (EGO), remember target location relative to a visual landmark (ALLO), and a nonspatial control, color report. We performed a graph theoretical analysis (GTA) on contrast reduced, time-series data during the memory delay period. GTA measures, including the hubness, clustering coefficient, and efficiency were found, as well as the organization of the network into modules. Dynamical measures of network connectivity (synchrony and complexity) were quantified for individual task modules. EGO and ALLO brain networks showed increased functional segregation & integration, relative to control. Contrary to expectations, there were no inferotemporal modules in both tasks, rather the network was largely segregated into occipito-dorsal-parietal (ODP) and & temporo-frontal (TF) networks modules. The ALLO network demonstrated significantly higher modularity and hubs in the ODP module, relative to the EGO network. When the subtracting the common baseline correlation, the EGO showed segregation of occipital brain areas from the OPD module, but ALLO did not. Our results demonstrate that rather than increased ALLO encoding of visual reach targets in the ventral stream, there is increased specialization in the interaction between early visual brain areas and dorsal parietal brain areas. There was also increase in desynchronization & complexity in the OPD module, in the ALLO task, indicating an increase in difficulty of information processing.

Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council (NSERC), Vision: Science to Applications (VISTA) program.

Talk Session: Wednesday, May 22, 2024, 8:15 – 10:00 am, Talk Room 2

Visual Memory

Moderator: Wilma Bainbridge, University of Chicago

Talk 1, 8:15 am, 61.21

THE EFFECTS OF VISUAL ENCODING SPEED ON ERP MARKERS OF SUBSEQUENT RETRIEVAL

Igor Utochkin[1] (iutochkin@uchicago.edu), Chong Zhao[1], Edward Vogel[1]; [1]University of Chicago

Our memory for meaningful visual stimuli is remarkable: Even when we see thousands of images, each presented for a few seconds, we can later recognize them among new images with high accuracy and in detail (Standing et al., 1973; Brady et al., 2008). However, recognition suffers if the images are encoded at a speed of 2 images per second or faster (Intraub, 1980; Potter, 1976; Potter et al., 2002). Presumably, this happens because the encoding of each new rapidly presented image disrupts relatively slow short-term memory consolidation which is essential for the instantiation of subsequent long-lasting episodic memory. Here, we studied how encoding speed impacts EEG markers of subsequent recognition, namely, ERP Old/New effects, that is, differences between ERP responses to earlier presented (old) and never presented (new) stimuli. In each block, participants memorized sequences of 20 real-world object images at a slow or fast rate (one image each 1,750 ms or 250 ms, respectively). Their memory was then tested with an "old/new" recognition task combined with EEG recording. Our analysis focused on two ERP Old/New components typically distinguished in the literature (Curran, 2000; Paller et al., 2007; Rugg & Curran, 2007): earlier frontal, FN400, and later parietal, LPC. Although observers showed significantly worse recognition at the fast compared to the slow encoding condition, we found almost equally pronounced FN400 in both conditions. In contrast, the LPC was much larger in amplitude in the slow than in the fast encoding blocks. One interpretation of this dissociation can be that fast encoding speed selectively impairs recollection-based memory (which reflects in reduced LPE) but not familiarity-based memory (little effect on FN400). However, other interpretations (e.g., that slower encoding speed causes a stronger confidence signal reflected in LPE) can also be discussed.

ONR N00014-22-1-2123

Talk 2, 8:30 am, 61.22

COLOR-SHAPE CONCEPTS AND THEIR REPRESENTATION IN MACAQUE MONKEY

Spencer R. Loggia[1,2] (spencer.loggia@nih.gov), Helen E. Feibes[1], Karthik Kasi[1], Noah Lasky-Nielson[1], Bevil R. Conway[1]; [1]Laboratory of Sensorimotor Research, National Eye Institute, [2]Department of Neuroscience, Brown University

Object concepts are important tools of cognition that often reflect the interaction of a color and a shape. So, "banana" is a yellow crescent. The brain areas that store color-shape interactions are poorly understood. Testing various hypotheses has been challenging because concepts differ between people, and the corresponding likelihood functions and priors about object shapes and colors are not precisely known. Moreover, functional brain patterns differ among individuals. To overcome these challenges, we raised two macaque monkeys to learn about the colors and shapes of a set of 14 objects. Shape was learned faster than color, as in humans. After the monkeys spent four years interacting with the objects, we scanned their brains while they held in mind the color or shape of the objects. We developed a search-light analysis inspired by convolutional networks that is more robust against noise and better generalizes to cross-cue decoding settings. Cross-cue decoding was significant throughout the cortical visual pathway, implying that color-shape concepts are stored in a distributed network. Overall, cross-cue decoding was best in the posterior parcel of inferior-temporal cortex (PIT) (Acc.=.36 +/-.04, chance=.17). Relative to within-cue decoding, cross decoding increased progressively from posterior to anterior inferior-temporal cortex (AIT) and rhinal cortex (r=.86, p=1.2e-16), suggesting the culmination of the ventral visual pathway in AIT/rhinal cortex is a key locus for generating color-shape concepts. Within PIT, color-decoded-from-shape was relatively greater than shape-decoded-from-color, while the opposite pattern was observed within AIT and rhinal cortex. These asymmetries suggest that PIT represents perceptual memory colors, while AIT and rhinal cortex, and their reciprocally connected targets, compute an abstract concept of colors associated with shapes that could be used to guide visual search.

TALK 3, 8:45 AM, 61.23

NEURAL CORRELATES OF BOUNDARY EXTENSION DURING VISUAL IMAGINATION

Timothy Vickery[1] (tvickery@udel.edu), Banjit Singh[1], Alyssa Levy[1], Kallie Sweetman[1], Zoe Cronin[1], Helene Intraub[1]; [1]University of Delaware

Participants typically remember seeing a greater expanse of a scene than was visible in a studied close-up (boundary extension, BE). Multivoxel pattern analysis (MVPA) was used to test the neural correlates of BE. For each participant, a classifier was trained using a whole-brain searchlight method to discriminate between close-up and wider-angle versions of 16 scenes during repeated perceptual exposures. Earlier in the experiment, each participant had studied either the close or the wide version of each scene and then, on cue, visually imagined it from memory. If a brain area reflects BE, then unlike classification during perception, visual images of close views should now be misclassified as wide (capturing false memory beyond the view) whereas visual images of wide views should be correctly classified. The classifier indeed revealed BE-consistent patterns during imagery in several high-level visual regions, especially in the posterior parietal cortex (cluster TFCE corrected for multiple comparisons). Importantly, this BE-consistent pattern did not reflect a brain-wide bias toward better classification of wider-angle views: (1) the pattern was constrained to visually-responsive regions (occipital, parietal, and inferior temporal); and (2) the pattern reversed (better classification of close views) in early visual cortex, suggesting a bias toward the object in these regions. Following the visual imagery task, participants were again shown their originally presented views and rated each one as closer or farther away than before (4-pt scale); rating analysis revealed the typically pattern: robust BE for close views and no directional error for the wide views, thus verifying BE with a common behavioral measure. We propose that our new method reflects active maintenance of boundary-extended scene representations in memory, and that it holds promise, not only for further exploration of BE, but as a general purpose tool for decoding false memory in the brain.

This work is supported by NIH COBRE 5P30GM145765-02 .

TALK 4, 9:00 AM, 61.24

INDIVIDUAL DIFFERENCES IN VISUAL MENTAL IMAGERY ASSESSED THROUGH STANDARD AND EVOLUTIONARY CLASSIFICATION IMAGES

Guido Maiello[1] (guido_maiello@yahoo.it), Ahmadreza Tajari[2], Katherine R. Storrs[3], Reuben Rideaux[4], Thomas S. A. Wallis[5,6], William J. Harrison[7], Roland W. Fleming[6,8]; [1]Justus Liebig University Giessen, [2]Sharif University of Technology, [3]University of Auckland, [4]The University of Sydney, [5]Technical University of Darmstadt, [6]Centre for Mind, Brain and Behaviour (CMBB), Universities of Marburg, Giessen and Darmstadt, [7]University of Queensland, [8]Justus Liebig University Giessen

Mental imagery—the ability to visualize images in the mind's eye—is associated with many perceptual and cognitive facilities that vary across individuals. Quantifying mental imagery abilities however is challenging and typically relies on subjective and self-report methods, as the world of the imagination is not directly measurable. In contrast to such approaches, here, we propose a method for directly visualising mental images using classification images. Despite their potential, classification images have not been adopted for evaluating individual differences in mental imagery ability due to two primary challenges: the time-consuming nature of traditional reverse correlation that requires many hours of testing, and the uncertainty about how to interpret the reconstructed images. To address these challenges, we first optimized a traditional reverse correlation paradigm to yield recognizable classification images in under 20 minutes, and then developed an additional "evolutionary" paradigm—based on genetic search. We used these methods in an experiment with 20 typical participants who underwent multiple sessions of "standard" and "evolutionary" reverse correlation tasks in which they detected the letter "S" in pure pixel noise images. We fed the generated classification images into deep neural network image classifiers trained at recognizing handwritten letters in noise. We took the networks' cross-entropy loss as a measure of the quality of the generated classification images, and thus of the mental imagery abilities of each participant. This approach exhibited substantial test-retest reliability within both standard ($r=.58$, $p<.01$) and evolutionary ($r=.42$, $p<.05$) reverse correlation sessions, as well as across paradigms ($r=.55$, $p<.01$), and the reliability of the estimated individual differences improved linearly with increasing number of trials ($r=.73$, $p<.001$). These results indicate that both "standard" and "evolutionary" reverse correlation methods consistently measured individual differences in mental imagery. This work thus paves the way for a more nuanced and objective understanding of this complex cognitive function.

TALK 5, 9:15 AM, 61.25

BEYOND 'GIST': THE DYNAMIC INTERPLAY OF CONCEPTUAL INFORMATION AND VISUAL DETAIL IN LONG-TERM MEMORY

Nurit Gronau[1] (nuritgro@openu.ac.il), Roy Shoval[1], Rotem Avital-Cohen[1]; [1]The Open University of Israel

Conceptual information plays an important role in visual LTM, however, the precise nature of such semantic-visual interactions is yet unclear. Here, we tested the effects of object meaning on memory for an arbitrary visual property, specifically, item location. Unlike studies examining object-location binding in visual WM, LTM's longer timescale might involve unique processes that are above and beyond those involved in VWM. According to 'Resource-limited' accounts, highly familiar items demand fewer encoding resources, allowing spare capacity for visual detail encoding (Popov & Reder, 2020). 'Schema-based' accounts, in contrast, suggest that conceptual knowledge may prioritize gist-based representations, at the expense of a visual representation (e.g., Bellana et al., 2021; Koutstaal et al., 2003). Namely, semantic information may hinder item-specific memory, particularly over long time lags. To test these opposing theories, participants encoded individual objects at arbitrary screen locations and were subsequently tested on their memory for these locations using a 4-AFC recognition test, encompassing both old/new items and old/new locations. As expected, overall memory was higher for meaningful (real-world) than for meaningless (scrambled) objects. Critically, given correct item identification, the relative correct location memory rates were significantly higher for the meaningful objects. A follow-up study employed only real-world objects that were independently rated for their 'meaningfulness' and 'visual complexity' factors. Once again, object meaning was positively associated with location accuracy, providing a more fine-tuned measure of conceptual influence on visual memory. Finally, using objects with color-meaningful (e.g., red wine) versus color-meaningless (red balloon) features, we found that in contrast to feature-independent theories (Utochkin & Brady, 2020), location memory was more heavily reliant on color memory when the latter was meaningful. Collectively, our findings align with resource-limited theories, suggesting that meaningful stimuli or features allow an enhanced LTM for arbitrary visual details. Follow-up studies will test semantic-visual memory dynamics over longer time-lags.

Talk 6, 9:30 am, 61.26

THE VISUAL MEMORABILITY OF NATURAL WARNING PATTERNS: INSIGHTS FROM HUMANS AND MACHINES

Federico De Filippi[1] (fdf1@st-andrews.ac.uk), Olivier Penacchio[1,2], Akira R. O'Connor[1], Julie M. Harris[1]; [1]University of St Andrews, St Andrews, United Kingdom, [2]Computer Vision Center, Universitat Autònoma de Barcelona, Barcelona, Spain

While some animals camouflage themselves, others advertise that they are toxic using bright colours and salient stripes and/or spots ('warning patterns'). Their striking appearance is thought to warn off predators: a memorable pattern may help predators learn about toxicity and discourage future attacks on similar prey. However, how warning patterns influence visual memory has never been documented. Research suggests that when glancing at a picture, people do not intuitively know what makes it memorable or forgettable, but they remember and forget the same images (i.e., there is high inter-subject consistency). This means that the likelihood of remembering a picture (its 'memorability') can be computationally predicted from the visual information contained in the picture. Memorable images also lead to stronger neural firing when processed by real and artificial visual systems. We used a database of Lepidoptera (butterfly/moth) images, some of which carry warning signals (aposematic: AP), and some which do not (non aposematic: nAP). We measured human memorability for both AP and nAP Lepidoptera and examined the sources of memorability variation across subjects. Observers studied images while providing subjective ratings (1-10) of memorability, followed by a recognition test ('Seen before?'). Memorability was computed as the proportion of subjects who remembered previously seeing each image. AP species appeared subjectively more memorable than nAP ones, but, on average, they were not better remembered. Remarkably, AP species led to high inter-subject consistency in memorability (Spearman's rho = .79), but consistency for nAP species was comparatively low (Spearman's rho = .37). When we exposed our Lepidoptera patterns to deep neural networks trained for object classification, we found that AP species that were memorable to humans also evoked stronger neural responses in some hidden layers. Taken together, these findings suggest that warning patterns might exploit shared visual mechanisms that underlie successes and failures in picture recognition.

This work is funded by the Biotechnology and Biological Sciences Research Council (United Kingdom Research and Innovation)

Talk 7, 9:45 am, 61.27

SEMANTIC AND VISUAL FEATURES DRIVE THE INTRINSIC MEMORABILITY OF CO-SPEECH GESTURES

Xiaohan (Hannah) Guo[1] (hannahguo@uchicago.edu), Susan Goldin-Meadow[1], Wilma A. Bainbridge[1]; [1]The University of Chicago

Co-speech gestures that teachers spontaneously produce during explanations have been shown to benefit students' learning. Further, prior work suggests that information conveyed through teachers' gestures is less likely to deteriorate than through speech (Church et al., 2007). However, how intrinsic features of gestures affect students' memory remains unclear. The memorability effect denotes a phenomenon where adults with different backgrounds consistently remember and forget particular visual stimuli (static images, dance moves, etc.), owing to the stimuli's intrinsic semantic and visual features. In this study, we investigate whether certain gestures are consistently remembered and, if so, which semantic and visual features are associated with these remembered gestures. We first created 360 10-second audiovisual stimuli by video recording 20 actors producing unscripted natural speech and gestures as they explained Piagetian conservation problems. Two trained experimenters extracted high-level semantics and low-level visual/acoustic features in speech and gesture for each audiovisual stimulus. We then tested online participants' memories in three different conditions using a between-subjects study-test paradigm: the audiovisual stimuli (gesture+speech condition), the visual-only version of the same stimuli (gesture condition), and the audio-only version of the stimuli (speech condition). Within each of the two experimental blocks, participants encoded nine random stimuli from an actor and made old/new judgments on all 18 stimuli from the same actor immediately after. We discovered that participants show significant consistencies in their memory for the gesture, gesture+speech, and speech stimuli. Focusing on the visual-only (gesture) condition, we found that (1) more meaningful gestures and speech predicted more memorable gestures; (2) using both hands led to more memorable gestures than using one hand. Our results suggest that both semantic (conveyed through speech and gestures) and visual (conveyed through gesture) features make co-speech gestures memorable.

Talk Session: Wednesday, May 22, 2024, 11:00 am – 12:45 pm, Talk Room 1

Scene Perception: Neural mechanisms, representations

Moderator: Danny Dilks, Emory University

Talk 1, 11:00 am, 62.11

A NEW SCENE-SELECTIVE REGION IN THE SUPERIOR PARIETAL LOBULE AND ITS POTENTIAL INVOLVEMENT IN VISUALLY-GUIDED NAVIGATION

Hee Kyung Yoon[1] (hee.kyung.yoon@emory.edu), Yaelan Jung[1], Daniel Dilks[1]; [1]Emory University

Growing evidence indicates that the occipital place area (OPA) – a scene-selective region in adult humans – is involved in "visually-guided navigation". Here, we present evidence that there is a new scene-selective region located in the superior parietal lobule – henceforth called the "superior place area" (SPA) – that may also be involved in visually-guided navigation. First, using functional magnetic resonance imaging (fMRI), we found that SPA responds significantly more to scene stimuli than to face and object stimuli across two different sets of stimuli (i.e., "dynamic" and "static") – establishing SPA as yet another scene-selective region. Second, we found that SPA, like OPA, responds significantly more to dynamic scene stimuli (i.e., video clips of first-person perspective motion through scenes, mimicking the actual visual experience of walking through a place) than to static scene stimuli (i.e., static images taken from the same video clips, rearranged such that first-person perspective motion could not be inferred) – suggesting that SPA, like OPA, is involved in visually-guided navigation. Such sensitivity to first-person perspective motion information through scenes cannot be explained by scene selectivity alone, domain-general motion sensitivity, or low-level visual information. And third, resting-state fMRI data revealed that SPA is preferentially connected to OPA, compared to other scene regions – again consistent with the hypothesis that the SPA, like OPA, is involved in visually-guided navigation. Taken together, these results demonstrate a new scene-selective region that may be involved in visually-guided navigation, and raise interesting questions about the precise role that SPA (compared to OPA) plays in visually-guided navigation.

Talk 2, 11:15 am, 62.12

DYNAMIC FUNCTIONAL CONNECTIVITY VIA IEEG - FMRI CORRELATION MAPS

Zeeshan Qadir[1] (qadir.zeeshan@mayo.edu), Harvey Huang[1], Morgan Montoya[1], Michael Jensen[1], Gabriela Ojeda Valencia[1], Kai Miller[1], Gregory Worrell[1], Thomas Naselaris[2], Kendrick Kay[3], Dora Hermes[1]; [1]Mayo Clinic, [2]University of Minnesota

Understanding neural computations of vision require studying how different brain regions interact with one another. However, functional connectivity across brain regions is often computed as stationary maps, concealing the rich neural dynamics that change at a finer timescale. To better understand how functional connectivity evolves over time, we propose a multimodal framework combining data from intracranial-EEG (iEEG) and fMRI. We recorded iEEG data from early visual (V1/V2) electrodes in 4 patients. Each patient was shown a subset of 1000 stimuli from the NSD-fMRI dataset. Electrodes with significant broadband (70-170 Hz) power increases w.r.t the baseline were considered for further analysis. From the NSD-fMRI dataset, we obtained average fMRI beta-weights for the 1000 stimuli that were repeated thrice across the 8 subjects. Next, for each iEEG electrode we computed a Pearson correlation map with all the fMRI vertices, across the 1000 stimuli, giving us a time x vertices correlation matrix. This provided us with a brain-wide temporally evolving correlation map for each electrode. In all 4 subjects, we observed that the iEEG broadband significantly correlates with the fMRI beta-weights in V1, and with V2/V3 about 5-10 ms later, followed by the ventral temporal regions around 170 ms. Other parietal and frontal brain regions also showed significant correlations after 100 ms. Further, we also observed that these correlations reduce around 450 ms, even though the stimuli were presented for 800 ms. These temporally resolved correlation maps show that V1 representations are not stationary but share representations with higher order visual areas over time. These results may suggest that connectivity to V1 evolves over time revealing feedback inputs from higher order ventral areas around 100-170 ms. Overall, we propose that our multimodal framework enables us to compute functional connectivity at high spatiotemporal resolution reflecting the rich dynamics of interaction across different brain regions.

We thank the patients in this study for their participation, Cindy Nelson and Karla Crockett for their assistance, and Peter Brunner for support with BCI2000. Research reported in this publication was supported by the NEI (R01EY035533, R01EY023384)

Talk 3, 11:30 am, 62.13

TOP-DOWN ALPHA DYNAMICS MEDIATE THE NEURAL REPRESENTATION OF COHERENT VISUAL EXPERIENCES

Daniel Kaiser[1,2] (danielkaiser.net@gmail.com), Lixiang Chen[1,3], Radoslaw M Cichy[3]; [1]Mathematical Institute, Justus Liebig University Giessen, [2]Center for Mind, Brain and Behavior, Philipps University Marburg and Justus Liebig University Giessen, [3]Department of Education and Psychology, Freie Universität Berlin

In order to create coherent visual experiences, our visual system needs to aggregate inputs across space and time in a seamless manner. Here, we combine spectrally resolved EEG recordings and spatially resolved fMRI recordings to characterize the neural dynamics that mediate the integration of multiple spatiotemporally coherent inputs into a unified percept. To unveil integration-related brain dynamics, we experimentally manipulated the spatiotemporal coherence of two naturalistic videos presented in the left and right visual hemifields. In a first study, we show that only when spatiotemporally consistent information across both hemifields affords integration, EEG alpha dynamics carry stimulus-specific information. Combining the EEG data with regional mappings obtained from fMRI, we further show that these alpha dynamics can be localized to early visual cortex, indicating that integration-related alpha dynamics traverse the hierarchy in the top-down direction, all the way to the earliest stages of cortical vision. In a second study, we delineate boundary conditions for triggering integration-related alpha dynamics. Such alpha dynamics are observed when videos are coherent in their basic-level category and share critical features, but not when they are coherent in their superordinate category, thus characterizing the range of flexibility in cortical integration processes. Together, our results indicate that the construction of coherent visual experiences is not implemented within the visual bottom-up processing cascade. Our findings rather stress that integration relies on cortical feedback rhythms that fully traverse the visual hierarchy.

This work is supported by the DFG (CI241/1-1, CI241/3-1, CI241/7-1, KA4683/5-1, SFB/TRR 135), the ERC (ERC-2018-STG 803370, ERC-2022-STG 101076057), the China Scholarship Council, and "The Adaptive Mind", funded by the Hessian Ministry of Higher Education, Science, Research and Art.

TALK 4, 11:45 AM, 62.14

NEURAL RESPONSES IN SPACE AND TIME TO A MASSIVE SET OF NATURAL SCENES

Peter Brotherwood[1] (peter.brotherwood@umontreal.ca), Emmanuel Lebeau[1], Mathias Salvas-Hébert[1], Marin Coignard[1], Shahab Bakhtiari[1], Frédéric Gosselin[1], Kendrick Kay[2], Ian Charest[1]; [1]CerebrUM, Université de Montréal, [2]Center for Magnetic Resonance Research, University of Minnesota

Understanding how neurons in the visual system support visual perception requires deep sampling of neural responses across a wide array of visual stimuli. Part of this challenge has been met by a recent large-scale 7T fMRI dataset, termed the Natural Scenes Dataset (NSD). This dataset provides extensive high-resolution spatial sampling of brain activity in eight observers while they view complex natural scenes. Here, we present the NSD-EEG, a large-scale electroencephalography (EEG) dataset that provides detailed characterisation of brain activity from a temporal perspective, thereby completing the characterisation of visual processing in the human brain. For this dataset, we optimised data quality by choosing 8 participants from a larger pool based on empirical signal-to-noise metrics and by using a high-density (164 channels) EEG system within a shielded Faraday cage. NSD images were shown for a duration of 250 ms, followed by a variable interstimulus interval of 750-1000 ms. Each participant viewed 10000 images 10 times, with a subset of 1000 images (common across participants) repeated 30 times. Preliminary analyses reveal remarkably consistent event-related potentials (ERPs) for each stimulus, with high inter-trial reliability even at a rapid one stimulus per second pace (max Pearson R: 0.8, $p<0.001$). Additionally, split-half representational dissimilarity matrices exhibit strong reliability (max Spearman R: 0.4, $p<0.001$), further affirming the robustness of our data. We plan to publicly release the NSD-EEG dataset in the near future, alongside an exhaustive battery of complementary behavioural and psychophysical data. In combination with the NSD dataset, this will enable a comprehensive examination of neural responses in space and time to complex natural scenes. Altogether, this will support the ongoing movement using machine learning, artificial intelligence, and other computational methods to characterise and understand the neural mechanisms of vision.

This work was supported by a UNIQUE postgraduate research grant (to PB), a Courtois Chair in Neuroscience (to IC), and an NSERC discovery grant (to IC).

TALK 5, 12:00 PM, 62.15

LESS IS MORE: AESTHETIC LIKING IS INVERSELY RELATED TO METABOLIC EXPENSE BY THE VISUAL SYSTEM

Yikai Tang[1], Wil Cunningham[1,2], Dirk Bernhardt-Walther[1]; [1]University of Toronto, [2]Vector Institute

What makes us like a particular scene or object and dislike another? A variety of visual properties, the observers' experience, familiarity, processing fluency, and self-relevance have been suggested to underlie aesthetic liking. Here we investigate whether the brain's goal to reduce energy costs (Olshausen and Field 1997; Friston, 2010) explains the construction of aesthetic appreciation. We propose a simple, straightforward approach to explaining neural responses to visual stimuli with different levels of aesthetic preference: the total metabolic cost of firing of neurons within relevant regions of interest. We test this hypothesis in an in-silico model of the visual system (VGG19) as well as human observers and find strong evidence in both. Specifically, we compare the metabolic cost incurred by 4914 images of objects and scenes from the BOLD5000 dataset for a VGG19 network pretrained for object and scene categorization with randomly initialized versions of VGG19. We find a strong inverse relationship between aesthetic preferences for the images and their metabolic cost, but only in the network trained for categorization. We then test the same hypothesis in the human visual system by comparing aesthetic liking of visual stimuli to the metabolic activity measured with functional magnetic resonance imaging. Crucially, we find strong evidence for the hypothesized inverse relationship between metabolic expense and aesthetic liking in both early visual brain regions (V1 and V4) and high-level regions (FFA, OPA, PPA). These findings represent the first direct evidence for a physiological basis of visual aesthetics at the level of energy consumption by the visual system. Aesthetic pleasure may function as an adaptive homeostatic signal to help conserve energy resources for survival. Our metabolic account for aesthetic liking unifies empirical evidence for visual discomfort with theories of processing fluency, image complexity, expertise, and prototypicality for aesthetic liking in a simple, physiologically plausible framework.

This work was supported by an NSERC Discovery Grant (RGPIN-2018-05946) to WC and an NSERC Discovery Grant (RGPIN-2020-04097) as well as a SSHRC Insight Grant (435-2023-0015) to DBW.

TALK 6, 12:15 PM, 62.16

IS VISUAL CORTEX REALLY "LANGUAGE-ALIGNED"? PERSPECTIVES FROM MODEL-TO-BRAIN COMPARISONS IN HUMAN AND MONKEYS ON THE NATURAL SCENES DATASET

Colin Conwell[1] (conwell@g.harvard.edu), Emalie McMahon[1], Kasper Vinken[3], Jacob S. Prince[2], George Alvarez[2], Talia Konkle[2], Leyla Isik[1], Margaret Livingstone[3]; [1]Johns Hopkins University, Department of Cognitive Science, [2]Harvard University, Department of Psychology, [3]Harvard Medical School, Department of Neurobiology

Recent advances in multimodal deep learning and in particular “language-aligned” visual representation learning have re-ignited longstanding debates about the presence and magnitude of language-like semantic structure in the human visual system. A variety of recent works that involve mapping the representations of “language-aligned” vision models (e.g. CLIP) and even pure language models (e.g. GPT, BERT) to activity in the ventral visual stream have made claims that the human visual itself may be “language-aligned” much like recent models. These claims are in part predicated on the surprising finding that pure language models in particular can predict image-evoked activity in the ventral visual stream as well as the best pure vision models (e.g. SimCLR, BarlowTwins). But what would we make of this claim if the same procedures worked in the modeling of visual activity in a species that doesn’t speak language? Here, we deploy controlled comparisons of pure-vision, pure-language, and multimodal vision-language models in prediction of human (N=4) AND rhesus macaque (N=6, 5:IT, 1:V1) ventral stream activity evoked in response to the same set of 1000 captioned natural images (the NSD1000 images). We find (as in humans) that there is effectively no difference in the brain-predictive capacity of pure vision and “language-aligned” vision models in macaque high-level ventral stream (IT). Further, (as in humans) pure language models can predict responses in IT with substantial accuracy, but perform poorly in prediction of early visual cortex (V1). Unlike in humans, however, we find that pure language models perform slightly worse than pure vision models in macaque IT, a gap potentially explained by differences in neural recording alone (fMRI versus electrophysiology). Together, these results suggest that language model predictivity of the ventral stream is not necessarily due to language per se, but rather to the statistical structure of the visual world as reflected in language.

Talk 7, 12:30 pm, 62.17

THE HIGH-DIMENSIONAL STRUCTURE OF NATURAL IMAGE REPRESENTATIONS VARIES SYSTEMATICALLY ACROSS VISUAL CORTEX

Raj Magesh Gauthaman[1] (rgautha1@jh.edu), Brice Ménard[1], Michael Bonner[1]; [1]Johns Hopkins University

The computational goal of the visual cortex is often described as systematic dimensionality reduction, where high-dimensional sensory input is gradually reduced to a low-dimensional manifold over multiple stages of processing. Recently, thanks to the unprecedented size of the Natural Scenes Dataset, we showed that the structure of human visual cortex representation is high-dimensional. We were able to reliably detect visual information encoded over many hundreds of latent dimensions. In an effort to reconcile these divergent theoretical predictions and empirical results, we set out to investigate how natural image representations are transformed along the visual hierarchy from a spectral perspective. Using a robust cross-decomposition approach, we estimated cross-validated covariance spectra of fMRI responses in several regions of interest in the visual cortex. In all of them, we observed power-law covariance spectra over hundreds of dimensions. Interestingly, we also noticed systematic trends: spectra decay more rapidly from earlier to later stages of visual processing. This could be seen from V1 to V4 and also from early- to mid- and late- stages of processing within the ventral, dorsal, and lateral visual streams. High-level functionally localized regions of visual cortex including face-, body-, scene- and object-selective cortex also show covariance spectra decaying more rapidly. Our findings demonstrate that while cortical representations of natural images are consistently high-dimensional across many stages of processing—thus using all available dimensions to encode visual information—there are, nonetheless, systematic regional variations in how information is concentrated along these dimensions. These differences in the representational structure of visual regions may provide insight into computational strategies in the human brain.

Talk Session: Wednesday, May 22, 2024, 11:00 am – 12:45 pm, Talk Room 2

Binocular Vision

Moderator: Jorge Otero-Milan, UC Berkeley

Talk 1, 11:00 am, 62.21

ADDRESSING THE VERGENCE-ACCOMMODATION CONFLICT IN VIRTUAL REALITY: A GEOMETRICAL APPROACH

Xiaoye Michael Wang[1] (michaelwxy.wang@utoronto.ca), Colin Dolynski[1], Michael Nitsche[2], Gabby Resch[3], Ali Mazalek[4], Timothy N Welsh[1]; [1]University of Toronto, [2]Georgia Institute of Technology, [3]Ontario Tech University, [4]Toronto Metropolitan University

Technologies on the mixed reality continuum, such as virtual reality (VR), commonly yield distortions in perceived distance. One source of such distortions is the vergence-accommodation conflict, where the eyes’ accommodative state is coerced to the fixed locations of a headset’s screen, while the angles at which the two eyes converge in virtual space continuously update. The current study conceptualizes the effect of vergence-accommodation conflict as a constant outward offset to the vergence angle of approximately 0.2°. Based on this conceptualization, a novel model was developed to predict and account for the resulting distance distortions in VR using the stereoscopic viewing geometry. Leveraging this model, an inverse transformation algorithm along the observer’s line of sight was applied to the rendered virtual environment to counter the effect of vergence offset. To test the effects of the transformation, participants performed a series of manual pointing movements on a tabletop with or without the inverse transformation algorithm. Results showed that the participants increasingly undershot the targets when the inverse transformation was not available, but were consistently more accurate when the algorithm was applied to the virtual environment. The results indicate that systematically transforming the rendered virtual environment based on perceptual geometry could ameliorate distance distortions arising from the vergence-accommodation conflict. The findings of the present study could be applied to designing VR-based applications, such as for medical/surgery training, to improve the accuracy when interacting with virtual objects.

This work was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC), the Canada Research Chair Program, the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Foundation for Innovation, and the Ontario Ministry for Research and Innovation.

TALK 2, 11:15 AM, 62.22

THE INFLUENCE OF SIMULATED OCULAR COUNTER ROLL ON STEREOACUITY

Stephanie M Reeves (stephanie_reeves@berkeley.edu), Jorge Otero-Millan[1,2]; [1]Herbert Wertheim School of Optometry & Vision Science, University of California Berkeley, [2]Department of Neurology, Johns Hopkins University

Stereopsis relies on precise binocular alignment to compute binocular disparity and infer 3D depth. When humans tilt their heads towards the shoulder, the two eyes rotate around the lines of sight in the opposite direction of head tilt. This ocular counter roll (OCR) only partially compensates for the head tilt. The torsion induced during OCR results in a misalignment of the horizontal meridians of the two eyes, which leads to vertical disparities between the retinas. The current work sought to investigate the effect of retinal image rotation due to OCR on stereoacuity while upright. We hypothesized that these vertical disparities will result in decreased stereoacuity. To investigate this research question, we recruited 8 participants to view stereoscopic random dot ring stimuli (spanning 2° to 3.5° peripherally, duration of 200 ms) with the use of a haploscope. Subjects reported whether a stimulus with crossed and uncrossed disparities of 0.1. 0.3, 0.5, 0.7, and 0.9 arcmins appeared in front or behind a fixation target with zero disparity. The stimulus rings were rotated by ±0°, 5°, 10°, and 30° to simulate OCR. Results revealed that stereoscopic thresholds during the 30° stimulus rotation were significantly worse than the 0° stimulus rotation thresholds ($t(7) = 3.00$, $p=0.02$). The reduction in stereoacuity at the 30° stimulus rotation was not worse than what is predicted by the reduction in horizontal disparity alone ($p=0.31$). Stimulus rotations of 0°, 5°, and 10° were not different from one another ($p>0.66$). Taken together, these results indicate that the limited amount of OCR (typically less than 10° in humans for any head tilt) may be optimized for stereopsis: with more torsion than is natural for the human body, stereoacuity gets worse, while modest amounts of torsion are tolerable for stereopsis.

TALK 3, 11:30 AM, 62.23

REVERSED DEPTH REPRESENTATION IN HUMAN AND ARTIFICIAL VISUAL SYSTEMS

Bayu Gautama Wundari[1], Hiroshi Ban[1,2]; [1]National Institute of Information and Communications Technology (NICT), Japan, [2]Graduate School of Frontier Biosciences, Osaka University, Japan

Stereopsis facilitates the brains of animals with front-facing eyes in linking the left and right retinal light patterns to disentangle the complex depth information in the sensory input. Extracting depth structures has been thought to be sufficiently explained by the binocular energy model in the primary visual cortex. However, engaging in real-world 3D tasks requires more complex stereo computations in higher visual areas. The locations and mechanisms through which the brain transforms the physical stimuli representation (binocular disparity) into perceptual representation (depth perception) remain unclear. To address the issue, we combined human psychophysics, neuroimaging, and deep neural network (DNN) simulations. We designed random-dot stereograms (RDSs) by varying the binocular dot contrast correlation to reverse the physical and the perceived depth: RDSs whose physical disparity elicited near were perceived as far, and vice versa. Participants (N=22) reported perceiving depth in reverse when presented with the engineered RDSs in a two-alternative forced choice (2AFC) task discriminating near and far. Decoding analysis on their fMRI voxel response patterns (V1-3, V3A, V3B, hMT, hV4) revealed that only V3A represented reversed depth, suggesting that reversed depth is an extrastriate phenomenon involving more complex stereo computation beyond early areas. We tested two DNNs to gain insight into the network architecture underlying reversed depth in V3A. One network could learn the geometry of contextual background for regressing disparity from a rectified pair of stereo images. The second network did not include such contextual information in its disparity estimation. Both networks were trained on the SceneFlow datasets, which provide accurate depth maps. We demonstrate that the network incorporating contextual information exhibited similar behavioral performance in the human's depth judgment tasks. We conclude that V3A may house the neuronal circuit that learns the spatial context to generate neuronal activity that gives rise to conscious reversed depth perception.

This research was supported by ERATO (JPMJER1801) and Ministry of Education, Culture, Science, Sports and Technology (21H00968)

TALK 4, 11:45 AM, 62.24

SURVIVING CONTINUOUS FLASH SUPPRESSION: A TWO-PHOTON CALCIUM IMAGING STUDY IN MACAQUE V1

Cai-Xia Chen[1], Dan-Qing Jiang[1], Xin Wang[1], Sheng-Hui Zhang[1], Shi-Ming Tang[1], Cong Yu[1]; [1]Peking University

Continuous flash suppression (CFS) has been widely used to study visual consciousness or awareness. Although the flashing Mondrian noise presented to one eye can suppress the perception of a stimulus presented to the other eye, some low-level visual information can survive the suppression and participate in downstream visual processing subconsciously. However, it remains elusive how the responses of V1 neurons, which receive stimulus inputs from two eyes, are affected by CFS. To address this issue, we used two-photon calcium imaging to record responses of superficial-layer V1 neurons to a target under CFS in two FOVs of an awake, fixating macaque. The target was a circular-windowed square-wave grating (d=1°, SF=3/6 cpd, contrast=0.45, drifting speed=4°/s). The flashing stimulus was a circular Mondrian noise pattern (d=1.89°, contrast=0.50, TF=10 Hz). The stimuli were presented for 1000-ms with 1500-ms intervals. The square grating at various orientations was first presented alone to either eye to identify oriented-tuned V1 neurons (~700 per FOV) and calculate each neuron's ocular dominance index (ODI). Then the grating target was presented to one eye and the flashing noise to the other eye to measure neuronal responses under CFS. With the presence of flashing noise, orientation responses of neurons preferring the noise eye (ODI>0.2), in the form of population orientation tuning function, were completely suppressed (by 96.5%) without measurable bandwidth, and those preferring both eyes (-0.2<ODI<0.2) were also severely suppressed (by 89.5%) with unmeasurable or very wide bandwidth. However, although the responses of neurons preferred the grating eye were also significantly suppressed (by 75.5%), the tuning bandwidth was still measurable, which increased from 11-13° to 19° (half-height half-width). These results indicate that only a small portion of the orientation responses in V1 neurons preferring the target eye can survive continuous flash suppression, while orientation responses of other neurons are mostly wiped out.

supported by the National Science and Technology Innovation 2030 Major Program (2022ZD0204600)

Talk 5, 12:00 pm, 62.25

COMPUTATIONAL MECHANISMS OF PERCEPTUAL TRAVELING WAVES

João Victor XAVIER CARDOSO[1], Hsin-Hung LI[2,3], David J. HEEGER[2,3], Laura DUGUÉ[1,4]; [1]Université Paris Cité, CNRS, Integrative Neuroscience and Cognition Center, F-75006 Paris, France, [2]Department of Psychology, New York University, New York, NY, [3]Center for Neural Science, New York University, New York, NY, [4]Institut Universitaire de France (IUF), Paris, France

Binocular rivalry is a perceptual phenomenon in which perception alternates between rival images presented to each eye. Under the right conditions, the dynamics of these alternations form a wave-like pattern starting where one rival image locally becomes the dominant percept. Studies have shown a link between these perceptual traveling waves and waves of brain activity in primary visual cortex (Lee et al., 2005). Here, we replicate and extend previous psychophysics studies of perceptual waves observed in binocular rivalry (e.g., Wilson et al., 2001), and fit a computational model to the behavioral data. A pair of orthogonal gratings, each windowed by an annulus and projected to one eye, were presented to human participants (n=21). Replicating previous results, a local contrast increment in one eye induced perceptual dominance that emerged locally and progressively expanded as it rendered invisible the stimulus presented to the other eye. Participants pressed a key when a perceptual wave reached a target area enabling us to measure propagation speed. We observed (1) slower speeds for more eccentric annuli, commensurate with differences in cortical magnification; (2) slower speed when crossing the vertical meridian, consistent with inter-hemispheric communication; (3) morning participants perceived faster waves than afternoon participants, interpreted as circadian variations in cortical excitability; (4) allocating attention to the annulus was necessary for perceptual waves to be perceived; and (5) rhythmic, local contrast increments induced rhythmic perceptual waves. Finally, we adapted a previously proposed binocular rivalry model (Li, et al. 2017) so it can reproduce both temporal and spatial patterns of perceptual waves. The model could replicate our main findings, along with features reported by other studies, such as changes in propagation speed as a function of attention, input strength and recurrent excitation. Together, our research aims to develop a computational framework for understanding perceptual traveling waves in binocular rivalry.

Talk 6, 12:15 pm, 62.26

DICHOPTIC CONTRAST INTEGRATION ACROSS THE HUMAN VISUAL CORTEX HIERARCHY USING FUNCTIONAL MRI

Kelly Chang[1] (kchang4@uw.edu), Xiyan Li[1,2], Kimberly Meier[1,3], Kristina Tarczy-Hornoch[1], Geoffrey M. Boynton[1], Ione Fine[1]; [1]University of Washington, [2]University of California, San Diego, [3]University of Houston

Introduction: A recent behavioral study by Meier et al. (2023) showed that when the contrast of a non-rivalrous grating is modulated independently in the two eyes, the perceived contrast of the combined stimulus roughly follows the maximum contrast over the two eyes. Here, in a similar paradigm using fMRI, we investigated the neural locus of this behavioral result. Methods: We measured BOLD fMRI signals in early visual cortex (V1 – V3) while participants (n = 10) viewed non-rivalrous dichoptic gratings (2-cpd) that varied slowly in contrast over time in each eye independently at 1/6 and 1/8 Hz. Observers provided a continuous report of perceived contrast over time by positioning a joystick lever. We fit a Minkowski mean [(L(t)^m + R(t)^m) / 2] ^ (1/m) to the behavioral and fMRI time-courses, where L(t) and R(t) are the contrast time-courses in each eye. An exponent parameter of m = 1 is simple averaging, and as m → ∞ the model increases towards a max response in which neural responses or perceived contrast is driven by the eye presented with the highest contrast. Results: The magnitude of m was smallest in V1 (m = 2.00) and increased across the visual hierarchy toward a max model in V2 (m = 5.19) and V3 (m = 8.12). Behavioral responses measured during scanning were consistent with a max model (m = 6.55) and the later stages of the visual hierarchy. Conclusion: Our fMRI results in V1 are similar to a previous fMRI study that used a normalization model (Moradi & Heeger, 2009) to predict V1 BOLD responses. However, the integration of contrast in V1 differs systematically from perceived contrast. BOLD signals in V2 and V3 were consistent with behavioral measurements, implicating these higher visual areas as the neural locus of perceived contrast.

Knights Templar Eye Foundation, Research to Prevent Blindness, UW Center for Human Neuroscience, Unrestricted grant from Research to Prevent Blindness to UW Department of Ophthalmology

Talk 7, 12:30 pm, 62.27

SELECTIVITY FOR BINOCULAR DISPARITY IN THE PRIMATE SUPERIOR COLLICULUS MAY NOT BE DIRECTLY INHERITED FROM V1

Incheol Kang[1] (incheollkang@gmail.com), Gongchen Yu[1], Leor Katz[1], Richard Krauzlis[1], Hendrikje Nienborg[1]; [1]National Eye Institue, NIH

The primate superior colliculus (SC) gets prominent inputs from V1 where selectivity for horizontal binocular disparity is well-established. Such disparity selective input could provide a direct route for depth information supporting orienting behaviors in 3D environments. Here, we used multichannel linear arrays to record from the superficial and intermediate layers of the SC of one rhesus macaque while presenting random-dot stereograms (RDSs) at the neurons' receptive fields (mean = 13.1°, range = 0.5° ~ 41.1°). We examined disparity tuning for both correlated and anti-correlated RDSs, in which corresponding dots shown to the left and right eye had opposite luminance polarities. Of the 393 isolated units, 272 (69%) were significantly selective for binocular disparity (Disparity Discrimination Index, $p < 0.05$). Units recorded in the same session tended to prefer similar disparities, suggesting clustering for disparity. Disparity tuning properties were comparable between neurons in the superficial (more visual) and intermediate (more visuomotor) layers. Consistent with the idea of pooling inputs from V1, the disparity selectivity emerged quickly after stimulus onset (~43 ms), typically showed even-symmetric tunings (78%) and had a broad tuning width. As in V1, the disparity tuning for anti-correlated RDSs was inverse to that for the correlated RDSs with a reduced amplitude compared to that for correlated RDSs. However, this amplitude reduction was substantially more pronounced in the SC (a median of 18%) compared to V1. Furthermore, the disparity selectivity was negatively correlated with the degree of monocularity ($r = -0.35$, $p < 10-8$), unlike previous findings in V1. Together, we find that most SC neurons are selective for binocular disparity providing a plausible neural substrate for how the SC supports visual orienting in 3D natural environments. Several properties of the disparity tuning appear incompatible with direct pooling of V1 and suggest that it is shaped by additional mechanisms.

Posters Sessions

SATURDAY MORNING POSTERS IN BANYAN BREEZEWAY

SATURDAY, MAY 18, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Scene Perception: Miscellaneous

23.301 BRIDGING PERSPECTIVES: A FOUNDATIONAL DATASET FOR THE EMPIRICAL AESTHETICS OF BRIDGE DESIGN

Mei Yang[1] (mei.yang@mail.utoronto.ca), Claudia Damiano[1], Paul Gauvreau[2], Dirk B. Walther[1]; [1]Department of Psychology, University of Toronto, [2]Department of Civil Engineering, University of Toronto

People readily recognize London when they see Tower Bridge, San Francisco by the Golden Gate Bridge, and Sydney when they see the Sydney Harbour Bridge. These bridges have emerged as iconic landmarks that shape their city's skyline by virtue of their aesthetic qualities. Bridges can also elicit a neutral aesthetic response and, sometimes, can be regarded as downright ugly. Bridges are designed as public infrastructure, which often shape their surroundings for centuries. Nonetheless, little is known about what shapes the aesthetic appeal of bridges. Here we explore how aesthetic judgements of bridges relate to engineering and design features. Our dataset comprises 318 images of 118 bridges from around the world, rated by 200 participants for aesthetic pleasure, interest, complexity, and safety. A civil engineering team annotated each image for type, depth, material, apparent age, and aesthetic premium. Using Factorial Analysis of Mixed Data (FAMD), we found two significant dimensions. The first "aesthetics" dimension shows strong correlations among aesthetic, complexity, and interest ratings and is connected to bridge type. The second "safety" dimension relates subjective ratings of safety to bridge age and material. Analysis of visual features of bridges, using the Mid-Level Vision (MLV) Toolbox, shows that contour length is a predictor of both bridge type and the aesthetic, complexity, and interest ratings. For example, truss bridges, made up of several interconnected beams, are represented by many short contours and are generally rated as more complex, interesting, and pleasing. On the other hand, the visually simple slab and girder bridges are often represented by a few long contours and are rated as uninteresting and not aesthetically pleasing. Our study offers the first attempt to systematically collect and analyze subjective ratings of bridge aesthetics, paving the way for empirically supported decisions for the design of bridges and, potentially, other public infrastructure projects.

This work was supported by NSERC Discovery Grant (RGPIN-2020-05778) to PG and an NSERC Discovery Grant (RGPIN-2020-04097) as well as a SSHRC Insight Grant (435-2023-0015) to DBW.

23.302 ENCOUNTERS WITH SEMANTIC VIOLATIONS DO NOT INTERFERE WITH IMMEDIATELY SUBSEQUENT SCENE-VIEWING BEHAVIOR

Alan Lu[1] (azlu@ucdavis.edu), John Henderson[1]; [1]University of California - Davis

Eye movements and visual attention are guided by expectations on semantic informativeness, such as relevance to scene category and to constituent objects in the scene. Expectations are generated based on implicit 'scene grammar' rules on what and where objects are likely to appear in a given scene. Given the pivotal role of semantic expectations, it is therefore unsurprising that unexpected elements like semantic violations have been found to be more difficult to recognize and to require more cognitive effort to process at fixation. Yet, little is known about the effects of semantic violations on subsequent viewing behavior. Here, we explored whether encountering a semantically inconsistent object has persistent effects on oculomotor programming and semantic guidance on the first three fixations post-exit from the object. Eye-tracking data were collected from 102 participants viewing 62 scenes within the SCEGRAM image database (Öhlschläger & Võ, 2017), each featuring either a consistent or inconsistent critical object insertion. We replicated previous findings and showed reliably elevated fixation rates and dwell times on the inconsistent objects compared to the consistent objects. However, post-exit eye movement characteristics in fixation durations and saccade amplitudes did not show any significant differences. Furthermore, inconsistent object semantics did not significantly influence subsequent attentional guidance. Encountering an inconsistent object neither pushed post-exit visual attention towards regions more conceptually similar to the object nor interfered with guidance from other sources of semantic scene information. Overall, the results showed that disruptions from semantic violations did not extend beyond their object borders and that inconsistent semantics were ignored or suppressed when deciding where next to attend. The current work is the first to study attention following fixation on a semantic violation and advances our understanding on how the visual system handles and adapts to unexpected elements in realistic visual environments.

This research was funded by: NSF BCS2019445 (PI: Henderson) NEI R01EY027792 (PI: Henderson) NEI T32-EY015387 (PI: Burns)

23.303 IMAGE CHARACTERISTICS, NOT TASK, INFLUENCE INTEROBSERVER CONSISTENCY

Sophie Heer[1] (s.heer@qmul.ac.uk), Marek Pedziwiatr[1], Antoine Coutrot[2], Peter Bex[3], Isabelle Mareschal[1]; [1]Queen Mary University of London, [2]LIRIS, CNRS, University of Lyon, [3]Northeastern University

Interobserver consistency (IOC) refers to the degree of similarity in the gaze patterns of different observers when viewing the same image. Here, we investigated the role of image characteristics and viewing task on IOC. We recorded the eye movements of 80 participants while they viewed 192 images of scenes from the ADE20K dataset (Zhou et al., 2017). To vary task demand, participants either freely viewed the images or answered a content-related question following each image. We also considered how two image characteristics – visual clutter and semantic complexity – might impact IOC. We calculated visual clutter using the Feature Congestion measure (Rosenholtz et al., 2005). Additionally, to calculate semantic complexity, we utilised the object labels available for the ADE20K images. First, we used a distributional semantic model (GloVe; Pennington et al., 2014) to compute the pairwise semantic dissimilarity between the object labels in each image. Then, we averaged these values to derive an overall 'semantic complexity' score per image, where images with greater dissimilarity among their objects were considered more complex. Interobserver consistency was assessed by how well the fixation heatmap of all observers, excluding one, predicted the fixation heatmap of the remaining observer, using Pearson correlation coefficient. We used a linear mixed-effects model to examine the relationship between visual clutter, semantic complexity and task condition on IOC. Our results revealed an effect of visual clutter ($p < .001$), with IOC decreasing as clutter increased. In contrast, semantic complexity showed a positive association with IOC ($p = .033$), suggesting increased consistency with greater semantic complexity. Intriguingly, we found no effect of task on IOC. Taken together, our study highlights the nuanced relationship between image characteristics and IOC. Furthermore, the surprising lack of task effect indicates that, in our study, image characteristics play a more prominent role in influencing interobserver consistency.

This research is funded by a Leverhulme Trust grant (RPG-2020-024) awarded to Isabelle Mareschal, Peter Bex and Antoine Coutrot.

23.304 INFORMATION REDUNDANCY FACILITATES EFFICIENT VISUAL PROCESSING

Kaiki Chiu[1] (kaikichiu@gmail.com), Emily Lo[1], Quinn O'Connor[1], Michelle R. Greene[1]; [1]Barnard College, Columbia University, Department of Psychology

Human visual processing is rapid and accurate even in the face of complex real-world scenes. A common framework for explaining this feat posits that the brain creates efficient representations of visual inputs by capitalizing on statistical redundancies (Attnaeve, 1954). A key implication of this framework is the testable hypothesis that images with higher redundancy (i.e., lower information content) undergo more efficient processing than less redundant (i.e., higher information content) counterparts. However, quantifying the information content of images may be an intractable challenge (Chandler & Field, 2007). In this study, we propose a novel approach by focusing on the relative information content within scenes, which we will estimate by measuring the relative compression efficacy of these images through widely accessible algorithms, such as JPEG and PNG. Specifically, our rationale is that more easily compressed images are likely to possess greater redundancy and thus less information. We amassed a database comprising ~1000 photographs of everyday scenes in RAW image format. We compressed each image in PNG (lossless) format and compared the file size differences between the original and compressed images. From this dataset, we selected the 100 most and 100 least compressible images. Observers (N=39) performed a rapid detection task in which they distinguished between scene images and 1/f noise (SOA: ~80 ms, with a dynamic pattern mask). Observers had higher scene detection sensitivity for the highly compressible images ($d'=3.15$ vs 2.88, $p<0.001$), indicating that images with lower relative information content were processed more easily, supporting our hypothesis. Our findings demonstrate that visual processing efficiency is influenced by the relative information content of scenes, as measured by compression algorithms. This aligns with the efficient coding hypothesis, suggesting that the visual system achieves efficiency by exploiting environmental regularities.

NSF CAREER 2240815 to MRG.

23.305 LEXICAL ACCESS TO SCENE FUNCTION RELIES ON ANCHOR OBJECT PRESENCE

Lea Alexandra Müller Karoza[1] (muellerkaroza@psych.uni-frankfurt.de), Sandro Luca Wiesmann[1], Melissa Le-Hoa Vo[1]; [1]Goethe University Frankfurt, Scene Grammar Lab, Germany

Effortless engagement with our surroundings relies on the purposeful arrangement of functional elements within a room for efficient use. Central to these functional clusters are so-called anchor objects, predicting the location of local objects (e.g., soap and toothbrushes near the sink-anchor). Additionally, scene categorization relies on scene function, or affordance, more than on objects (Greene et al., 2016). To examine how anchors affect affordance understanding, we primed a lexical decision task (LDT) on action words (i.e., “washing hands”) with scene images lacking related (REL; sink) or unrelated (UNREL; shower) anchors, or random objects (RAND; shelves). Images from other categories (e.g., a kitchen) served as controls. In Experiment 1, stimuli comprised real photos with pixel masks hiding objects. Participants were quickest in the RAND condition and slower when an anchor was missing, regardless of its action relevance (REL or UNREL). In Experiment 2, using 3D-rendered scenes with whole objects removed, participants were fastest in the RAND condition. Notably, removing a related anchor impeded lexical decision more than removing unrelated anchors. To ensure that visually sparser rendered scenes were still identifiable when anchors were absent, participants categorized the scenes in a control experiment. Scene type (real versus 3D rendered) and missing object (random versus anchor) impacted categorization, with no interaction, suggesting that reported differences between experiments are not solely due to varying scene categorization, but might rely on realism or the information remaining in scenes (i.e., prevalent objects in rendered scenes versus clutter, context- and texture-cues in photographs of real scenes). Experiment 1 implies scene-level affordance understanding using photos and pixel masks, while Experiment 2 suggests object-specific affordance understanding with sparser 3D-rendered scenes. Conclusively, understanding scene affordances flexibly involves both specific object-level information and broader scene context depending on their diagnosticity for scene affordance assessment and the context provided by the scene.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 222641018 SFB/TRR 135 TP C7 granted to MLHV and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project ‘The Adaptive Mind’).

23.306 TRACKING THE DEVELOPMENTAL TRAJECTORIES OF SEMANTIC AND SYNTACTIC ASPECTS OF VISUAL COGNITION IN CHILDREN

Dilara Deniz Türk[1] (dilaradeniz.turk@gmail.com), Daniela Bahn[2], Christina Kauschke[2], Melissa Le-Hoa Vo[1]; [1]Goethe University Frankfurt, Scene Grammar Lab, Germany, [2]Philipps University Marburg, Germany

Humans learn to associate objects and their locations in everyday scenes based on some rules and they develop expectations about what objects (semantics) should be where (syntax). For instance, we expect to find a pot in a kitchen rather than in a bathroom and we expect to see certain objects placed around a pot. The current study aims to explore the developmental trajectory of visual cognition in children. The eye movements of 39 children, aged 6-10, were tracked during a free viewing task involving scenes with consistent, semantically inconsistent, or syntactically inconsistent objects, as well as during a search task with objects at expected or unexpected locations. In addition to the implicit eye movement measures from these tasks, explicit measures were obtained as children furnished a dollhouse with 61 objects. Our findings replicated previous findings of consistency effect involving young adults and children aged 2-4, demonstrating a reduction in dwell time, first fixation time, and reaction time for consistent objects in both tasks. This reduction indicates strengthened predictions and faster attentional disengagement for objects in their familiar context or location. Moreover, as the dwell time for syntactic violations and first fixation time for inconsistently placed objects increased, the performance in the dollhouse task increased (i.e. smaller distance between related objects). No such relationship was observed between dwell times on semantic violations and placing the objects in correct rooms. Additionally, age-related trends were evident, with children exhibiting improved object placement performance, and also greater fixation delay on unexpectedly placed objects as they grew older. These findings suggest that scene knowledge still becomes more refined between 6 to 10 years. In summary, this study sheds further light on the impact of scene knowledge on implicit and explicit behaviors of children, providing valuable insights into the developmental aspects of visual cognition.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 222641018 SFB/TRR 135 TP C7 granted to MLHV and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project 'The Adaptive Mind').

23.307 UNVEILING CORE, INTERPRETABLE IMAGE PROPERTIES UNDERLYING MODEL-BRAIN SIMILARITY WITH GENERATIVE MODELS

Yingqi Rong[1], Colin Conwell[1], Dianna Hidalgo[2], Michael Bonner[1]; [1]Johns Hopkins University, [2]Harvard Medical School

Deep Neural Networks (DNNs) are now capable of predicting the hierarchy of natural images representations in human visual cortex with substantial accuracy. However, a key challenge in the use of these networks to predict representations in the brain is discerning the specific properties of these networks that underlie their predictive accuracy. In this work, we developed an approach for leveraging high-throughput generative vision models to run targeted, hypothesis-driven experiments on the key image properties that drive DNN predictions of brain representation. Specifically, we used diffusion models to create diverse image variations while preserving targeted image information. This targeted information included specific visual features (e.g. edges, background) as well as semantics from captions and categories. Using our synthesized image variations, we quantified the impact of each interpretable manipulation on the representational similarity between AlexNet activations and image-evoked fMRI responses in early visual and occipital temporal cortex (EVC, OTC). We found that representational similarity to high-level OTC (but not EVC) was stable so long as the synthesized images retained their semantic content, and this effect was robust to substantial structural variations in the synthesized images. To demonstrate the broad utility of this method, we quantified the influence of objects, backgrounds, shapes, and other visual details on model performance, and we performed analogous targeted experiments on aspects of higher-level scene semantics (e.g. object relations). Overall, these findings highlight the promise of employing generative models to probe brain-model similarities. Our work provides insight into how specific forms of image information shape the relationship between computational models and brain responses, and it paves the way for a deeper understanding of how models approximate biological visual processing.

23.308 ARE YOU SEEING THIS? AN INVESTIGATION INTO THE VARIABILITY OF MENTAL IMAGERY AND PROPENSITY OF APHANTASIA IN A NON-SELF-DIAGNOSED POPULATION

Patsy E. Folds[1] (pefold8941@ung.edu), Courtney G. Nutt[1], Tanner L. Lumpkin[1], Erin A. Conway[1], Ralph G. Hale[1]; [1]University of North Georgia

Mental imagery is the experience of visual representations in the "mind's eye" without external stimulation. A portion of the population, ranging somewhere between 0.7-15.3% (Betts, 1909; Faw, 2009; Zeman et al., 2020; Dance & et al., 2022; Beran & et al., 2023) has an inability to voluntarily generate mental images, a phenomenon referred to as aphantasia. Studies investigating the prevalence of aphantasia have used measures such as the Vividness of Visual Imagery Questionnaire-2 (VVIQ) (Marks, 1995), Spontaneous Use of Imagery Scale (SUIS) (Nelis & et al., 2014), and Questionnaire upon Mental Imagery (QMI) (Sheehan, 1967). However, the vast majority of prior research has relied on self-selected sampling, i.e., individuals diagnosing themselves with aphantasia. In the present study, participants from a healthy population were sampled. All completed the VVIQ-2, SUIS, and QMI to measure mental imagery ability. Afterward, they reported any familiarity with aphantasia and whether they felt they had the condition. A wide range of mental imagery abilities were found between individuals in our sample, only a minority of which falling in the aphantasia range. However, a significant difference was found between each of these measures indicating variations related to convergent and divergent validity. Participant knowledge of aphantasia and questions related to "self-diagnosis" will be discussed as it pertains to aphantasia generally and to each of our measures. The present study highlights the extensive range of mental imagery abilities as well as the variability between measures used to assess imagery and aphantasia. Unlike previous studies examining aphantasia, participants here were not aware we were looking for individuals with aphantasia and they were not asked to self-diagnose. This study is the first of its kind to examine the impacts self-diagnosis and familiarity with aphantasia may have on reporting mental imagery ability.

23.309 EXCESSIVE NOISE EXPLAINS THE IMPAIRED VISUALLY-GUIDED NAVIGATION ABILITIES IN ADULTS WITH WILLIAMS SYNDROME: A COMPUTATIONAL APPROACH

Zvi Shapiro[1] (zrshapi@emory.edu), Alex Weigard[2], Quyen Cao[1], Daniel Dilks[1]; [1]Emory University, [2]University of Michigan

Recent evidence indicates our ability to recognize and navigate through places are causally dissociable: Adults with Williams syndrome (WS) – a genetic disorder – are impaired on "visually-guided navigation" (VGN) yet spared on "scene categorization". But why? Here, we evaluated three, computationally derived mechanisms that may explain this VGN impairment: i) too little signal, ii) too much noise, or iii) an emphasis on speed (over accuracy). Twenty adults with WS and twenty mental age (MA) matched controls completed a VGN task and a scene categorization task (as a control). Then using each group of participants' response times and accuracy for each task, we fit a hierarchical Linear Ballistic Accumulator model, an evidence accumulation model of decision-making, and operationalized the rate of information accumulation to the correct response (i.e., correct drift rate) as "signal," the rate of information accumulation to the incorrect response (error drift rate) as "noise," and the amount of information needed before making a decision (boundary) as "speed-accuracy trade-off." For the VGN task, we found a significantly greater correct drift rate in WS compared to MA controls, revealing that the WS adults have more – not less – signal than controls, and thus cannot explain the VGN impairment. By contrast, we found both a significantly greater error drift rate and a significantly lower boundary in WS compared to MA controls, suggesting that the WS adults have more noise and a greater emphasis on speed than controls, consistent with their VGN impairment. Interestingly, however, the "excessive noise" effect was specific to the VGN task (i.e., it was not found in the scene categorization task), while the "speed emphasis" effect was found in both tasks. Taken together, these results reveal that the VGN impairment in WS, relative to scene categorization, is driven by high levels of noise.

23.310 LIMITED SCENE UNDERSTANDING AND ALTERED EYE MOVEMENTS WITH SIMULATED CENTRAL AND PERIPHERAL VISION LOSS

Byron A. Johnson[1] (byron_johnson@ucsb.edu), Shravan Murlidaran[1], Ansh Soni[1], Michael Beyeler[1], Miguel P. Eckstein[1]; [1]University of California, Santa Barbara

Macular degeneration and retinitis pigmentosa are two leading causes of vision loss in which complementary parts of the retina are degraded or spared over time. Previous research has shown that vision loss impairs visual search, reading, and navigation, however, less is known about how these conditions affect scene perception. In order to understand how central and peripheral vision loss affects scene perception, fourteen observers were presented with 120 scenes in one of three conditions: central vision loss (CVL), peripheral vision loss (PVL), and no simulation. The scene images were categorized into two groups: 50% social interaction scenes and 50% MS COCO images. The CVL condition presented scenes with a gaze contingent central scotoma (Gaussian, SD = 5 degrees). The PVL condition presented scenes with a clear central aperture (5 degrees diameter) and Gaussian blurred periphery to simulate tunnel vision. The simulation tracking latency used a maximum update speed of 7 ms. Trials were terminated after one or three saccades. Scene description ratings were calculated with Sentence BERT Scores on participant description's semantic similarity to gold standard descriptions generated with unlimited viewing time. Results with stop words removed show that both CVL and PVL conditions led to descriptions that were rated significantly lower than descriptions generated without impairment (ANOVA, $F = 8.60$, $p = 0.006$). The effect of viewing condition significantly interacted with scene type, suggesting that descriptions made with CVL and PVL were worse than the no impairment condition for scenes with social interactions ($F = 7.38$, $p = .009$). Although Tukey's HSD revealed no difference in ratings between CVL and PVL, the CVL condition led to increased saccade amplitudes compared to PVL and no impairment ($F = 9.05$, $p = .006$). These findings show how CVL and PVL can influence eye movements for rapid scene understanding.

23.311 PERIPHERAL VISUAL INFORMATION IS NECESSARY FOR VISUALLY-GUIDED NAVIGATION

Rebecca Rennert[1] (rebecca.rennert@emory.edu), Avital Okrent Smolar[2], Jose Cijin[2], Deepta Ghate[2], Daniel Dilks[1]; [1]Emory University, Department of Psychology, [2]Emory University, Department of Ophthalmology

It has been recently hypothesized that "visually-guided navigation" relies on peripheral visual input, while "scene categorization" relies on central visual input. Indeed, finding one's way through the immediately visible environment should require only coarse representations of boundaries and obstacles to avoid, as opposed to fine-grained representations needed to ultimately identify the spatial layout (e.g., open versus closed) and/or scene content (i.e., objects in the room) necessary to recognize a place as a particular kind of place. Here we provide the strongest test of this hypothesis by investigating visually-guided navigation and scene categorization abilities in individuals with loss of peripheral vision and sparing of central vision – that is, individuals with glaucoma. If peripheral visual input is necessary for visually-guided navigation, then individuals with glaucoma will be impaired on a visually-guided navigation task, while spared on a scene categorization task. Indeed, we found that individuals with glaucoma performed disproportionately worse on the visually-guided navigation task compared to the scene categorization task, relative to typically-sighted controls. These results provide the first causal evidence that peripheral visual information is necessary for visually-guided navigation.

23.312 EXAMINING THE RETINOTOPIC ORGANIZATION OF THE VISUAL HIERARCHY IN AUTISTIC AND NEUROTYPICAL INDIVIDUALS

Yeo Bi Choi[1] (yeo.bi.choi.gr@dartmouth.edu), Adam Steel[1], Brenda Garcia[1], Edward Silson[2], Caroline Robertson[1]; [1]Dartmouth College, [2]University of Edinburgh

As information ascends the visual hierarchy, smaller receptive fields that are selective for simple visual features combine to form larger ones that are selective for more complex features. Previous research demonstrates different receptive field sizes in lower levels of the autism visual hierarchy, which vary with symptom severity (Schwarzkopf et al., 2014). In this study, we examined whether such group differences propagate to the category-selective levels atop the visual hierarchy. We used population receptive field (pRF) analysis to test whether pRF sizes are atypical in both lower (early visual cortex, EVC) and higher levels (parahippocampal place area, PPA; occipital place area, OPA) of the visual hierarchy in individuals with autism spectrum conditions (ASC; n=6) compared to neurotypical controls (n=17). fMRI participants viewed traversing bar stimuli of scene fragments (4 orientations x 2 directions, each 36s, diameter=11.4°) and a traditional category localizer rapidly presenting blocks of scene, face, and object images. In both groups, pRF sizes grew with eccentricity in the EVC (one-way ANOVA, ASC $p<0.001$, Con $p<0.001$) with no interaction (Diagnosis*Eccentricity $p>0.05$). From EVC to category-selective areas, pRF sizes increased in both groups. In both groups, the size of pRFs in PPA was larger than those in EVC (two-sample t-test, ASC $t(10)=-4.43$, $p<0.01$, Con $t(32)=-9.06$, $p<0.001$). However, in both PPA and OPA (permutation-test, $p<0.05$) the ASC group had a smaller pRF size than that of controls, when comparable in EVC. These results show that the fine-grained visual architecture in autism differs from that of controls in category-selective areas. Specifically, we observe smaller pRFs in category-selective scene areas. Future work should explore pRF sizes in other category-selective areas, and how these relate to autistic sensory and cognitive traits.

Saturday, May 18, 8:30 am – 12:30 pm, Banyan Breezeway

Spatial Vision: Neural mechanisms

23.313 VISUAL ACTIVITY IN PRIMATE SUPERIOR COLLICULUS DEPENDS ON VISUAL CORTEX

Leor N Katz[1] (leor.katz@nih.gov), Gongchen Yu[1], Richard J Krauzlis[1]; [1]Laboratory of Sensorimotor Research, National Eye Institute, Bethesda, Maryland 20892 USA

In the primate, the primary target of retinal ganglion cells is the lateral geniculate nucleus (LGN), the main relay station from retina to cortex. A secondary target is the superior colliculus (SC), which receives an estimated 10% of retinal projections. These direct retinotectal projections have been speculated to support a range of visual functions including blindsight, express saccades, and rapid threat and face detection, but their functional significance in visual processing is mostly unknown. Here we used linear arrays to record 182 SC neurons in two rhesus macaques performing simple visual tasks, before and during LGN inactivation, to determine which aspects of visual and movement-related processing were preserved when retinal inputs remained, but visual cortical inputs were blocked. Inactivation of LGN was performed by injecting ~1μl of muscimol and confirmed by documenting a visual scotoma—a region of the visual field in which monkeys could not detect stimuli—measured during a saccade task on all eight inactivation sessions. Before LGN inactivation, SC neurons exhibited strong visual responses to stimuli in their RF. During LGN inactivation, visual responses were largely eliminated. This loss of responsiveness was evident across all SC neurons regardless of the form of visual stimulus (e.g., static objects, motion, looming) and regardless of functional class: neither "visual" nor "visual-movement" neurons (typically associated with the superficial and intermediate layers of SC, respectively) responded to the onset of a stimulus in their RFs. Movement-related activity measured during spontaneous saccades directed into the neurons' RF, in contrast, remained largely unchanged before versus during LGN inactivation, indicating that non-visual inputs could still elicit activity. These results show that SC visual responses in the awake primate are primarily dependent on signals routed through the LGN and visual cortex, and that retinotectal inputs by themselves have a limited functional role in visual processing.

National Eye Institute Intramural Research Program at the National Institutes of Health ZIA EY000511

23.314 MAGNOCELLULAR MODULATION OF SPATIAL VISION

Kristian Paul Skoczek[1], Ljubica Jovanovic[2], Paul McGraw[1], Neil Roach[1], Alan Johnston[1]; [1]University of Nottingham, [2]Ecole Normale Supérieure, PSL University & CNRS, Paris

Perceived spatial separation between two points can be influenced by adaptation to a dynamic texture pattern, such as moving or flickering dots. We utilised this adaptation effect to probe magnocellular influence on spatial perception by manipulating the spatial and temporal frequency of the adapting stimulus and isolating the contribution of colour and luminance mechanisms. Neurones in the magnocellular layers of the lateral geniculate nucleus respond at higher temporal frequencies than neurones in primary visual cortex. Magno-cells also respond most at high temporal and low spatial frequencies, and are less sensitive to isoluminant chromatic patterns. After viewing an adapting pattern in one hemifield, participants reported which of two pairs of dots (presented in adapted and un-adapted hemifields) appeared to have greater separation. The separation of one of the dot pairs was varied to derive the point of subjective equality (PSE) and provide a measure of the compression effect in comparison to a baseline condition with no adaptor stimulus. In a sequence of experiments, we found that: (1) Adapting to luminance-defined dot patterns flickering at high temporal frequency (60Hz) induced significant spatial compression despite the adaptor being invisible; (2) After adaptation to a Gabor array, compression was strongest for arrays with lower (0.5 cpd) spatial frequency carriers; (3) Adapting to colour-defined isoluminant dots produced significant compression at low (3Hz) but not high (60Hz) flicker rates; (4) Compression induced by a pattern of moving coloured dots with varying luminance content produced compression after-effects that were weakest at isoluminance, and coincided with a loss of perceptual motion coherence. Across all experiments, adaptor properties preferentially targeting the magnocellular pathway produce stronger compression after-effects. Specifically, the compression effect is maximal within spatially low-pass, temporally band-pass transient luminance channels, indicating that metric properties of spatial vision are encoded by adaptable neural processes reflecting magnocellular pathway specialisation.

This work was funded by a Leverhulme Trust Research Project Grant

23.315 EXAMINING THE PRECISION OF SPATIAL REPRESENTATIONS WITHIN VISUAL CORTEX

Ilona M. Bloem[1,2] (i.bloem@nin.knaw.nl), Janneke F.M. Jehee[3], Serge Dumoulin[1,2,4]; [1]Netherlands Institute for Neuroscience, Amsterdam, the Netherlands, [2]Spinoza Centre for Neuroimaging, Amsterdam, the Netherlands, [3]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands, [4]Vrije University Amsterdam, Amsterdam, the Netherlands

The representation of information across space is fundamental to visual perception. It is well known that the visual field is not uniformly organized within visual cortex. How does the organization of visual cortex contribute to the precision in spatial representations? In the present study, we aim to extend the current understanding of spatial representations, by examining the precision of these representations along the continuum of angular location. Participants viewed a small checkerboard (0.5 degree visual angle) briefly presented (500 ms) at random locations along an iso eccentric circle (2.5 degree eccentricity). After a short delay, participants reported the target's location by moving a probe along the circle to the perceived location as precisely as possible. Analysis of the behavioral results shows that this simple spatial localization task resulted in large, systematic misrepresentations of angular location (up to ± 10 degrees mean angular error), consistent with prior reported categorical biases. In order to examine the precision in these behavioral judgements, we removed these location-dependent repulsive biases from the data, grouped trials into bins based on the presented location, and computed the variance of the behavioral errors across trials within each bin. We found that behavioral variability varied as a function of angular location: behavioral judgements had the greatest precision at the horizontal meridian, and the least precision, or largest perceptual uncertainty, in angular position at off-cardinal locations. In ongoing work, we will investigate the degree to which these behavioral values of precision can be linked to fMRI measures of precision in the visual cortex. Together, these findings will provide insights into the neural implementation of spatial representations in the human visual cortex.

016.Vici.185.050

23.316 SPATIAL FREQUENCY ADAPTATION MODULATES SPATIAL TUNING IN HUMAN VISUAL CORTEX

Ecem Altan[1] (altan.ecem@hotmail.com), Catherine Morgan[1], Steven Dakin[1,2], D. Samuel Schwarzkopf[1,2]; [1]The University of Auckland, New Zealand, [2]University College London, United Kingdom

Neurons in the early visual areas are selectively sensitive to spatial frequency (SF). However, the relationship between SF and the spatial tuning of neuronal populations has not been directly studied in humans. Here, we explored the interplay between SF adaptation and its effect on the size of population receptive fields (pRFs) in the human visual cortex. We reasoned that prolonged exposure to high and low SFs would lead to a selective decrease in the sensitivity of neurons with small and large receptive fields, thereby affecting overall pRF sizes, as measured via functional magnetic resonance imaging (fMRI). We first conducted a psychophysical experiment to quantify the subjective perceptual changes after adaptation to bandpass-filtered isotropic noise stimuli with SFs of 0.5 and 3.5 cpd. Next, we performed a fMRI experiment that integrated the SF adaptation paradigm into a standard pRF mapping procedure. This enabled the measurement of pRF size changes after adaptation to two noise stimuli with relatively high and low SFs. The perceptual aftereffect confirmed significant over- and underestimations of SF after adaptation to low and high SFs. Most importantly, our fMRI results showed that adaptation to a certain SF modified the spatial tuning of neuronal populations. As predicted, low and high SF adaptation resulted in smaller and larger pRF sizes, respectively. Our results provide the most direct evidence to date that the spatial tuning of the visual cortex, as measured by pRF mapping, is directly linked to the spatial frequency selectivity of visual neural populations. Our study has implications for our understanding of size perception, visual acuity, and sensitivity to blur.

Supported by the Marsden Fund Council from Government funding, managed by Royal Society Te Apārangi

23.317 COVARIATION IN THE SURFACE AREA OF HUMAN PRIMARY VISUAL CORTEX AND CORTICAL SPATIAL FREQUENCY TUNING

Marc M. Himmelberg[1], Yuna Kwak[1], Marisa Carrasco[1], Jonathan Winawer[1]; [1]New York University

Introduction. The surface area of human primary visual cortex (V1) is highly variable, with a 2-to-3-fold range among healthy adults, and a coefficient of variation (CoV) of around 0.17, much greater than the surface area of the cerebral cortex, with a CoV around 0.06. This variation likely has implications for visual encoding. Here, we used fMRI to ask how the surface area of V1 is related to its spatial frequency tuning. Methods. In each observer (n=34), we used fMRI to derive retinotopic maps and quantified overall V1 surface area (to 12° eccentricity) and cortical magnification as a function of eccentricity and polar angle. In a separate fMRI session, we measured V1 spatial frequency tuning, estimating preferred frequency as a function of eccentricity and polar angle. We also derived an overall spatial frequency tuning metric for each observer –cycles per V1– by integrating the functions over eccentricity and polar angle. Results. Our data showed substantial variability in V1 surface area (CoV=0.15), and even larger variability in overall spatial frequency tuning (CoV=0.36). Across observers, V1 surface area positively correlated with cycles per V1 (r=0.50); indicating that larger V1s preferred higher spatial frequencies. At the group level, the pattern of V1 cortical magnification and preferred spatial frequency systematically covaried with visual field location. Both were greatest near the fovea and decreased with eccentricity. Moreover, cortical magnification and preferred spatial frequency were greatest at the horizontal, intermediate at the lower vertical, and lowest at the upper vertical meridian. Conclusions. These data reveal a link between the surface area of V1 and its spatial frequency tuning –at the individual level, when summed over visual field location, and at the group level, as a function of visual field location.

Funding: NIH R01-EY027401 to MC and JW

23.318 FEEDBACK TO V1 STRONGLY INFLUENCES BOLD SIGNAL DURING CONTEXTUAL MODULATION: EVIDENCE FROM LAMINAR FMRI

Joseph Emerson[1] (emers245@umn.edu), Karen Navarro[2], Cheryl Olman[1,2,3]; [1]University of Minnesota, Graduate Program in Neuroscience, [2]University of Minnesota, Department of Psychology, [3]University of Minnesota, Center for Magnetic Resonance Research

Human visual perception relies on contextual information to make inferences about spatially localized features in visual scenes. Similarly, neural representations of local features in primary visual cortex (V1) are shaped by broad spatial context through long-range lateral connections and feedback from higher-order visual areas. However, it is unclear exactly how and to what extent lateral and feedback connectivity individually contribute to contextual modulation of neural responses in V1. Ultra-high-field fMRI have enabled non-invasive imaging of cortical layers in humans, which can be exploited to examine the cortical origins of neural signals underlying blood-oxygenation-level-dependent (BOLD) contrast. We analyzed data from five participants using 7T fMRI at 0.6 mm isotropic resolution to measure the influence of visual context on BOLD response profiles across cortical depth in V1. Participants viewed sine-wave grating disks embedded in large surround gratings with matched spatial frequency and contrast. Segmentation cues were provided by either an offset in relative orientation or an offset in relative phase between target and surround gratings for a total of three contextual conditions plus a surround-only condition to measure the effects of cortical feedback in the absence of feedforward input. Our analysis isolated the effects of orientation-tuned surround suppression (OTSS) from figure-ground modulation (FGM). Consistent with contextually-driven responses measured in mice and monkeys, we found significant modulation of BOLD signal in target-selective voxels in the absence of feedforward input. While we found strong signatures of FGM in superficial and deep layers, we did not find significant modulation of the BOLD signal due to OTSS. Our results suggest that the mechanisms responsible for OTSS have a weaker impact on the BOLD signal. We conclude that a large proportion of the BOLD signal measured in V1 depends on feedback from higher-order visual cortex, which is reflected in contextually-dependent changes in laminar profiles.

NIH Grant 5T32MH115886-04, NIH Grant 5R01MH111447-05, NIBIB P41 EB027061, NIH Grant P30 NS076408, NIH Grant 1S10OD017974-01, NIH Grant NS123482

23.319 INVESTIGATING ON AND OFF PATHWAY SPATIAL TUNING IN THE UPPER AND LOWER VISUAL FIELDS USING THE SSVEP

Martin Scott[1], Alexandra Yakovleva[1], Anthony Norcia[1]; [1]Stanford University

The ON and OFF retino-cortical processing pathways encode increments and decrements (respectively) in perceived luminance. Recent evidence suggests that the ON and OFF pathways may be differently affected by selective vision loss, encouraging the diagnostic perimetric assessment of ON & OFF pathway function. This calls for an improved understanding of normative ON and OFF pathway function across the visual field. Here, using electroencephalography (EEG), we aim to investigate whether spatial tuning in the ON and OFF pathways differs between the upper and lower visual fields. In 39 healthy observers, we measured ON- and OFF-biased visually evoked potentials (VEPs) using a 128 channel EEG system. Responses from the upper and lower visual fields were simultaneously measured using a spatial array of flicker-frequency tagged hexagonal probes (3.75Hz and 3Hz, respectively). This low-contrast (20%) luminance flicker had a saw-tooth profile, the polarity of which was set to bias responses towards the ON or OFF pathways. Across 8 conditions, we manipulated the polarity and spatial scale (size & number) of the stimuli elements, clamping probe surface area. After a process of dimension reduction, VEP amplitudes were fit with a linear model at the group-level for statistical inference. Reliable upper and lower visual field neural responses for the ON and OFF pathways were found at all spatial scales. As previously reported, OFF pathway responses were larger than ON pathway responses. However, the OFF pathway demonstrated a striking preference for smaller and more numerous probes, while the ON pathway was relatively scale invariant. These preferences were similar in the Upper and Lower visual fields. This difference in scale tuning could be formulated as the ON pathway comprising of sub-units with a broader range of receptive field sizes, perhaps produced by weaker centre-surround inhibition in this pathway. This possibility will be investigated in future work.

Research to Prevent Blindness, NIH (R01-EY030361-01, P30-EY026877)

23.320 BUILDING A COMPREHENSIVE TOOLKIT FOR HUMAN VISUAL CORTEX PARCELLATION

Fernanda L. Ribeiro[1] (fernanda.ribeiro@uq.edu.au), Torin Bambridge-Lozan[1], Noah C. Benson[2], D. Samuel Schwarzkopf[3], Alexander M. Puckett[1], Steffen Bollmann[1]; [1]University of Queensland, [2]University of Washington, [3]University of Auckland

The visual system comprises several functionally specialized cortical visual areas, where adjacent neurons represent adjacent retinal locations. These retinotopic maps are typically defined in polar coordinates, resulting in two orthogonal coordinate maps: one representing polar angle and the other eccentricity. While the retinotopic organization of early visual areas (V1, V2, and V3) in the human visual cortex is generally assumed to be organized according to a universal topological template that is similar across people, recent investigations have revealed compelling evidence of interindividual topological differences (Ribeiro et al., 2023; DOI:10.7554/eLife.86439). These differences cast doubt on the traditional template of early visual cortex organization. Therefore, we propose a unified, automated solution for retinotopic mapping and visual cortex parcellation based only on anatomical data derived from a T1-weighted image and that is not dependent on any single template of retinotopic organization. Our toolkit integrates (1) standard neuroimaging software (FreeSurfer 7.3.2 and Connectome Workbench 1.5.0) for anatomical MRI data preprocessing, (2) a deep-learning model (Ribeiro et al., 2021; DOI:10.1016/j.neuroimage.2021.118624) for predicting retinotopic maps at the individual level, and (3) an efficient implementation of the visual field sign analysis (Sereno et al., 1994; DOI:10.1093/cercor/4.6.601) for early visual areas parcellation. These components are packaged into Docker and Singularity software containers, which can be easily downloaded for local use and are available on Neurodesk (Renton et al., 2023; DOI:10.1101/2022.12.23.521691). Our toolkit can generate detailed, individual-specific retinotopic maps. Moreover, with polar angle and eccentricity maps, our toolkit generates visual field sign representations with unambiguous boundaries between early visual areas. These results demonstrate the potential of our open-source toolbox (https://github.com/felenitaribeiro/deepRetinotopy_TheToolbox) for individual-specific visual cortex parcellation.

This work was supported by the Australian Research Council (LP200301393) and the UQ AI Collaboratory.

23.321 TOWARDS A DETAILED FUNCTIONAL NEUROANATOMY OF THE VENTRAL VISUAL CORTEX

Josefine A. Zerbe[1,8], Maggie Mae Mell[1,8], Juliane Damm[2], Marianna Schmidt[2], Nikolaus Weiskopf[2,3], Evgeniya Kirilina[2], Tomas Knapen[4,5,6,8], Martin N. Hebart[1,7,8]; [1]Vision and Computational Cognition Group, Max Planck Institute of Human Cognitive and Brain Sciences, [2]Department of Neurophysics, Max Planck Institute of Human Cognitive and Brain Sciences, [3]Felix Bloch Institute for Solid State Physics, Leipzig University, [4]Spinoza Centre for Neuroimaging, [5]Netherlands Institute for Neuroscience, [6]Experimental and Applied Psychology, Vrije Universiteit Amsterdam, [7]Department of Medicine, Justus Liebig University Giessen, [8]equal contribution

Previous research of human visual cortex has revealed functionally distinct and hierarchically organized processing pathways, each consisting of sequences of retinotopic maps. However, beyond early visual cortex the extent and even presence of these maps is not generally agreed upon. Specifically, it is unclear if variability in retinotopic structure reflects true idiosyncratic effects or noise in the retinotopic mapping data. Moreover, beyond the functional organization, the fine-grained structural connectome underlying human visual cortex is largely unknown and until recently had to be inferred indirectly from post-mortem studies. This has made it challenging to relate the macroscale structural organization of visual cortex to its functional topographic organization. To address these challenges, here we provide a densely-sampled dataset of 7 individuals, combining ultrahigh-resolution functional, structural, and diffusion data across 12 scanning sessions. We rigorously preselected individuals based on the reliability of an initial retinotopic mapping scan. Across 7 sessions of 7T MRI, we collected 192 minutes of retinotopy, as well as extensive object and motion-specific localizer and resting state data supplemented by multi-parameter mapping anatomy. Across three retinotopic mapping tasks, we (1) used moving bar apertures of varying width, (2) focused on the foveal representation (central 4 dva), and (3) focused on the periphery by moving the fixation cross to each corner of the screen. Across 5 additional sessions of 3T Connectom MRI, we further collected structural and diffusion data at 0.8mm isotropic resolution, allowing for a detailed mapping of short association fibers between adjacent brain regions. Data quality analyses of the functional data revealed minimal head motion and high noise ceilings, offering detailed, individually specific retinotopic maps. Paired with the densely sampled ultrahigh-resolution diffusion data, this dataset promises a highly detailed understanding of the functional neuroanatomy of the human visual system.

This work was supported by a Max Planck Research Group Grant and the ERC Starting Grant COREDIM awarded to MNH.

23.322 CORTICAL DEPTH-DEPENDENT POPULATION RECEPTIVE FIELD SIZE VARIATION IN HUMAN V1, V2 AND V3

Maya Jastrzebowska[1] (maya.jastrzebowska@gmail.com), Nicolás Gravel[1], Polina Iamschchinina[1,2], Daniel Haenelt[3], Nikolaus Weiskopf[3,4], Radoslaw Cichy[1]; [1]Department of Education and Psychology, Freie Universität Berlin, Germany, [2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Germany, [3]Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [4]Felix Bloch Institute for Solid State Physics, Faculty of Physics and Earth Sciences, Leipzig University, Germany

The early visual cortex is organized in accordance with well-established principles of retinotopy and cortical magnification. Receptive field (RF) size increases systematically with eccentricity from fovea to periphery and along the visual hierarchy. However, the organization of RFs across cortical depths is not yet well understood. While non-human primate neurophysiology shows RF size variation across cortical laminae, evidence in humans is lacking. Here, we used submillimeter fMRI to map RF properties in vivo at the scale of cortical laminae. We measured gradient-echo blood oxygenation level-dependent (BOLD) responses to a drifting bar stimulus using 7 Tesla fMRI and population receptive field (pRF) mapping in four human participants. We projected the fMRI data to eight equivolumetric cortical surfaces based on white matter and pial surface reconstructions. Fitting a pRF model to the BOLD time series of each vertex of each surface, we estimated the location in visual space and pRF size that best explain visual field selectivity. We computed the pRF size at 2 degrees of eccentricity in three early visual regions of interest (ROIs) V1, V2 and V3. For each participant and ROI, we characterized the cortical depth profile of pRF size, as well as the profile of the surround suppression to center excitation ratio (suppression index). We replicate previous findings of a U-shaped relation between pRF size and cortical depth in V1. Moreover, we extend these findings by demonstrating depth-dependent patterns in V2 and V3. Similarly to V1, pRF sizes in V2 are largest in deep layers, followed by superficial and middle layers, with the reverse pattern in V3. The suppression index remains flat across depths, consistent with previous reports. Our findings demonstrate that pRF size variation across cortical depth is robustly quantifiable in humans in vivo, lending support to future examinations of feedforward and feedback mechanisms of spatial vision.

Saturday, May 18, 8:30 am – 12:30 pm, Banyan Breezeway

Development: Clinical and high-level

23.323 DETECTION OF VISUAL IMPAIRMENTS USING CNN AND RED-EYE REFLEX IMAGES

Alexander Lichtenstein[1], Bob Williams[1]; [1]Health Access LLC

Amblyopia, also known as “lazy eye”, is the most common cause of visual impairment in children and affects 3-6% of the population. It is caused by incomplete development of vision due to factors such as refractive errors, strabismus and deprivation of visual signals. Treatment consists of correcting the underlying problem in the visual system, e.g. by correcting refractive errors, realigning the eyes or removing opacities coupled with “penalization therapy” where the better eye is blurred or occluded to force development of vision in the amblyopic eye. However, the effectiveness of treatment decreases with age and is poor after the age of six. Even with early intervention, complete restoration of normal vision is rare and some degree of stereopsis impairment usually persists. Photovision screening based on the Bruckner test, in which the red reflex at the back of the eye is examined, can indicate amblyogenic factors. Advances in smartphone technology with its high-resolution cameras and high computing power offer the possibility of earlier detection of amblyogenic factors by enabling early evaluation of children by parents or caregivers without the need to consult a professional. We present KidsVisionCheck, an app that is suitable for vision screening and enables parents to check their children on a regular basis for eye abnormalities. It is based on the model which uses a convolutional neural network (based on a ResNet model) to detect visual abnormalities from red-eye reflex images. We trained our model using data collected from children's vision screenings and labeled by a trained ophthalmologist. As the result, we were able to achieve good performance in detecting abnormalities based on the red-eye reflex.

23.324 EXAMINING VARIABILITY OF HIGH-LEVEL VISUAL CATEGORIES ACROSS DEVELOPMENT

Kelly Hiersche[1] (hiersche.1@buckeyemail.osu.edu), Anna Quatrale[1], Zeynep Saygin[1]; [1]The Ohio State University

There exist numerous high-level category-selective areas within ventral temporal cortex (VTC), including the word-selective visual word form area (VWFA), face-selective fusiform face area (FFA), and object-selective posterior fusiform sulcus (PFS). Although the existence and functions of these areas have been thoroughly investigated, it remains unclear whether these regions vary more in their selectivity and spatial location or extent in childhood, and how this variability changes across development. Here, using a sample of 42 children, some scanned across multiple timepoints, we attempt to answer this question. We scanned children ages 3-9 on a functional MRI experiment where they saw line drawings of faces, words, scrambled words, and everyday objects. We defined bilateral VWFA, FFA, and PFS fROIs in each participant and calculated selectivity to the preferred category in independent fMRI runs. Children were divided based on age (below or above 6) and matched based on motion on the task. We calculated variability across children by computing distance of each child's selectivity and center of gravity (in fROIs projected to template space) to each other child in their age-group, and also coefficient of variation (CV) of the standard deviation of this distance divided by mean distance. We find that while selectivity increases with age for many fROIs, as reported previously, distance and CV of selectivity are relatively stable with age; center of gravity was more variable in older children for all fROIs. Ongoing investigations include comparisons of spatial location (with respect to anatomical landmarks) of these fROIs across development, as well as comparisons to a motion-matched adult sample to further test the theory that variability increases with greater experience.

Alfred P. Sloan Foundation (to Z.M.S); NSF GRFP (to K.J.H.)

23.325 EFFECT OF CEREBRAL VISUAL IMPAIRMENT ON FUNCTIONAL-STRUCTURAL COUPLING OF THE ATTENTION NETWORKS

Negin Nadvar[1] (nnadvar@mgh.harvard.edu), Marie Drottar[1], Claire Manley[2], Peter Bex[3], Lotfi Merabet[2], Corinna Bauer[1,2]; [1]The Lab of Neuroimaging and Vision Science, Gordon Center for Medical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, [2]Department of Ophthalmology, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, [3]Department of Psychology, Northeastern University, Boston, MA

Visual attentional processes, including selective attention, are often impaired in individuals with Cerebral Visual Impairment (CVI) [1]. Yet, the underlying neural correlates thereof remain unclear. This study investigated potential differences in the relationship between white matter integrity and resting state functional connectivity (rsFC) of the dorsal (DAN) and ventral (VAN) attention networks underlying visual attention [2] in individuals with CVI compared to controls. Resting-state fMRI, diffusion, and METRICS [3] data were collected on a 3T Philips scanner for 25 participants (16 control, 9 CVI). Average rsFC was calculated within the DAN and VAN, and also between the DAN, VAN, thalamus, and early visual regions. Tract volume, fractional anisotropy (FA), radial/axial diffusivity (RD/AD), and myelin water fraction (MWF) and intra-extracellular T2 (IET2) [4] were calculated for the superior longitudinal fasciculus (SLF) (divisions I, II, and III), inferior longitudinal fasciculus (ILF), inferior fronto-occipital fasciculus (IFOF), and the anterior (ATR), superior (STR), and posterior (PTR) thalamocortical radiations. Spearman partial correlations within each group (adjusting for age) followed by FDR correction for multiple comparisons were used to investigate the functional-structural coupling of the attention networks. In the control group, within-network rsFC for the DAN was positively correlated with FA of the SLF I. Between-network connectivity of the DAN was significantly positively correlated with FA and negatively correlated with RD of the left ILF and FA of the bilateral IFOF. Between-network connectivity of the VAN was significantly positively correlated with FA of the left ILF and IFOF. Functional connectivity between the thalamus and VAN was significantly negatively associated with PTR IET2. No significant associations were observed in the CVI group. Overall, the results suggest that CVI may be associated with aberrant functional-structural coupling of the attention networks. The relationship with the previously observed selective attention deficits in individuals with CVI needs to be further investigated.

National Eye Institute

23.326 MODELING THE FORMATION OF EXTRASTRIATE PRIMATE VISUAL FIELD MAPS

Yujia Xie[1], Michael Arcaro[2], Nabil Imam[1]; [1]College of Computing, Georgia Institute of Technology, [2]Department of Psychology, University of Pennsylvania

Topographic maps of visual space are pervasive throughout the primate visual system. The molecular factors responsible for generating the precise topographic map of primary visual cortex, V1 are well studied. However, our understanding of the mechanisms governing postprimary cortical development remains limited. It is assumed that extrastriate maps self- organize around an established V1 map (Rosa 2002). Recent research, modeling the development of extrastriate maps as iterative propagation from the V1 map, have demonstrated that activity-based connections and wiring density limits are sufficient to generate a hierarchy of mirror-symmetric maps of visual space (Imam & Finlay 2020). Here, we extend this model by incorporating the specific cortical topology and retinotopic organization observed in primates. We measured the retinotopic organization of visual cortex in macaque monkeys using fMRI (Arcaro et al. 2017). For each subject, we segmented the grey matter and reconstructed the cortical surface from high-resolution anatomical MRIs. We then constructed a network model in which distances between nodes corresponded to the measured anatomical distances along the cortical surface. Connectivity within the model was established using an activity-based developmental program. As the program unfolded, nodes in V1 preferentially formed connections with nodes outside of V1 based on activity-dependent correlations and edge density limits within the network. Our preliminary results reveal that this activity-based modeling approach can produce a series of mirror-symmetric maps in the approximate locations corresponding to visual areas V2, V3, and V4 found in individual monkeys. Ongoing work aims to assess the impact of different distance measures (cortical vs. volumetric) and species-specific (human vs. macaque) factors related to cortical folding. Rather than requiring area-specific specifications, this work supports the view that a stereotypical layout of cortical areas unfolds within a given species through iterative self-organizing rules, requiring only a prespecification of the primary cortical map.

23.327 DEVELOPMENT OF FEATURE-BINDING IN INFANTS

Shuma Tsurumi[1,2], So Kanazawa[3], Masami Yamaguchi[2]; [1]Hokkaido University, [2]Chuo University, [3]Japan Women's University

Visual perception is established by the integration of multiple features. This mechanism benefits from the ongoing interplay between feedforward and feedback loops, yet this causal connection remains unclear. Researchers have explored the role of recurrent processing in feature integration by studying an illusion called 'misbinding,' wherein visual characteristics are erroneously merged, resulting in a perception distinct from the originally presented stimuli (Wu et al., 2004). Anatomical investigations have revealed that the neural pathways responsible for recurrent connections are underdeveloped in early infants. Therefore, there is a possibility that younger infants could potentially perceive the physically presented visual information that adults miss due to misbinding. We investigated this question by employing a familiarization/novelty preference procedure. In the familiarization phase, two sheets of dots, with one sheet moving upward and the other moving downward, were presented. Importantly, the dots within the right and left regions as well as those in the remaining area on both sheets were displayed in distinct colors, either red or green. After familiarization phase, two types of trials (coherence and segregation) were conducted during the testing phase. Coherence trials showed that dots of either color moved coherently in the same direction as in the familiarization phase. Segregation trials showed that dots in the right and left areas moved in opposite directions in the center. If infants perceived misbinding, infants would look at segregation longer because of its novel stimuli. In contrast, they would look at coherence longer if no misbinding occurred. We found that older infants exhibited misbinding, whereas younger infants showed no such illusions. This means that infants older than half a year perceive incorrectly integrated visual stimuli; however, infants younger than half a year perceive physically presented stimuli without failure of feature-binding. These results suggest that the development of feedback processing contributes to feature-binding.

This research was financially supported by Grant-in-Aid for JSPS Fellows (No.19J21422), Research Activity Start-up (No. 23K18965), Scientific Research (B) (No.19H01774) and Grant-in-Aid for Scientific Research on Innovative Areas (No.17H06343).

23.328 DEVELOPMENT OF CROWDING IN AMBLYOPIA DEPENDS ON HOW YOU MEASURE IT.

Sarah J Waugh[1] (s.j.waugh@hud.ac.uk), Louisa A Haine[2], Monika A Formankiewicz[2], Maria Fronius[3]; [1]University of Huddersfield, UK, [2]Anglia Ruskin University, UK, [3]Goethe University, Germany

Crowding is extensive in young children and strabismic amblyopia, so clinical assessment is important. Crowding magnitude (logMAR) is calculated from acuities, with closely-separated optotypes thought to enhance amblyopia screening. Test separations are fixed in optotype proportion (FOP), or in arcmin (FAM). We examined impact of test design on crowding development in normal and amblyopic vision. Crowding magnitude (logMAR) was assessed with close-separation FOP and FAM tests. Crowding magnitude (FOP) = "crowded" HOTV-acuity (0.2 optotype-width separation) – "uncrowded" HOTV-acuity (isolated). Crowding magnitude (FAM) = "crowded" Landolt C-acuity (~2.6' separation) – "uncrowded" Landolt C-acuity (≥35' separation). Participants for FOP test were N=44 amblyopes and 100 healthy controls: ≤8yr N=21 anisometropic amblyopes (AA), N=20 strabismic amblyopes (SA) and N=59 controls (C). For FAM test there were N=100 amblyopes and 76 controls: ≤8yr N=14 AAs, N=39 SAs and N=36 Cs. Crowding magnitude reduces with age in control eyes (≤8 yr) for FOP (slope±1SD of -0.02±0.01, $P<0.05$) and FAM (-0.05±0.01, $P<0.05$) tests. Patterns of development differ in amblyopia. With the FOP test in SA (≤8 yr), crowding magnitudes are parallel (slopes FE -0.01±0.01 and AE -0.01±0.02; magnitudes±1SE are 0.28±0.02 and 0.42±0.03 logMAR, respectively). In AAs, magnitudes are greater ($P<0.05$) in fellow (0.34±0.03 logMAR) than amblyopic (0.26±0.02 logMAR) eyes but this pattern reverses <5yrs (0.20±0.1 vs 0.34±0.05 logMAR). With the FAM test, in SAs, crowding reduces in fellow eyes (slope -0.06±0.01, $P<0.05$; 0.17±0.02 logMAR) but arrest-of-development appears for amblyopic eyes (slope=0.0±0.0; 0.31±0.02 logMAR). In AA, magnitudes are greater ($P<0.05$) in fellow (0.21±0.03 logMAR), than amblyopic (0.12±0.03 logMAR) eyes, but this pattern reverses >8yr (0.04±0.02 vs 0.09±0.03 logMAR). Development of crowding in strabismic and anisometropic amblyopia is different depending on crowded-acuity test used, due to crowding/masking contributions. A crowding-distance test with minimal/no masking may clarify development of crowding in normal and amblyopic vision.

URN020-01 from University of Huddersfield, ARU VC Studentship to Waugh for Haine. R01 EY027964-01A1 to Pelli. ERA-NET Neuron (BMBF01EW1603B) and Augenstern-e.V. to Fronius.

23.329 THE EFFECT OF LOW ACUITY IN INFANCY ON DEVELOPING VENTRAL VISUAL STREAM REPRESENTATIONS

Áine T. Dineen[1] (dineenai@tcd.ie), Cliona O'Doherty[1], Anna Truzzi[1], Anna Kravchenko[1], Lorijn Zaadnoordijk[1], Alex R. Wade[2], Graham King[1], Enna-Louise D'Arcy[1], Jessica White[1], Tamrin Holloway[1], Chiara Caldinelli[1], Adrienne Foran[3,4], Eleanor J. Molloy[1,5], Rhodri Cusack[1]; [1]Trinity College Dublin, [2]University of York, [3]The Rotunda Hospital, Dublin, [4]Children's Health Ireland at Temple Street, Dublin, [5]The Coombe Hospital, Dublin

The seemingly deficient blurry beginnings of infant visual experience have been hypothesised to have adaptive value. Initial low acuity could serve as an inductive bias strengthening more holistic representations, contributing to the robustness of the developed visual system. Computational models of visual recognition using deep neural networks (DNNs) have supported this in principle, showing that training DNNs initially with blurred images broadens their receptive fields and improves face recognition performance. However, whether low acuity actually affects developing representations in humans has not been tested. We therefore characterised representational geometry in the ventral visual stream of infants using functional MRI (fMRI), and compared it to DNNs trained with various levels of smoothing. fMRI was acquired in awake 2- and 9-month-old infants (N=134) and adults (N=24) as they viewed 36 looming pictures, comprising 3 examples of 12 categories. To model the effect of initial low-acuity, we trained DNNs (ResNet-50) on 1000 categories of objects (ImageNet), with six DNNs receiving images smoothed to different degrees (gaussian blur with sigma ranging from 0-6 pixels). Representational similarity analysis was used to compare the representational geometry of each layer of the DNNs with that of the early and late regions of the ventral visual stream. For both regions, the infant groups and adults shared a considerable portion of representational geometry with the DNNs (Spearman r~0.2-0.5). The degree of image smoothing during DNN training substantially modulated the correspondence while the degree of smoothing during DNN inference had little effect. At 2-months, representational geometry was more similar to DNNs trained with greater smoothing compared to older age groups, supporting the hypothesis that developing ventral visual representations are shaped by initial low acuity. More generally, we demonstrate that awake infant fMRI and computational modelling synergise to provide unique insights into the developmental origins of the visual system.

This research was supported by an ERC Advanced Grant 2017 FOUNDCOG 787981 and an Irish Research Council Government of Ireland Postgraduate Scholarship GOIPG/2023/2479. We wish to thank Sojo Joseph, radiographer at the Trinity College Institute of Neuroscience for his contribution to MRI acquisition.

Saturday, May 18, 8:30 am – 12:30 pm, Banyan Breezeway

Plasticity and Learning: Models, neural mechanisms

23.330 FEAR CONDITIONING PROMPTS SPARSER REPRESENTATIONS OF CONDITIONED THREAT IN PRIMARY VISUAL CORTEX

Lihan Cui[1] (lihancui@ufl.edu), Andreas Keil[2], Mingzhou Ding[1]; [1]J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, [2]Department of Psychology and NIMH Center for Emotion and Attention, University of Florida

Threat detection and valuation is an important function of the visual system. It has been proposed that repeated exposure to threatening stimuli alters sensory responses. Our previous study has found that neural representations of conditioned threat in the primary visual cortex become sparser with aversive learning. We examined this issue further in this study. Simultaneous fMRI and EEG data were recorded from 18 participants viewing the random appearance of two Gabor patches with the 45-degree Gabor patch (CS+) occasionally paired with a loud scream (US) (25% reinforcement rate) and the 135-degree Gabor patch (CS-) never paired with the US. Applying the MVPA decoding method to fMRI data in a sliding trial window fashion, we found that (1) decoding accuracy between CS+ and CS- was significantly higher than chance level in all visual areas throughout the conditioning session, (2) as aversive learning progressed, the number of voxels contributing to the representation of CS+ decreased in primary visual cortex but not in other visual areas, and (3) the number of voxels contributing to the representation of CS- remained the same. Analyzing the concomitantly recorded EEG data, we found that (1) the latency of the event-related potential evoked by CS+ became progressively shorter with aversive learning and (2) the latency of the event-related potential evoked by CS- became progressively longer. These results confirmed that the neural representation of conditioned threat became sparser in the primary visual cortex and suggested that the sparsified representation facilitated the detection and evaluation of threat.

23.331 ISOLATING THE KONIOCELLULAR CONTRIBUTION TO AVERSIVE LEARNING IN HUMAN VISUAL CORTEX

Katherine J. McCain[1] (mccain.k@ufl.edu), Nathan M. Petro[2], Andreas Keil[1]; [1]University of Florida, [2]Institute for Human Neuroscience, Boys Town National Research Hospital

Prior work using steady-state visual evoked potentials (ssVEPs) to measure visuocortical activity related to the koniocellular visual pathway has shown a unique pattern of enhanced ssVEP amplitude. Specifically, the ssVEP tends to be heightened when viewing tritan stimuli compared to luminance stimuli. The tritanopic approach (Cavanagh et al., 1992) used to isolate the koniocellular pathway via activation of S-cones is a robust technique enabling the comparison of tritanopic and luminance conditions. It remains unclear how characteristics of tritan stimuli, like conditioned aversive valence, modulate this effect. To address this question, a series of experiments explored the impact of aversive conditioning on visuocortical responses to tritan and luminance stimuli, using ssVEP frequency tagging. In experiments 1 and 2, participants (N=27) completed a differential classical conditioning task with habituation, acquisition, and extinction phases. Each trial included a pre-stimulus adaptor period, containing only a central fixation point over a uniform yellow background (tritan stimulus) or a uniform black background (luminance stimulus). Following the adaptor, a random dot kinematogram (RDK)—white dots superimposed on adaptor—flickered at driving frequencies of 14 Hz (N=13) and 17.14 Hz (N=14). Moreover, the RDK moved in one of two directions (135 or 225 degrees) that signaled the presence (CS+) or absence (CS-) of threat (92 dB white noise burst). It was observed that tritan stimuli (CS+ and CS-) drove stronger ssVEP signals than luminance stimuli. This main effect increased in the acquisition phase, consistent with arousal-based modulation of the koniocellular pathway. Furthermore, luminance and tritan CS+ activity showed the expected selective amplitude enhancement late in the viewing epoch. A third ongoing study aims to replicate these effects with oriented gratings using the same manipulations. Results are consistent with findings showing heightened ssVEP for tritan stimuli and with theoretical notions emphasizing the role of the koniocellular pathway in processing motivational relevance.

Research was supported by NIH grant R01MH125615 to Dr. Andreas Keil

23.332 THE IMPLICIT PENALTY EFFECT: IMPLICIT CUE VALIDITY REGULATES PERFORMANCE IN AN EXOGENOUS CUEING PARADIGM.

Akira Sarodo[1] (chelsea3636@akane.waseda.jp), Katsumi Watanabe[1], Shao-Min (Sean) Hung[1,2]; [1]Faculty of Science and Engineering, Waseda University, [2]Waseda Institute for Advanced Study, Waseda University

Numerous evidence suggests that the human visual system both implicitly and explicitly extracts statistical regularity. However, the interaction between the implicit extraction and explicit knowledge of the statistical regularity remains unclear. In the present study, we adopted an exogenous cueing paradigm and manipulated cue-target congruency across the blocks (50%, 60%, 70%, 80%, 90%, 100%) to examine whether the explicit knowledge regarding the cue-target congruency affects the implicit cueing effect. Participants were asked to perform an orientation discrimination task on target. In Experiment 1, we increased the cue-target congruency from 50% to 100% and found that reaction time to the invalid cued target increased as the cue-target congruency increased, whereas the congruency had no impact on the reaction time to the valid cued targets, showing an incongruency penalty. This incongruency effect was observed even when the congruency was at chance (50%) and regardless of the explicit awareness of the cue-target congruency. Experiment 2 replicated this incongruency effect with decreasing cue-target congruency from 100% to 50% across the blocks. Two follow-up experiments with a fixed congruency (70%) demonstrated that explicit knowledge regarding the cue-target congruency from either an explicit prime or self-estimation did not alter the size of the penalty effect. That is, the penalty effect tracked the actual congruency rather than the perceived contingency. These results further showed that the explicit knowledge played only a trivial role in the incongruency penalty effect. This series of psychophysical experiments implies that the human visual system tracks the statistical regularity to optimise behaviour even without explicit awareness of the regularity.

23.333 INVESTIGATING THE REPRESENTATIONAL TRANSFORMATIONS UNDERLYING THE LEARNING OF EXCEPTIONS IN VISUAL CATEGORIES

Yongzhen Xie[1] (yongzhen.xie@mail.utoronto.ca), Emily Q. Wang[1], Yao Chen[1], Michael L. Mack[1]; [1]Department of Psychology, University of Toronto

Categories in the real world often contain perceptually incoherent objects, such as structural exceptions that resemble members of competing categories and oddball exceptions that encompass features distinct from any known categories. Prominent theories suggest that learning exceptions depends on flexible transformations of category representations, yet evidence of such representational dynamics in the brain is limited. Here, we had participants learn competing visual categories that included both structural and oddball exceptions while fMRI data was collected before, in the middle of, and after category learning. Representational similarity analysis of neural patterns for category stimuli at each learning stage revealed that exception learning induced unique transformations in object representations in distinct brain regions. Specifically, the introduction of exceptions led to an increase in feature-specific information in visual cortex representations. In contrast, representations in the prefrontal cortex exhibited an increase in prototype information consistent with coding for category regularities. Notably, subfields of the hippocampal formation also showed distinct transformations—feature-specific information increased in dentate gyrus representations and decreased in CA1 representations. These results align with the dentate gyrus' theorized role in constructing item-specific representations and CA1's role in generalization across related experiences. Moreover, we found that exception learning induced distinct representational transformations for structural and oddball exceptions. Particularly, within the representational spaces of the prefrontal and temporal cortices, structural exceptions became uniquely more differentiated from regular category members through learning. This finding aligns with the expectation that learning structural exceptions relies on distinguishing them from perceptually confusable items in the competing category via differentiation. Altogether, our results demonstrate that object representations can be flexibly and selectively transformed across the brain to support the learning of category regularities and their exceptions.

NSERC Postgraduate Scholarship (Doctoral Program) to Yongzhen Xie; Future Leaders in Canadian Brain Research Grant to Michael L. Mack; NSERC Discovery Grant to Michael L. Mack

23.334 A HIGH FIDELITY QUANTIFICATION OF THE TIME COURSE OF LEARNING CONSOLIDATION FROM A MASSIVE VISUAL SEARCH DATASET

Patrick Cox[1] (pac323@lehigh.edu), Chloe Callahan- Flintoft[2], Michelle Kramer[3], Stephen Mitroff[4], Dwight Kravitz[4]; [1]Lehigh University, [2]United States Army Research Laboratory, [3]Transportation Security Administration, [4]The George Washington University

Visual search behavior adapts to the statistics of local contexts, as shown in cognitive effects like priming (e.g., Maljkovic & Nakayama, 1994), statistical learning (e.g., Geng & Behrmann, 2005), contextual cueing (e.g., Chun & Jiang, 1998), and selection history (e.g., Awh et al., 2012). Previous work has described a general evidence accumulation function that may underlie such learned adaptations in behavior across a range of features (Kramer et al., 2022), but questions remain about the time course of the consolidation of this learning. Here we used a massive dataset (~15.7 million users, ~3.8 billion trials) of human behavioral data from a mobile app (Airport Scanner, Kedlin Co.) to quantify the consolidation of learning in a visual search task over a range of time delays, and to describe the impact of sleep during the consolidation period. The size of this dataset allowed for an extremely high fidelity characterization of the time course of consolidation from seconds to days with a high degree of temporal precision. Linear modeling of the effects of prior experience (proportion of trials containing a target), time delay (amount of time since the last block of search) and their interaction showed significant main effects of prior experience and time delay, as well as a complex nonlinear interaction for all of our dependent variables (hit rate, target present correct response time, correct rejection rate, target absent correct response time). A secondary analysis on the effect of sleep on visual search performance revealed main effects of sleep on search accuracy but no interaction of sleep with accumulation of evidence about target prevalence. These characterizations of the temporal dynamics of consolidation behavior provide much needed constraints on hypotheses about the myriad of neural mechanisms underlying learning and their characteristic time courses.

US Army Research Laboratory Cooperative Agreements #W911NF-21-2-0179, #W911NF-23-2-0210, & #W911NF-23-2-0097; IC Postdoc Research Fellowship Program administered by ORISE through an interagency agreement between the DOE and the ODNI

23.335 HIERARCHICAL BAYESIAN AUGMENTED HEBBIAN REWEIGHTING MODEL OF PERCEPTUAL LEARNING

Zhong-Lin Lu[1,2] (zhonglin@nyu.edu), Shanglin Yang[1], Barbara Dosher[3]; [1]NYU Shanghai, [2]New York University, [3]University of California, Irvine

The Augmented Hebbian Reweighting Model (AHRM; Petrov et al., 2005) has successfully modeled various phenomena in perceptual learning. Fitting the AHRM to data presents a significant challenge because, as a sequential learning model, it must be simulated to generate performance predictions with sequential trial-by-trial updates, and estimation of the AHRM parameters is generally done using hierarchical grid-search methods. In this study, we introduce three modeling technologies to facilitate AHRM fitting: A Hierarchical Bayesian AHRM (HBAHRM) that incorporates population, subject, and test levels to estimate the joint posterior hyperparameter and parameter distribution while considering covariance within and between subjects; vectorization techniques with PyTensor to drastically speed up simulations involving multi-dimensional arrays; and pre-computed the likelihood function of the AHRM. We fit the data from Petrov et al. (2005), which investigated perceptual learning in an orientation identification task with 13 subjects in two external noise orientation contexts. We found that the HBAHRM provided significantly better fits to the data than the Bayesian Inference Procedure that inferred AHRM parameters for each subject independently. At the population level, the HBAHRM generated fits with an Rsq of 0.852 and RMSE of 0.031 (in d' units). In a simulation study, we found that the HBAHRM exhibited excellent parameter recovery and fit the simulated data with an Rsq of 0.982 and RMSE of 0.010 (d' units). Additionally, the HBAHRM made excellent predictions of the performance of a new simulated observer with no data, 300 trials (all in one context), and 2700 trials (300 in one and 2400 in the other context) of data. The HBAHRM and the new modeling techniques can be readily applied to analyze data from various perceptual learning experiments and provide predictions of performance of new observers with no or limited data.

The research was supported by NEI grant EY017491.

23.336 LEARNING AND TRANSFER BETWEEN DIFFERENT EXTERNAL NOISE LEVELS IN ORIENTATION PERCEPTUAL LEARNING

Jiajuan Liu[1] (jiajuanl@gmail.com), Zhong-Lin Lu[2], Barbara Dosher[3]; [1]University of California, Irvine, [2]New York University, [3]University of California, Irvine

Orientation identification tasks often show improvement with practice, but training with zero external noise may be more efficient: multi-session training in zero external noise showed essentially full transfer to high external noise, while training with high external noise showed limited transfer to low external noise (Dosher & Lu, 2005). Here, we examined how training in zero or in high external noise transferred to multiple external noise levels as assessed in each session. In four sessions, observers were trained with feedback in a peripheral (5.4 deg) two-alternative orientation task (-55+/-10 deg) in either zero or high external noise (in two separate groups of n=7), and were also assessed at four external-noise levels (0, 0.8, 0.17, 0.33) without feedback at the beginning (pretest) and end (posttest) of each session. Contrast thresholds tracked a low accuracy level (65%) to minimize learning without feedback (Liu et al, 2012). Transfer to different retinal locations was tested in the fifth session. Results: 1) The contrast threshold improved in both training groups under their respective training conditions, with faster learning observed in the zero external noise group (log-log threshold vs block slope: -0.28 vs -0.07); 2) training improved performance in all external noise levels in both groups; 3) within-session improvement from pretest to posttest was larger in the first few sessions when most learning occurred and overnight consolidation (from previous-day posttest to current day pretest) was also most apparent early in training; 4) there was substantial location transfer. The integrated reweighting theory (IRT, Dosher et al, 2013) successfully captured these results with the same learning rate, because external noise in the stimulus adds noise to and perturbs learned weights from stimulus representations (encoding) to decision (decoding)(Lu, et al, 2010). Empirically, assessing multiple external noise levels throughout training revealed transfer across external noise conditions throughout the training process.

Supported by the National Eye Institute Grant # EY–17491.

23.337 IDENTIFYING COMPONENT PROCESSES IN PERCEPTUAL LEARNING WITH NON-PARAMETRIC BAYESIAN MODELING OF THE LEARNING CURVE IN A YES-NO TASK WITH METHOD OF CONSTANT STIMULI

Yukai Zhao[1] (zhaoyukai@nyu.edu), Jia Yang[2], Fang-Fang Yan[2], Chang-Bing Huang[2], Barbara Anne Dosher[3], Zhong-Lin Lu[1,4]; [1]New York University, [2]Chinese Academy of Sciences, [3]University of California, Irvine, [4]New York University Shanghai

Perceptual learning is a multifaceted process that may involve general learning, between-session forgetting or consolidation, and within-session rapid relearning and adaptation (Yang et al., 2022). The traditional learning curve, often derived from aggregated data in blocks or sessions comprising tens or hundreds of trials in most perceptual learning studies, may have obscured certain component processes. In a previous study, we developed three non-parametric inference procedures to estimate fine-grained contrast threshold learning curves in a Gabor orientation identification task, measured with the staircase procedure. In this work, we introduce a non-parametric Bayesian inference procedure to estimate the posterior distribution of the block d' learning curve in Yes-No tasks measured with the method of constant stimuli, incorporating varying block sizes. The model assumes the decision criterion as a constant likelihood ratio across all blocks for each subject. We applied the method with three block sizes (10, 35, and 100 trials/block) to a global motion same-different judgement task conducted over 3500 trials across five sessions (Yang et al., 2022). The goodness of fit to the data increased with the temporal resolution of the analysis. Model comparisons, based on the Bayesian Predictive Information Criterion (BPIC), identified the 10 trials/block model as the best fit. When fitting a multi-component generative model of perceptual learning (Zhao et al., submitted) to the average d' learning curves at the group level, we uncovered general learning, between-session forgetting and within-session rapid relearning with 10 and 35 trials/block. In contrast, the original study with 100 trials/block only identified general learning and within-session rapid relearning. The non-parametric Bayesian inference procedure offers a versatile framework for high-temporal resolution assessment of the component processes in perceptual learning across diverse tasks and testing paradigms.

National Eye Institute (EY017490)

Saturday, May 18, 8:30 am – 12:30 pm, Banyan Breezeway

Face and Body Perception: Bodies

23.338 LATERALIZED PERCEPTION OF STATIC AND DYNAMIC SOCIAL INTERACTIONS IN LEFT AND RIGHT VISUAL CORTEX

Violette Munin[1,2] (violettemunin@hotmail.fr), Etienne Abassi[1,2], Liuba Papeo[1,2]; [1]Institute of Cognitive Sciences Marc Jeannerod, CNRS UMR 5229, [2]University Claude Bernard Lyon 1

The visual system has extensively been studied in relation to its function in object and action recognition. Latest research findings show that it also plays a specific role in representing social interactions (an agent acting towards another), hosting specialized neural structures for this task. In current studies, static representations of social interactions (two people facing towards vs. away from one another) and dynamic representations (video-clips of interacting/non-interacting people) yielded stronger activity for facing/interacting in left extrastriate body area (EBA) and right posterior superior temporal cortex (pSTS), respectively. We asked whether different localization and lateralization of the effect might depend on the stimuli being static or dynamic. We reanalyzed two fMRI datasets, where the same 15 female and male adults saw video-clips and static images of facing – seemingly interacting– and non-facing people. First, whole-brain analysis replicated higher activity for facing (vs. non-facing) bodies in visual cortex, which was overall stronger for dynamic stimuli. For both static and dynamic stimuli, the effect was stronger in left areas. Next, we individually localized the body-selective EBA, the motion-selective MT/V5, the biological motion-selective pSTS, the so-called social-interaction pSTS (SI-pSTS)¬ –and other (control) visual areas. Region-of-interest analysis showed that the facing>non-facing effect in EBA occurred for both static and dynamic stimuli, and was stronger in the left. MT/V5 and pSTS showed the same left-lateralized effect but only for dynamic stimuli. The SI-pSTS showed a third response pattern with a selective, bilateral effect for dynamic stimuli. Challenging the common view that allocates social stuff to right visual areas, these results support a prominent role of left regions in social-interaction processing. Moreover, they suggest that within the hub for social processing in pSTS, there are two different regions, biological-motion pSTS and SI-pSTS, with different response profiles, and presumably different functions in the representation of socially related/interacting agents.

This research was funded by the European Research Council Starting Grant to Liuba Papeo (Grant number: THEMPO-758473).

23.339 MACAQUES SHOW AN UNCANNY VALLEY IN BODY PERCEPTION

Lucas Martini[1] (lucas.martini@uni-tuebingen.de), Anna Bognár[2], Rufin Vogels[2], Martin Giese[1]; [1]University of Tuebingen, [2]KU Leuven

Highly realistic virtual stimuli from avatars have been essential for the investigation of face processing and will likely be similarly important for the study of body perception. In face avatars, humans as well as macaques show an uncanny valley, i.e. when such stimuli are made increasingly realistic the emotional preference for them does not linearly increase with realism, but is drastically reduced for almost realistic presentations. An uncanny valley has also been found for humans for the observation of robots. Whether such an uncanny valley exists for body stimuli in non-human primates is not known. METHODS: We developed a novel computer animation pipeline that allows the animation of a commercial monkey body avatar from markerless multi-camera motion capture recordings of real monkeys. Based on a minimal number of hand-labeled keyframes, our method generates highly realistic dynamic body animations. In a behavioral experiment with eight male rhesus monkeys, we validated our macaque model by monitoring the subjects' eye movements on renderings of a submissive turning and a neutral walking movement. These stimuli were shown from different viewpoints, incorporated realistic background, and degraded variants of the most realistic avatar. RESULTS: The analysis of the eye movement data revealed an uncanny valley effect in coverage of the stimulus, number of image fixations, and mean fixation duration ($p=0.002$, $p=0.02$, and $p<0.001$), where the gazing behavior for the non-degraded avatar did not significantly deviate from the real video. CONCLUSION: Monkeys exhibit an uncanny valley for the observation of monkey bodies. This shows the universality of this effect across different social stimuli and primate species.

European Research Council (2019-SyG-RELEVANCE-856495). International Max Planck Research School for Intelligent Systems (IMPRS-IS).

23.340 PREDICTIVE FEATURES IN DEEP NEURAL NETWORK MODELS OF MACAQUE BODY PATCH SELECTIVITY

Alexander Lappe[1,2], Anna Bognár[3], Ghazaleh Ghamkhari Nejad[3], Rajani Raman[3], Albert Mukovskiy[1], Lucas Martini[1,2], Rufin Vogels[3], Martin A. Giese[1]; [1]CIN & HIH, University Clinics Tübingen, [2]International Max Planck Research School for Intelligent Systems, [3]KU Leuven

INTRODUCTION: Previous work has shown that neurons from body patches in macaque superior temporal sulcus (STS) respond selectively to images of bodies. However, the visual features leading to this body selectivity remain unclear. METHODS: We conducted experiments using 720 stimuli presenting a monkey avatar in various poses and viewpoints. Spiking activity was recorded from mid-STS (MSB) and anterior-STS (ASB) body patches, previously identified using fMRI. To identify visual features driving the neural responses, we used a model with a deep network as frontend and a linear readout model that was fitted to predict the neuron activities. Computing the gradients of the outputs backwards along the neural network, we identified the image regions that were most influential for the model neuron output. Since previous work suggests that neurons from this area also respond to some extent to images of objects, we used a similar approach to visualize object parts eliciting responses from the model neurons. Based on an object dataset, we identified the shapes that activate each model unit maximally. Computing and combining the pixel-wise gradients of model activations from object and body processing, we were able to identify common visual features driving neural activity in the model. RESULTS: Linear models fit the data well, with mean noise-corrected correlations with neural data of 0.8 in ASB and 0.94 in MSB. Gradient analysis on the body stimuli did not reveal clear preferences of certain body parts and were difficult to interpret visually. However, the joint gradients between objects and bodies traced visually similar features in both images. CONCLUSION: Deep neural networks model STS data well, even though for all tested models, explained variance was substantially lower in the more anterior region. Further work will test if the features that the deep network relies on are also used by body patch neurons.

ERC 2019-SyG-RELEVANCE-856495; SSTeP-KiZ BMG: ZMWI1-2520DAT700

23.341 RELATIVE DEPTH DISCRIMINATION IN NATURAL IMAGES OF PAIRED HUMAN BODY JOINTS

Jiaqi Liu[1] (liu00687@umn.edu), Daniel Kersten[1]; [1]University of Minnesota Twin Cities

Humans have the ability to perceive three-dimensional depth from a two-dimensional image plane. An illustrative example is the ability of humans to recognize body pose and extrapolate the three-dimensional spatial arrangement of joints given a human body image. While past studies have indicated that the internal representation of the human figure can sometimes impose constraints on the depth discrimination of static stimuli, the precise integration of local and structural information among body parts for inferring depth remains unclear. Here we investigated human ability to identify relative depth between pairs of body parts given limited spatial context from natural images. In the experiment, 20 observers viewed a series of pairs of body parts, each recognizable above chance, and displayed through a circular aperture. Observers were then asked to identify which part was closer to them. The manipulation of structural information involved varying the spatial relationship (retained/original position and disrupted/side by side) and different types of body pairs (same side and cross side). Each condition comprised 100 trials, with images sourced from the Leeds Sports Dataset. The performance of human depth perception was evaluated against the ground truth established by the Unite the People dataset. We found that retained spatial relations significantly enhance the discrimination of relative depth between body parts compared to disrupted spatial relations. Furthermore, the accuracy in depth discrimination was higher in elbow-elbow pairs compared to elbow-wrist pairs. Additionally, an investigation into how Euclidean distance between parts could influence depth discrimination revealed that, in contrast to elbow-elbow pairs, a closer distance between the wrist and elbow resulted in heightened accuracy, suggesting a potential grouping mechanism between adjacent parts. Our study underscores that humans efficiently employ both structural knowledge and low-/mid-level grouping cues to infer depth information given limited spatial context.

This research was funded in part by NIH grant 5R01EY029700 Towards a Compositional Generative Model of Human Vision.

23.342 COMPARING HUMAN FACE AND BODY RECOGNITION AT VARIOUS DISTANCE AND ROTATION VIEWING CONDITIONS.

Matthew Groth[1] (mjgroth@mit.edu), Michal Fux[1], Hojin Jang[1], Suayb S. Arslan[1], Walt Dixon[2], Joydeep Munshi[2], Pawan Sinha[1]; [1]Massachusettes Institute of Technology, [2]GE

Deep networks trained on large face datasets have achieved impressive performance on recognition tasks. However, as we reported in last year's meeting (Fux et al., 2023), humans still outperform DNNs when viewing conditions are challenging as with large distances, non-frontal regard, and atmospheric turbulence. In the current study, we investigate recognition performance of humans and deep networks with images of whole bodies. This task is akin to the 'person re-identification' challenge of great interest to the machine vision community. We worked with a large database of images acquired at a variety of distances and from multiple yaw/pitch angles. We ran an online behavioral study in which participants were asked to rank whole-body images of people from most to least likely to be the same identity as the person in the query image. The query image depicts individuals in three conditions: Whole body, head occluded, and body occluded. Distance to the camera ranged from 10m to 500m. The results enable an analysis of the relative contribution of the head and body to recognition as a function of viewing distance. A comparative analysis of humans against a whole-body trained DNN establishes a clear superiority of human performance across all tested distances. Moreover, this comparison reveals differences in the relative importance of head and body regions, with humans deriving significant identity information from the head region, unlike the DNN. These preliminary results indicate potential divergent strategies employed by humans and DNNs, offering insights into distant person identification and implications for the design of future machine models.

This research is supported by ODNI, IARPA. The views are of the authors and shouldn't be interpreted as representing official policies of ODNI, IARPA, or the U.S. Gov., which is authorized to reproduce & distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

23.343 PERCEPTION OF HUMAN ACTIONS ON THE FOUR FUNDAMENTAL DIMENSIONS OF FORMIDABLENESS, FRIENDLINESS, INTENTIONALITY AND ABDUCTION

Nick E Barraclough[1] (nick.barraclough@york.ac.uk), Laura C Vinton[1], Alessia Vlasceanu[1], Gabriel Mackie[1], Stephan de la Rosa[2], Catherine Preston[1], Stephen P Tipper[1]; [1]University of York, [2]IU University of Applied Sciences, Erfurt, Germany

We evaluate the actions of other individuals based upon their movements and postures that reveal critical information to guide our decision making and behavioural responses. These signals convey a range of information about the actor, including their goals, intentions and internal mental states. In order to understand the structure of the conceptual space underlying our perception of actions we assessed which action qualities were fundamental, and how individuals perceived actions on these dimensions. We recorded 240 different actions using motion-capture and used these data to animate a volumetric avatar. 230 participants then viewed these actions and rated the extent to which each action demonstrated 23 different action characteristics (e.g., avoiding-approaching, weak-powerful etc.).

Exploratory Factor Analysis showed that action space was four-dimensional, with the factors (dimensions) of: friendly-unfriendly, formidable-feeble, intentional-accidental, and abduction-adduction. The first two factors of friendliness and formidableness explained approximately 22% of the variance each, compared to intentionality and abduction, which each explained approximately 7–8% of the variance. We developed an action morphing method that used source actions located at different positions within action space to generate novel actions that lay along the different action dimensions. Morphed actions that varied along each action space dimension were rated as varying monotonically along that dimension. Action discrimination along the action space dimensions was tested using adaptive 2-AFC procedures. We found considerable variation in perceptual thresholds, varying by up to ~1100%; this interindividual variation in action perception could not be explained by autistic traits. Together, our results show that actions are perceived on 4 fundamental quality dimensions. Friendliness and Formidableness appear similar to the principal factors underlying our evaluation of facial traits and emotions, while Intentionality and Abduction appear unique to actions. The isolation of these dimensions allows us to investigate the mechanisms underpinning our conceptual representation of human actions.

23.344 PERSONALITY TRAIT INFERENCES FROM THREE-DIMENSIONAL BODIES IN AMERICAN VERSUS CHINESE INDIVIDUALS

Ying Hu[1,2] (huy@psych.ac.cn), Natalie Kate Bendiksen[3], Xiaolan Fu[1,2], Alice J. O'Toole[3]; [1]State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China, [2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, 100049, China, [3]The University of Texas at Dallas, School of Behavioral and Brain Sciences

A person's body shape spontaneously elicits first impressions of personality traits (e.g., lazy, extraverted). Previous research on Americans identified a trait space structured by valence (positive vs. negative) and agency (active vs. passive) (Hu et al., 2018). Here, we explore the generalizability of this structure for Chinese participants. American (41 females, 39 males) and Chinese (40 females, 40 males) participants rated 140 three-dimensional bodies across 30 personality traits, categorized by the Big Five domains (Extraversion, Conscientiousness, Openness, Agreeableness, Neuroticism). For Americans, dimensionality reduction analyses of trait ratings replicated Hu et al. (2018)'s valence (explained variance: 60.73% for female bodies, 53.71% for male bodies) and agency (female: 8.23%, male: 8.82%) dimensions. The trait space for Chinese participants, however, exhibited dimensions of valence (female: 61.34%, male: 59.26%) and Extraversion (female: 10.32%, male: 9.01%). Multiple linear regression with cross validation showed that body shape parameters predicted the pattern of trait inferences, as well as specific traits for both cultures. Conscientiousness and Extraversion were best predicted for Americans; Conscientiousness, Openness, and Neuroticism were best predicted for Chinese. Cross-culture correlations on same bodies showed large-scale correlations (ρ ranges: 0.24 to 0.81) that were higher for Conscientiousness and Openness than for Extraversion. Americans associated skinnier bodies with both Conscientiousness and Extraversion, whereas Chinese individuals associated skinnier bodies with Conscientiousness and sturdier body shapes with Extraversion. These findings underscore the cultural similarities in valence/Conscientiousness inferences and the cultural differences in agency and Extraversion inferences. Valence/Conscientiousness generalized across culture, suggesting its adaptive function for selecting healthy and successful partners. Conversely, agency and Extraversion varied culturally, showing social motivation in selecting higher agency and more extraverted partners for social exchange in Americans compared to Chinese individuals. This study pioneers the exploration of cultural influences on body trait impressions, contributing to the development of a theory with broader global applicability.

This research was supported by the National Natural Science Foundation of China (NSFC) (No.E3JJ032520 and No.62276259). The author Y.H. was also supported by the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (No. E1CX4625CX) while preparing this manuscript.

SATURDAY, MAY 18, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Face and Body Perception: Neural mechanisms of social cognition

23.345 IDIOSYNCRATIC FACIAL MOTIONS: UNCOVERING IDENTITY INFORMATION IN FACIAL MOVEMENTS THROUGH A LANDMARK-BASED ANALYSIS

Hilal Nizamoğlu[1,2], Katharina Dobs[1,2]; [1]Justus Liebig University in Giessen, Germany, [2]Center for Mind, Brain and Behavior (CMBB), University of Marburg and Justus Liebig University Giessen

Previous studies on dynamic faces have shown that both rigid and non-rigid facial movements contribute to identity recognition, and that the identity-specific information inherent in these motions varies depending on the type of facial expression. Therefore, it becomes clear that individuals exhibit distinct idiosyncratic patterns in their facial movements, which can serve as cues towards their identity. However, the specific features of facial movements that contribute to this uniqueness remain unclear. Here, we employed machine learning techniques to measure and quantify motion information in facial expressions, using a dataset of six basic emotional facial expressions (anger, disgust, fear, joy, sad, surprise) performed by 12 German and 12 Turkish lay actors. An automated facial landmark detection tool was applied to measure the positional changes of landmarks at the peak of each expression relative to a neutral baseline. We then trained a Linear Discriminant Analysis (LDA) model with these landmark shifts to classify the emotional expressions. This first LDA model was able to classify the type of emotional expression (Accuracy: 44%, $p<0.001$), independent of the actor's identity. More strikingly, another LDA model, trained to classify the identities of the 24 actors across different expressions, successfully predicted their identity (Accuracy: 45%, $p<0.001$). Furthermore, the landmark positional changes provided useful information for classifying actors' gender (Accuracy: 59%, $p<0.01$) and country-of-origin (Accuracy: 71%, $p<0.001$), supporting previous studies on cultural and gender-based variations in facial expressions. In conclusion, our study shows the richness of information embedded in facial motion features, extending beyond emotional expression to contain aspects of the actor's identity, gender, and cultural background. This landmark-based approach emerges as a promising tool to unravel the nuances of idiosyncrasies in facial movements, offering valuable insights into the intricate interplay of expression, identity, and cultural factors.

This work was supported by DFG, Germany SFB/TRR 135 (grant number 222641018) TP C9 and S; and by the Research Cluster "The Adaptive Mind", funded by the Hessian Ministry for Higher Education, Research Science and the Arts. We also thank Prof. Dr. Benjamin Straube for the video dataset.

23.346 LEARNED TRUSTWORTHINESS AND FACE TRUSTWORTHINESS DO NOT HAVE THE SAME INFLUENCE ON IMPLICIT RESPONSES TO FACES MEASURED VIA FAST PERIODIC VISUAL STIMULATION (FPVS)

Sara Verosky[1] (sverosky@oberlin.edu), Hannah Nishiura[1], Lucie O'Brien[1], Huanting Liu[1], Nada Aggadi[1]; [1]Oberlin College

Although people readily form impressions based on facial appearance, they also form impressions based on other sources of information, such as information about a person's past actions. Past research has demonstrated that it is possible to detect implicit responses to face trustworthiness using fast periodic visual stimulation (FPVS) in conjunction with electroencephalography (EEG). Because people readily retrieve affective associations with faces, the current study investigated whether learned trustworthiness would yield similar responses to face trustworthiness as measured via FPVS. In support of this possibility, FPVS has been found to be sensitive to both face familiarity and to semantic categorization. Over the course of a week, participants learned to associate faces with negative or positive behaviors. Later, while EEG was recorded, participants completed three separate FPVS tasks. In each of these tasks, participants viewed oddball sequences of faces where a single base face was presented repeatedly at a rate of 6 Hz and oddball faces with different identities were presented every fifth face (6 Hz/5 = 1.2 Hz). Reproducing and extending prior findings, we observed a robust response at 1.2 Hz and its harmonics to faces with learned associations as compared to novel faces over right occipitotemporal cortex, as well as over a number of other electrode sites, including fronto-central sites. In addition, also reproducing earlier findings, we observed a stronger response at 1.2 Hz and its harmonics for sequences with less trustworthy-looking versus trustworthy-looking oddball faces over right occipitotemporal cortex and a number of other sites. However, contrary to our predictions, we did not observe a significant influence of learned trustworthiness on the oddball response. These data indicate that impressions based on learning are treated differently than impressions based on appearance, and they raise questions about the types of design and stimuli that yield responses that are measurable via FPVS.

23.347 FROM POINT LIGHT DISPLAYS TO RICH SOCIAL NARRATIVES: NEURAL REPRESENTATIONS OF VISUAL SOCIAL PROCESSING IN THE SUPERIOR TEMPORAL SULCUS

Hannah Small[1] (hsmall2@jhu.edu), Haemy Lee Masson[2], Ericka Wodka[3,4], Stewart H. Mostofsky[3,4], Leyla Isik[1]; [1]Johns Hopkins University, [2]Durham University, [3]Kennedy Krieger Institute, [4]The Johns Hopkins University School of Medicine

Social perception is used ubiquitously in daily life. Prior work has revealed that a region in the right posterior superior temporal sulcus (STS) selectively supports social interaction perception in controlled stimuli such as point light displays. In contrast, both the left and right STS have been shown to support social interaction perception in naturalistic stimuli. However, previous work did not account for the rich verbal signals that occur simultaneously with social visual signals in natural settings and did not compare with controlled experiments in individual subjects. Do social interaction and language selectivity generalize across simple, controlled experiments and a more naturalistic movie stimulus? In an fMRI experiment, 12 participants completed controlled tasks previously shown to identify social interaction and language selective voxels in the STS (viewing interacting versus independent point light figures and listening to spoken versus scrambled language). Participants also viewed a 45 minute naturalistic movie. We fit an voxel-wise encoding model that included low- and mid-level visual and auditory features, as well as higher-level social and language features, including the presence of a social interaction and language model embeddings of the spoken language in the movie. Despite the drastically different nature of our controlled versus movie experiments, voxel-wise preference mapping and variance partitioning revealed spatial and functional overlap between the movie and controlled experiments for both social interaction and language-selective voxels. However, there were some differences between the controlled experiments and the natural movie. The movie stimuli enabled a richer characterization of voxel-wise processing and also elicited stronger bilateral social responses in the pSTS, even when accounting for spoken language. Overall, these results show that controlled and naturalistic stimuli recruit similar areas for social processing, but naturalistic stimuli can give a richer understanding of the neural underpinnings of simultaneous visual and verbal social processing in real-word settings.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE2139757. This work is also supported by grant R21MH129899.

23.348 INVESTIGATING THE NEURAL COMPUTATIONS UNDERLYING VISUAL SOCIAL INFERENCE WITH GRAPH NEURAL NETWORK AND INVERSE PLANNING MODELS

Manasi Malik[1] (mmalik16@jhu.edu), Tianmin Shu[1], Josh Tenenbaum[2], Leyla Isik[1]; [1]Johns Hopkins University, [2]Massachusetts Institute of Technology

Recognizing social interactions in visual scenes is a crucial human ability, however, the neural computations that enable this remain undetermined. Prior work has shown that two distinct computational models, a bottom-up graph neural network (SocialGNN) based solely on visual information and a generative inverse planning model (SIMPLE) based on mental state inference, both uniquely explain significant variance in human judgements. Here, we compare both models to neural data to understand how the brain combines these two types of computations for social scene understanding. We collected fMRI data from adults while they watched videos of two animated agents interacting. We compared neural representations with human behavior judgements and each computational model of social interaction recognition. Preliminary analysis using whole-brain searchlight RSA showed a significant correlation between neural RDMs and behavioral RDM in the visual cortex, lateral occipital temporal cortex and the superior temporal sulcus (STS). With the computational models, we find that SocialGNN exhibited a significantly higher correlation than SIMPLE in more posterior and dorsal regions including the lateral occipital cortex (LOTC) and posterior STS - regions previously implicated in social perception. On the other hand, SIMPLE demonstrated significantly higher correlation than SocialGNN in more anterior regions, including anterior STS and medial prefrontal cortex (mPFC). Further, both SocialGNN and SIMPLE explain significant variance in posterior and mid regions of the STS suggesting these regions as a potential site of integration of social perception and mental state inference. This work provides a novel framework for testing computational theories of social perception and cognition, as well as preliminary evidence for how the brain combines bottom-up vision and mental state inference during social scene understanding.

This work was funded by NIMH R01MH132826 awarded to L.I.

23.349 SEPARATE NEURAL REPRESENTATIONS FOR PHYSICAL AND COMMUNICATIVE SOCIAL INTERACTIONS ALONG THE LATERAL VISUAL PATHWAY: EVIDENCE FROM DATA-DRIVEN VOXEL DECOMPOSITION

Yuanfang Zhao[1] (distancejay@gmail.com), Emalie McMahon[1], Leyla Isik[1]; [1]Johns Hopkins University

Recognizing social interactions is remarkable for both its adaptive significance and visual complexity. Previous studies have suggested that the lateral visual cortex and superior temporal sulcus (STS) are involved in social interaction perception. However, it has been difficult to disentangle neural responses of different types of social interaction with hypothesis-driven approaches, due to challenges with feature labeling, sampling and experimenter bias. To overcome these issues, we employ a data-driven voxel decomposition technique (i.e., non-negative matrix factorization) to a largescale naturalistic fMRI dataset of participants freely viewing two hundred 3-second video clips. These naturalistic videos depicted two individuals engaging in various social and nonsocial activities sampled from everyday scenes. Our analysis of the lateral visual cortex and STS revealed two components with distinct functional profiles that were shared across all participants. We used extensive dataset annotations and free-response captions to characterize these components. The first component responds strongly to videos that are rated as highly communicative in feature annotations and are captioned primarily as “talking”. Voxel weight analysis revealed that anterior STS is most highly weighted by this component. Conversely, the second component responds strongly to joint physical actions between people in the videos. This component has a significant correlation with the “joint action” feature annotations, even after controlling for motion energy, and top-responding videos are captioned as physical interactions such as “dancing”. These results are particularly noteworthy since neural responses to the labeled feature “joint action” have not been identified before. Voxel weight analysis indicates that this component is most strongly weighted in mid-level regions of the lateral stream, including middle temporal area (MT) and extrastriate body area (EBA). Together, our findings suggest that joint action and communication represent two distinct forms of social interaction that are encoded differently in posterior to anterior regions along the lateral visual pathway.

This work is supported by R01 grant (No. NIH R01MH132826)

23.351 FACE RACE MODULATES NEURAL INHIBITORY EFFICIENCY

Viola Benedetti[1,2] (viola.benedetti@unifi.it), Peter de Lissa[1], Fabio Giovannelli[2], Gioele Gavazzi[2], Maria Pia Viggiano[2], Roberto Caldara[1]; [1]University of Fribourg, Switzerland, [2]University of Florence, Italy

Race is a universal social construct that can be quickly extracted from facial features. How race information modulates face processing has been extensively investigated in face identification and categorization, resulting in thousands of studies. However, no studies have directly investigated the interaction of face race with the neural inhibitory executive function yet. This is surprising given that response inhibition is key for appropriate and effective human multicultural and multiracial social interactions. To address this issue, we tested Western-Caucasian observers using a face categorization by race task with a Go/No-Go paradigm, while recording high-density electrophysiological signals. Participants had to quickly categorize one race (Go stimulus), while inhibiting responses to the other race (No-Go stimulus), with executive inhibitory demands directly tapping into face race categorization. In one condition, Western-Caucasian (WC) faces (i.e., same-race) served as Go stimuli and East Asian (EA) faces (i.e., other-race) were the No-Go stimuli that had to inhibit. In the other condition, the Go race was inverted and presentation order was balanced across observers, with all face images normalized for spatial frequency and contrast. The rate-correct score (RCS) revealed a higher frequency of accurate responses specifically when EA faces were presented as No-Go stimuli, which highlights a stark behavioral other-race inhibition advantage. Using a mass-Bayesian EEG analysis, we then revealed an early race effect with higher occipital P1 amplitude for EA. Only in the inhibitory condition (No-Go) we observed a post-perceptual effect: a late central No-Go P3 for EA faces. Interestingly, this component is positively correlated with behavioral performance, reflecting a psychophysiological correlate for the other-race inhibition advantage. Altogether, our data show a crucial neural interplay between face race and executive inhibition, providing a novel functional signature on its temporal dynamics. These findings reveal how face race shapes our human interaction and the subsequent social categorization processes.

This work was supported with funding from the Swiss National Science Foundation awarded to RC (100019_189018)

23.352 FACES EVOKE SOCIAL NETWORK INFORMATION IN AMYGDALA AND ENTORHINAL CORTEX

Ji Young (Julie) Hwang[1], Allen Chen[1], Heejae Kim[1], Jerry Jin[1], Nadiat Adedoyin[1], Ed Connor[1]; [1]Johns Hopkins University

Face perception is crucial for understanding social networks and maintaining social relationships. Here, we investigated how face stimuli evoke neural representations of social network information in amygdala and entorhinal cortex of macaque monkeys. To characterize social network knowledge for individual monkeys, we collected surveillance videos of their home, neighboring, and unfamiliar social groups at the Johns Hopkins breeding farm. Videos collected across a three-month time frame were analyzed for interactive and solitary behaviors in four stable social groups, ranging in size from 7–18 individuals. We used this behavioral data to construct multi-edge social network graphs for these groups based on dominance relationships,

facial threat behavior, physical aggression, dominance mounting, stealing, submissive facial signals, physical submission, affiliative facial gestures, grooming, physical proximity, social play, and knowable genetic relationships. Two subject monkeys from the same group were studied with linear array probe recording in amygdala and entorhinal cortex while viewing photographs of monkeys from the home, neighboring, and unfamiliar groups. We analyzed neural coding of personal social knowledge about home and neighboring groups, using unfamiliar monkeys as a control for social information based solely on visual appearance. We found that many neurons in both amygdala and entorhinal cortex encode personal social knowledge about relationships involving the subject monkey, relationships involving other monkeys, and relationships across the entire groups (home and neighboring), including dominance hierarchies, physical and symbolic agonism, physical and symbolic affiliation, and the knowable tree of genetic relationships.

Kavli NDI at Johns Hopkins University

Saturday, May 18, 8:30 am – 12:30 pm, Banyan Breezeway

Multisensory Processing: Neural coding

23.353 CLASSIFICATION OF SOUND-SYMBOLIC SPEECH SOUNDS IN VISUAL CORTEX

stefano ioannucci[1] (stefano.ioannucci@unifr.ch), Carole Jordan[1], Anne-Sophie Carnet, Carolyn McGettigan, Petra Vetter; [1]Visual and Cognitive Neuroscience Lab, University of Fribourg, Switzerland, [2]UCL Speech Hearing and Phonetic Sciences, University College London, UK.

Sound symbolism, also known as the "bouba-kiki effect", is a psychological phenomenon revealing a non-arbitrary link between meaningless speech sounds and visual features. For instance, individuals consistently match rounded, curvy shapes with the word "bouba" and jagged, angular shapes with "kiki" (Lockwood & Dingemanse, 2015; Ramachandran & Hubbard, 2001), suggesting a cross-modal association between speech sounds and visual shape. Despite extensive behavioral evidence demonstrating sound symbolism, neural evidence in the human brain remains limited (Peiffer-Smadja & Cohen, 2019; Revill et al., 2014). Building on research indicating early visual cortex's response to auditory stimuli (Vetter et al., 2014, 2020), our study investigated whether visual cortices, especially those sensitive to shape, show distinguishable neural activity patterns to the implicit shape conveyed by sound symbolic speech sounds. We collected 3T fMRI BOLD signals from 8 blindfolded participants listening to sound symbolic words (rounded, spiky, and mixed) and matching them with visual shapes (on a 4-point scale from more spiky to more round). Beta weights from visual regions (V1, V2, V3, and LOC) were extracted, and multi-variate pattern analysis assessed if sound-symbolic word categories yielded differential neural patterns. Behavioural results replicated the "bouba-kiki" effect – "rounded" words were associated with rounded shapes and vice versa for "spiky" words. Preliminary fMRI analyses indicated successful decoding of "round" and "spiky" sounds in visual cortical regions for 5 participants. These findings provide evidence that fMRI activity in early visual and shape-selective cortical regions can contain information about sound symbolic word l associations, in the absence of visual stimulation. This supports sound symbolism as a genuine cross-modal effect in the human brain and suggests that visual brain regions represent also high-level auditory information, emphasizing the depth of audiovisual interactions in visual cortex.

This work was supported by a research grant from the BIAL Foundation (No. 238/20) to P.V. and C.M and by a PRIMA grant (PR00P1_185918/1) from the Swiss National Science Foundation to P.V.

23.354 SPATIOTEMPORAL DYNAMICS OF NEURAL REPRESENTATIONS DURING PERCEPTION OF NATURALISTIC AUDIOVISUAL EVENTS

Yu Hu[1] (yhu584@uwo.ca), Yalda Mohsenzadeh[1]; [1]Western University

What we see and hear carry different physical properties, but we are able to integrate the distinct information to form a coherent percept. The cross-modal integration is observed at many brain regions including primary and non-primary sensory areas as well as high-level cortical areas. Most previous studies on audiovisual integration used flash/tones or image/sound pairs, which are easy to manipulate the experimental conditions but lack ecological relevance. Under more natural scenarios when audiovisual events are perceived, however, where and when different levels of information are processed and integrated across brain areas and over time remain less investigated. To address that, we selected sixty 1-second naturalistic videos with representative visuals and sounds of three categories - animals, objects, and scenes. We recorded both functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) data when participants (N=19) viewed videos and listened to the accompanying sounds while doing an orthogonal oddball detection task. With multivariate pattern analysis and representational similarity approach, we found that the visual and acoustic features were processed almost simultaneously, with the onset at ~60 ms and the first peak at ~100 ms. The acoustic information was represented not only in auditory areas, but also in visual areas including the primary visual cortex and high-level visual regions, demonstrating the early cross-modal interactions. However, the visual features were only represented in visual cortices, suggesting asymmetrical neural representations of modality information during multisensory perception. The high-level categorical and semantic information emerged later in time with the onset at ~ 120 ms and the peak at ~210 ms and was observed at high-order visual and association areas as well as the parietal and frontal cortex. By fusing the representations from fMRI and EEG, we also resolved the neural processing during audiovisual perception at each voxel and at each millisecond.

23.355 DISSOCIABLE DECODING OF PREDICTIVE SENSORY PROCESSING FROM EEG OSCILLATIONS

Soukhin Das[1,2] (skndas@ucdavis.edu), Dr. Mingzhou Ding[3], Dr. George (Ron) Mangun[1,2,4]; [1]Center for Mind and Brain, [2]Department of Psychology, University of California, Davis, [3]Pruitt Family Department of Biomedical Engineering, University of Florida, [4]Department of Neurology, University of California, Davis

Studies have established that attention is capable of operating across diverse sensory modalities, such as vision and audition, playing a pivotal role in our ability to integrate and process multisensory information. Despite this, the neural mechanisms that underlie cross-modal attention remain largely elusive. In this investigation, we utilized event-related potentials (ERPs) to probe the neural basis of cross-modal attention through a 2x2 cue-target design. Auditory (HEAR or SEE) or visual cues (H or S) were employed to indicate the modality (visual/auditory) of the to-be-attended target. After a random delay, auditory tones or visual gratings were presented as target stimuli in the cued modality in 80% of the trials. Conversely, in 20% of the trials, the targets were presented in the un-cued modality, constituting invalid trials. Participants (n=32) were instructed to discriminate the frequency (wide versus narrow) of visual gratings or the tone (high versus low) of auditory stimuli across all trials, regardless of cue validity. Decoding alpha power using SVM uncovered distinctive patterns in early and late latencies during the cue-to-target period. Alpha oscillations exhibited unique cortical patterns based on the to-be-attended target modality. We found robust decoding accuracies for the to-be-attended modality within respective sensory areas, i.e., central electrodes for the auditory and parieto-occipital electrodes for the visual modality. Temporal generalization further illustrated the evolving nature of alpha patterns over time. For both modalities, our findings indicated the sustained representation of sensory information in a serial manner across the hierarchy, emphasizing the maintenance of predictive processing. Furthermore, an alignment between cortical alpha patterns during stimulus processing and the response window suggested a connection between prediction signals and decision-making processes. Our findings contribute to understanding the role of alpha oscillations in cross-modal attentional control. This work extends the current framework for decoding the neural mechanisms of cross-modal attention.

NIMH117991

23.356 TOPOGRAPHIC MAPPING OF VISUAL SENSATIONS IN A COMPLETELY BLIND INDIVIDUAL

Jesse L. Breedlove[1] (jbreedlo@umn.edu), Logan T. Dowdle[1]; [1]University of Minnesota

Non-optic sight (NOS) is a phenomenon in which blind individuals have visual sensations of things they touch, hear, or otherwise sense in their surroundings. The current study is exploring the neural correlates of NOS in a woman (subject NS) who lost her sight to retinal degeneration. The images that NS sees are primarily triggered through touch and proprioception (e.g., if she picks up a mug, she sees a mug). Unlike imagery, these images are determinate, involuntary, and persist as long as she infers the object remains in her line-of-sight. In a previous fMRI experiment, NS placed simple 3D shapes in her field-of-view. The resulting BOLD signals hinted at visual cortical topographic mapping of the "seen" objects (smaller shapes evoked more foveal activation). To test whether this activity really follows "retinotopic" organization, we designed and 3D-printed tactile versions of retinotopic mapping stimuli (3-dimensional wedge and bar) which could be used within a 7T scanner. Every 8s, NS was cued to either rotate the wedge around a fixed point (16 positions) or shift the bar up/down or left/right (9 positions each). We then performed population receptive field mapping using the brain data and the stimuli translated into 2D apertures. Receptive field parameters across NS's visual cortex showed a striking topographic organization resembling typical sighted retinotopy: increasing eccentricity moving anteriorly from the occipital pole and voxels tuned to portions of the visual field that are contralateral and flipped relative to their respective cortical locations. Fascinatingly, some phase reversals can even be observed occurring roughly at the anatomical locations where transitions between V1/2/3 are typically seen. These results represent the first topographic mapping of stimuli experienced as visual perceptions in a completely blind individual and demonstrates that the visual cortex can support concrete visual experiences that accurately interpret non-retinal sensory input, even in blindness.

Research reported in this presentation was supported by the University of Minnesota's Neuroimaging Postdoctoral Fellowship Program, College of Science and Engineering's, and the Medical School's NIH funded T32 Neuroimaging Training Grant. NOGA: 1T32EB031512-01

23.357 SOMATOSENSORY MODULATION OF VISUAL CORTEX RESPONSES DURING TRACE CONDITIONING UNDER ANESTHESIA

Mykyta M. Chernov[1] (mykytachernov@gmail.com), Pauline Audurier[1], Dave S. Jacobs[1], Robert M. Friedman[1], Vincent D. Costa[1]; [1]OHSU

Learning and memory enables organisms to make predictions about the outcomes associated with specific stimuli. Although anesthesia inhibits the formation of explicit memories, it does not hinder the formation of implicit memories (Samuel et al., 2018).The neural mechanisms underlying implicit learning and memory formation under anesthesia are understudied, especially for multisensory associations. We recorded neural activity in early visual cortex of rhesus macaques under propofol anesthesia to determine how neuronal responses are modulated during a trace conditioning procedure in which naturalistic visual stimuli were flickered at 10 Hz (CS; e.g. faces, objects, and symmetrical visual textures) and probabilistically associated with delayed somatosensory stimulation of the fingers (US). One set of visual cues was associated with a high probability of somatosensory stimulation while the other was associated with a low probability. Pupillometry was used to verify that conditioned associations were learned (n = 5). Trial-by-trial changes in the predictability of each cue to signal subsequent somatosensory stimulation were encoded by changes in pupil size. The most prominent outcome was pupillary constriction immediately after experiencing somatosensory US, regardless of whether it was preceded by a high or low probability visual cue. Multi-contact neurophysiology probes were used to record in areas V1, V2, and V4 (n = 3). In each area, 30-50% of the recorded neurons exhibited modulations in ongoing neural responses elicited by the flickering visual cues depending on whether the somatosensory US was delivered. Complementary experiments used intrinsic optical imaging (IOI) to measure changes in neural activation at the mesoscale similarly found somatosensory related modulation of hemodynamic responses in areas V1, V2, and V4. These results imply that under anesthesia implicit multisensory associations can modulate early visual cortical responses. These findings provide a framework to a better understand the neural circuitry underlying implicit learning and the neurocognitive adverse effects of anesthesia.

Multisensory Processing: Development, clinical

23.358 VISUAL DISCRIMINATION AFTER THE MULTISENSORY REHABILITATION OF HEMIANOPIA

Benjamin Rowland[1] (browland@wakehealth.edu), Naomi Bean[2], Barry Stein[1]; [1]Wake Forest School of Medicine, [2]Duke University School of Medicine

Unilateral cortical blindness is a common consequence of damage to visual cortex on one side of the brain. This condition is often permanent and intractable to rehabilitative techniques. Recently, we have developed an effective, non-invasive multisensory 'training' paradigm that can ameliorate this condition within weeks. The paradigm involves repeated exposure to spatiotemporally congruent cross-modal (visual and auditory) stimuli in the blind (i.e., contralesional) hemifield. It is believed to restore visual responsiveness to neurons in the midbrain superior colliculus (SC) that were rendered unresponsive as a secondary consequence of the cortical lesion. Here we asked whether the functional restoration was restricted to the visuomotor detection/approach behaviors known to be subserved by the SC or also involved a capacity to distinguish spatial patterns. Cats were trained in a visual detection/approach task as well as a battery of visual pattern discriminations. They were required to indicate (via button press) whether a pair of simultaneous visual stimuli were the same or different along a given dimension such as size, the direction of movement, orientation, or shape. The two stimuli were either presented within or across the two hemifields while the animal fixated centrally. During prelesion testing animals discriminated patterns everywhere in space. After all contiguous areas of unilateral visual cortex had been removed, animals were blind across the entire contralesional hemifield. However, after weeks of cross-modal training, visual detection/approach was restored. The training paradigm also restored their ability to discriminate visual patterns presented together in the previously blinded hemifield or across the two hemifields. Post-rehabilitation performance was well above chance but still below pre-lesion performance levels. These data reveal that the visual processing capabilities restored by this multisensory training paradigm extended far beyond the visually-guided behaviors commonly associated with the SC to include the visual discrimination capabilities that are commonly associated with the neocortex.

NIH R01 EY031532

23.359 SENSORIMOTOR DYNAMICS OF ECHOACOUSTIC TARGET ACQUISITION IN BLIND HUMANS

Arshil Patel[1], Giovanni Fusco, Santani Teng[1]; [1]Smith-Kettlewell Eye Research Institute

Some blind humans actively echolocate, analyzing reflections from self-generated vocalizations to perceive the environment. Similar to visual scene exploration, real-world echolocation is an active-sensing process, requiring the goal-directed interplay of motor actions (e.g. locomotion and oral click production) and the resulting sensory percepts. To explore this dynamic, here we recorded head movements and clicking behavior of participants as they echoacoustically localized a target object. In each trial, participants used echolocation to guide their head orientation toward the target, positioned at a 1-meter distance and random azimuth in the frontal hemifield. Target size was Large or Small to manipulate difficulty; control blocks omitted active clicks. Participants ended the trial by holding their head still and pressing a button. Analysis of video recordings extracted frame-by-frame information such as target azimuth, head pose, azimuth error (head/target relative angle), and the participant's response (head orientation during button press). To evaluate performance and sensorimotor dynamics for each trial, we examined the final azimuth error, head orientation time course, and click timestamps extracted from the audio recording. Initial results from an early-blind proficient echolocator show clear advantages to echolocating a larger target, with average angular error of ~9.5°, vs. ~24.5° for the small target (35.8° in the passive no-click control condition). Compared to large targets, small-target trials lasted ~14s longer and contained ~11.5 more head direction reversals, both factors of 2.5, suggesting that head kinematics remain relatively invariant to task conditions. Small-target trials contained ~3.3x as many clicks on average, reflecting a moderate but significant increase in average click rate (0.44 Hz, $p<.001$). Pooled across trials, small-target clicking behavior converged steadily toward the target center, with over 75% of clicks directed within 20° of the target, suggesting an intensity-maximization strategy. Together, these results illustrate naturalistic sensorimotor dependencies underlying auditory active sensing in the absence of vision.

Foundation for Ophthalmology Research and Education International (to ST) National Eye Institute 1R21EY032282-01 (to ST) Smith-Kettlewell Rehabilitation Engineering and Research Center

23.360 INTEGRATING IMPAIRED VISION AND HEARING TO IMPROVE SPATIAL LOCALIZATION

Yingzi Xiong[1,2] (yxiong36@jh.edu), Quan Lei[3], Shirin Hassan[4], Daniel Kersten[2], Gordon Legge[2]; [1]Johns Hopkins University, [2]University of Minnesota, [3]Wichita State University, [4]Indiana University

Introduction. Spatial localization, which is critical for safe mobility and social interactions, relies heavily on vision and hearing. When vision and/or hearing impairment occurs, integrating vision and hearing may maximize the use of the residual senses. However, such impairment is often associated with degraded sensory input and unstable sensory status, which may influence the integration process. Here we investigated the integration of vision and hearing in a spatial localization task in individuals with heterogeneous vision and hearing impairment. Methods. Eighty-five participants completed a spatial localization task: 36 younger and 13 older controls with normal vision and hearing, 10 with hearing impairment only, 13 with vision impairment only, and 13 with dual vision and hearing impairment. Participants verbally reported the directions of visual (200ms, 3 deg diameter, 90% contrast target), auditory (200ms, pink noise with 200-8000 Hz, 60 dB Hearing Level), or audiovisual targets (simultaneous from the same location) across 17 locations spanning 180 degrees in the horizontal plane. Spatial biases (offsets) and uncertainties (variability) were obtained for each location in each condition. Results. Vision and hearing impairments were each associated with increased biases and uncertainties in unimodal localization, resulting in large variations across locations and individuals. To reconcile these variations, we identified individualized integration zones and segregation zones based on whether the audiovisual discrepancies support a common cause inference. Across all locations, people with sensory impairment, especially those with dual sensory impairment, showed less integration zones than controls. However, the benefit of integration (reduced uncertainty in the bimodal condition) in the integration zones, or lack thereof in the segregation zones, were consistent across all groups. Conclusion. Impairments in vision and hearing reduce the likelihood of making a common cause inference while localizing a bimodal target. However, the advantage of integration persists when the criteria for a common cause are satisfied.

NIH R00 EY030145

23.361 PRIOR VISUAL EXPERIENCE INCREASES CHILDREN'S USE OF EFFECTIVE HAPTIC EXPLORATION STRATEGIES IN AUDIO-TACTILE SOUND-SHAPE CORRESPONDENCES

Shibo Cao[1] (shibo.cao001@umb.edu), Julia Kelly[1], Cuong Nguyen[1], Hiu Mei Chow[1,2], Brianna Leonardo[1], Aleksandra Sabov[1], Vivian Ciaramitaro[1]; [1]University of Massachusetts Boston, [2]St. Thomas University

Sound-shape correspondence refers to the preferential mapping of information across the senses, such as associating a nonsense word like “bouba” with rounded abstract shapes, and “kiki” with spiky abstract shapes (e.g., Ramachandran & Hubbard, 2001, Spence 2011). Such associations are found between auditory and visual stimuli (AV), and between auditory and tactile stimuli, which are touched but not seen (AT). Previous research highlights how atypical visual experience in the blind (e.g. Fryer et al., 2014; Sourav et al., 2019) and naive visual experience in children (Chow et al., 2021) can yield weak or absent AT associations. Such findings imply that visual experience influences AT association, yet, the mechanisms underlying such influence remain unclear. Here we investigate one potential mechanism: seeing the abstract shapes improves haptic exploration by increasing the use of effective haptic strategies and/or decreasing the use of ineffective haptic strategies in children. We examined haptic exploration from videos of a previous study (Chow et.al., 2021). Thirty 6-8 year-old children completed 16 AV and 16 AT trials, with order counterbalanced across participants. Children picked which of two, side-by-side, visual shapes (AV trials) displayed on a screen, or tactile shapes (AT trials) presented hidden from view inside a box, matched a nonsense sound. We quantified the proportion of duration of each of 5 haptic exploration strategies (effective: contour following, clawing, pinching, and poking, and ineffective: sweeping) for each participant, as they explored tactile shapes during AT trials. We found that with prior visual experience (AV trials first), poking, an effective strategy, was dominant whereas without prior visual experience (AT trials first), uncategorizable, ineffective, strategies were dominant. These findings suggest that prior visual experience of abstract shapes in 6-8 year-olds can increase the of haptic exploration, potentially explaining why prior visual experience can strengthen audio-tactile sound-shape correspondences early in development.

23.362 THE DEVELOPMENT OF MULTISENSORY FACILITATION FOR SIMPLE AUDIO-VISUAL INFORMATION CONTINUES THROUGH ADOLESCENCE.

Armando Bertone[1] (armando.bertone@mcgill.ca), Alexia Ostrolenk[1,2]; [1]McGill University, [2]Unity Health Toronto

Background. Significant variability characterizes the reported age ranges regarding the maturation or attainment of 'adult-like' ability for audio-visual (AV) multi-sensory integration (MSI). In part, this variability originates from diverse characteristics of unimodal audio and visual stimuli (e.g., social vs. non-social) used to assess MSI and the inconsistent developmental periods selected for cross-sectional evaluations. This study assessed the development of MSI for simple, non-social information from the early school-age years through to adulthood. Methods. One-hundred and eleven (n=111) neurotypical participants were placed into either child (6-9 ys, Mage=8.2), school-age (10-13 ys, Mage =12.0), adolescent (14-17 ys, Mage=15.0) or adult (18+ ys, Mage=22.7) age groups. Participants completed an audio-visual target detection task (Hershenson, 1962), presented with either a highly salient auditory (A) stimulus (beep), visual (V) stimulus (flash), or both stimuli at the same time (AV). Reaction times (RT) were recorded for a total of the 240 trials, with outliers (<100 or >1500 ms) excluded from analysis. Race Model analysis was conducted for each age group to compare audiovisual RTs with those expected from redundant stimuli, and a redundancy gain was calculated for each participant. Results. Increasing violations of the race model inequality as a function of age group were found, suggesting a progressive increase of AV multisensory facilitation for simple non-social information across developmental periods. Notably, the race model was violated through the 65th percentile of the RT distribution for the adolescent (14-17 ys) age group, which differed from that of the adult group (85th percentile). Conclusion. AV multisensory facilitation is not fully developed at adolescence, even when unimodal stimuli features are simple. This finding supports a later rather than earlier maturation period for AV MSI. These results also underscore the importance of including all development periods when assessing the development of AV MSI ability, including the adolescent age period.

Fonds de recherche du Québec - Santé Research Scholar Award J2 to AB

23.363 THE RELATIONSHIP BETWEEN COGNITIVE ABILITY AND SUSCEPTIBILITY ON THE SOUND-INDUCED FLASH ILLUSION (SIFI) ACROSS DEVELOPMENT

Samantha Scholes[1] (samantha.scholes@mail.mcgill.ca), Nathalie Jack[1], Armando Bertone[1]; [1]McGill University

Introduction. Multisensory integration (MSI) is the ability to combine multiple sources of sensory information into one coherent whole. While understanding the developmental trajectory of MSI is important, it is imperative to account for other variables which may contribute to this ability, such as cognitive functioning (or IQ). MSI performance has been most frequently compared to other attributes (i.e., language), with IQ being scarcely considered as an associated attribute. The aim of this study was to examine the age at which MSI for non-social information matures and the role of IQ. Methods. 101 neurotypical participants were placed into child (6-9ys, Mage=7.6), school-age (10-13ys, Mage =11.8), adolescent (14-17ys, Mage=15.0) and adult (18+ys, Mage=22.3) age groups. All participants completed the

Sound-Induced Flash Illusion (SIFI), where they were asked to indicate how many flashes were perceived while ignoring auditory cues. Performance on the fusion (2F1B) and fission (1F2B) conditions were examined. Participants also completed a cognitive assessment (Wechsler Abbreviated Scale of Intelligence-2; WASI-2), with a primary interest on perceptual reasoning (PRI) and verbal comprehension (VCI) indices. Results. A 2 (illusion conditions) x 4 (groups) ANOVA revealed a significant main effect of condition ($p<.01$) and interaction effect ($p<.01$), with post-hoc analyses demonstrating increased overall susceptibility for the fission condition; however, adults were significantly less susceptible on this condition compared to the other age groups. Hierarchical multiple regressions revealed that higher PRI scores were associated with increased susceptibility to the fission condition across age groups. Conclusion. Results suggest decreasing susceptibility solely to the fission illusion with increasing age, and that PRI may be related to susceptibility to this illusion. These results highlight the need to consider non-verbal IQ in assessing performance on non-social MSI tasks across development. Future studies should continue to explore the relationship between IQ and MSI, including social MSI and verbal IQ.

23.364 THE NOISY ENCODING OF DISPARITY (NED) MODEL PREDICTS PERCEPTION OF THE MCGURK EFFECT IN NATIVE JAPANESE SPEAKERS

Anastasia Lado[1] (nastialado@gmail.com), John Magnotti[1], Michael S. Beauchamp[1]; [1]University of Pennsylvania

The McGurk effect is widely used to demonstrate the importance of visual information in speech perception. Observers view a face enunciating a syllable while hearing an auditory recording of an incongruent syllable, evoking a fusion percept different than either component (McGurk & MacDonald, 1976). The Noisy Encoding of Disparity (NED) model uses probabilistic inference to predict perception by assuming that different McGurk stimuli have fixed stimulus disparities which observers encode with noise and compare to an internal threshold (Magnotti & Beauchamp, 2015). The NED model has been shown to accurately predict McGurk perception in diverse populations, including U.S. college students, children in the United Kingdom (Hirst et al., 2018), and cochlear implant users in Germany (Stropahl et al., 2017). To examine whether the NED model could predict McGurk perception in a non-Western population, 82 native Japanese-speaking participants were recruited from two Japanese universities. Using online testing, 15 different McGurk stimuli were presented 10 times each, along with 30 audiovisual congruent stimuli, in random order. Participants provided typed responses and no feedback was ever given. For control congruent stimuli, accuracy was uniformly high (96%). In contrast, the McGurk effect was highly variable across stimuli and participants. The weakest stimulus evoked the fusion percept on 3% of trials compared with 77% for the strongest stimulus. For more than half of participants, the strongest McGurk stimulus induced the illusion on every trial. Across participants, the NED model accurately predicted the rate of fusion responses for held-out stimuli, mean r^2 = 0.71 ± 0.09, p < 10^-17. The effects of cultural and linguistic factors in shaping the McGurk effect remain a subject of debate. These results demonstrate similarly high levels of variability in the McGurk effect for Western and non-Western populations and show that the NED model can accurately predict perception in both populations.

The research was funded by NIH NS065395 and NS113339.

SATURDAY, MAY 18, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Motion: Higher-order

23.365 MAY WE CHANGE THE COLOR OF THE VOLLEYBALL?

Seung min Han[1,2] (hse001209@naver.com), Hyeonjin Kim[1,2], Hoon Choi[1]; [1]Department of Psychology, Hallym University, [2]Graduate School of Digital Contents, Hallym University

Color media colored balls in sports. In the past, volleyball games were played with white balls, but FIVB introduced colored balls to activate TV broadcasts. Thus, the colors and patterns applied to the balls are diversifying. Balls with various colors and patterns spinning rapidly can cause color mixing, further creating an optical illusion of size. Volleyball players use their eyes to see the flying ball and their hands to receive it. This harmonious interaction of the body for effective sports performance is called motor coordination. Players in sports that use balls respond to visual stimuli and exert motor coordination skills. However, if optical illusions occur, their visual perception may be distorted, negatively impact performance. Our experiment explored the effects of color and speed on perceived size of balls. The experiment employed a 3x5 within-subject design with color (black+white, blue+red, yellow+red) and speed (36,72,108,144,180 deg/sec) as independent variables, involving 22 adults. The stimuli were circles with two-color striped patterns, tilted clockwise from vertical by 45°. The experiment is total of 300 trials, organized into 5 blocks of 60 trials each. In each trial, participants saw a moving circle with a size of 11.58°x11.58° on the screen. They then reported their perceived size by expanding a 0.46°x0.46° white circle. Repeated Measures 2-way ANOVA analysis revealed a significant main effect of color ($p=.047$). When the circle had color, it was perceived as larger than when there was no color. The main effect of speed was also significant ($p<.001$). It showed an inverted U-shaped curve in which the perceived size increased as the speed increased up to 108 deg/sec and then decreased thereafter. The interaction between color and speed showed a marginal effect ($p=.098$). These results show that the color and speed of a moving ball can influence the perceived size of the ball.

23.366 ACTION CLASSIFICATION FROM MOTION FLOW

Uriel Gonzalez[1] (uriel.gonzalezbr@rutgers.edu), Nathan Destler[1], Jacob Feldman[1]; [1]Rutgers University

How do we classify other people's actions, and what information do we use to do it? Ever since Johansson's introduction of Point Light Displays (PLDs) it has been known that human observers can integrate a few moving points into a human form and, in many cases, infer actions, intentions, and emotions. However it is not clear what kind of representation supports this ability---dynamically changing shape, or just motion flow? Some studies have suggested that motion flow suffices to achieve action classification, but the evidence for this is based on very limited action sets, such as distinguishing walking from noise. This raises questions about the sufficiency of various kinds of visual information to support the recognition of a broader range of actions. To address this gap, we used OpenPose to generate reduced video stimuli from multiple action videos drawn from naturalistic scenes. This research uses 78 actions, including everyday activities like brushing teeth, drinking water, etc. For these 78 actions, three distinct types of videos were created: PLDs (consisting of dots at joint locations), Stick Figures (joint positions joined in an anatomically correct body plan), and Motion Flow Videos (flow fields based on the Lucas-Kanade algorithm). Participants were asked to identify the actions in each video (free text description). We employed a Natural Language Processing model to estimate the semantic similarity of each participant's response to that of others, allowing us to automatically estimate intersubjective agreement. Further analysis was conducted using a Hierarchical Bayesian Model, which compared the posterior and predictive distributions of the semantic similarities across each video condition. Intersubjective agreement was highest with Stick Figures, followed by PLDs, and lowest for Motion Flow videos, suggesting that dynamic pose representations are indeed required for accurate action classification, and that motion flow supports at most a coarser classification.

23.367 ARE RICH PERCEPTS FROM POINT-LIGHT DISPLAYS SPECIFIC TO BIOLOGICAL MOTION?: A CASE STUDY OF DYNAMIC POINT-LIGHT CLOTHS

Merve Erdogan[1] (merve.erdogan@yale.edu), Wenyan Bi[1], Ilker Yildirim[1], Brian Scholl[1]; [1]Yale University

Visual processing seems specialized in several ways for the perception of other agents, and one of the best examples of this is biological motion—as when displays of moving dots ('point-light walkers'; PLWs) give rise to rich percepts of locomoting agents. Despite thousands of experiments over decades of research, one of the most foundational questions about such phenomena remains unsettled: just how specific are they to *biological* motion? Addressing this question is challenging, largely because of the absence of non-biological comparison stimuli — since the translation or rotation of rigid objects (as in familiar structure-from-motion displays) lacks the rich characteristic *relative* motion of the points from PLWs. Here, to fill this gap, we introduce a novel visual phenomenon: the perception of rich behavior from dynamic point-light *cloths* (PLCs) — as when a sheet on a clothesline is waving in the wind. Across many experiments, we found broad similarities between the perception of PLWs and PLCs — in terms of both experimental results and phenomenological demonstrations. In the first place, people readily perceive the dynamic shape and behavior of cloths from point-light displays, and (as with PLWs) this depends critically on the points' relative motions, since such rich percepts disappear both in static displays and in dynamic displays with spatially scrambled points. This applies not only to the perception of cloth structure itself, but also to more nuanced properties, such as a fabric's stiffness. Moreover, the perception of PLCs is also highly robust, persisting (as do PLWs) e.g. even in displays with limited-lifetime points, or when embedded in noise. These results collectively demonstrate how the perception of rich internal structure from dynamic point-light displays transcends biological motion, and also applies to the behavior of other sorts of non-rigid entities: even beyond biology, the visual system extracts rich structure from surprisingly limited input.

23.368 VISUAL EXPERIENCE AND SENSITIVITY TO MOTION SEQUENCES, FROM HUMAN MOVEMENT PATTERNS TO ANIMATED SHAPES

Dhun Verma[3], Shlomit Ben-Ami[1,2], Alish Dipani[3], Naviya Lall[3], Suma Ganesh[4], Chetan Ralekar[1], Sharon Gilad-Gutnick[1], Pawan Sinha[1]; [1]MIT Department of Brain and Cognitive Sciences, Cambridge, MA, USA, [2]Sagol School of Neuroscience, School of Psychological Sciences, Tel-Aviv University, Tel-Aviv, Israel, [3]The Project Prakash Center, New Delhi, India, [4]Department of Ophthalmology, Dr. Shroff's Charity Eye Hospital, Delhi, India

Humans effortlessly interpret complex motion sequences and are especially attuned to the kinematics and dynamics of human actions and embedded social cues. Remarkably, as demonstrated by Heider and Simmel animations (H&S, 1944), even very simple interacting geometric shapes evoke rich social narratives. The extent to which sensitivity to interaction sequences relies on high-level inferences versus more bottom-up visual processes remains unclear. Exploring the role of visual experience in the development of this sensitivity can provide valuable insights. In prior work with patients treated for congenital cataracts as a part of Project Prakash, we probed their post-surgery capacity to recognize and interpret motion sequences, revealing a diverse spectrum of abilities. On the one hand, we observed recognition of human point-light displays immediately upon eye-opening following surgery (Ben-Ami, Neuropsychologia, 2022). On the other hand, we found compromised extraction of goal-directed inter-shape interactions from H&S sequences, even years after treatment (Verma, JOV, 2023). This study addressed this observed disparity by examining sensitivity to point-light displays of dyadic physical interactions (e.g., hugging, fighting). Like H&S animations, these displays capture coordinated movements between two interacting individuals. Additionally, they encompass the structural and kinematic information of naturalistic human movement found in human walker point-light displays but missing from H&S animations. Patients were tested before and up to one year after sight-restoring surgery (longitudinal, n=10) or five years post-surgery (cross-sectional, n=8). Employing a free-report design, patients described observed interactions and extracted their valence ('aggressive' or 'friendly'). Cross-sectional findings reveal fully developed recognition and valence extraction, while longitudinal results indicate a gradual onset of these abilities over time after surgery. The distinct dependency of point-light versus H&S interaction recognition on visual experience implies potentially divergent underlying mechanisms. We shall offer hypotheses to account for this intriguing distinction between simple visual agents and those anchored in biological kinematics.

Acknowledgments: NEI (NIH) grant R01 EY020517 to PS

23.369 PERCEPTUAL SWITCHES DURING STRUCTURE-FROM-MOTION DO NOT ELICIT PUPIL RESPONSE

Bobicheng Zhang[1] (zhangb42@msu.edu), Jan Brascamp[1]; [1]Michigan State University

During perceptual bistability, the observer's percept typically switches unpredictably between two alternative interpretations despite the physical properties of the stimulus staying the same. Such perceptual switches have been suggested to be linked to norepinephrine-based neuromodulation in the brain, as pupil dilations have been found to occur in association with perceptual switches. However, the pupil dilates in response to task-relevant events in general, and in existing work perceptual switches were explicitly reported and thus task-relevant. As such, observed switch-related dilations may have reflected non-specific task relevance rather than switch-specific processes. Here we measured pupil responses to perceptual switches that were not task-relevant. Observers viewed a structure-from-motion rotating sphere consisting of equilateral triangles that inverted at semi-random intervals (approximately every 3 seconds). In separate conditions observers either reported perceptual switches (rendering them task-relevant) or reported changes in the triangles' orientation (rendering the perceptual switches task-irrelevant). We tracked participants' eyes and used an algorithm based on optokinetic nystagmus (OKN) to infer perceptual switch moments, even when observers did not report them. Control analyses confirm the reliability of this algorithm. We found that task-relevant switches elicited pupil-dilations, but task-irrelevant ones did not. These results suggest that pupil-associated neuromodulation, while closely linked to task-relevant events, may not have any specific tie with perceptual bistability. These results are consistent with results we recently reported for binocular rivalry, and thereby indicate that similar results hold across distinct forms of perceptual bistability.

23.370 DETECTING PURSUIT IN DYNAMIC VISUAL SCENES

Maria Kon[1] (maria.kon.ctr@nrl.navy.mil), Sangeet Khemlani[1], Andrew Lovett[1]; [1]U.S. Naval Research Laboratory

To determine whether one object is pursuing another, people must track them over time and compare their locations. Currently, no computational theory exists to explain this behavior. To support creating such a theory, we developed a novel paradigm for tracking how people detect pursuit. Participants (n = 94) observed videos of uniquely colored circles moving on a black screen and a) judged whether a red circle was pursuing any other circle, and b) indicated the color of the pursued circle. This task is innovative in that it both: a) permitted participants to respond as soon as they made a judgment, which provided information about the timecourse of processing; and b) had participants select the circle they thought was being pursued even when they incorrectly judged that pursuit was occurring, which provided information about the types of errors made. This information was compared to task performance by a computational model we developed within a cognitive architecture. This model operates by detecting the red circle, tracking it and computing its trajectory, scanning along that trajectory to detect candidates for pursuit, and integrating that information over time. The model makes accuracy and reaction time predictions, and model results are largely consistent with empirical findings: it, like humans, is slower for absent than present trials, and it, like humans, is slower for larger sets of circles. Likewise, it predicts when a video is reliably more difficult, and it makes human-like detection errors. The study and modeling results suggest that observers who erroneously detect that a pursuit relation is present do so because some non-pursued circle happens to move parallel to the red circle for a short duration; and those who fail to detect a pursuit relation do so because they render a judgment too early. Future work will examine how systematic these errors are.

This research was funded by a National Research Council Research Associateship awarded to MK, and data were collected by Knexus Research Corporation.

23.371 FROM VISUAL FEATURES OF MOVING OBJECTS TO SUBJECTIVE IMPRESSIONS OF CAUSALITY

Lina Eicke-Kanani[1] (lina.eicke-kanani@tu-darmstadt.de), Yunyan Duan[1], Thomas S. A. Wallis[1,2]; [1]Technical University of Darmstadt, Germany, [2]Centre for Mind, Brain and Behaviour (CMBB), Universities of Marburg, Giessen and Darmstadt, Germany

In daily life, we instinctively attribute causes and effects to objects in our environment based on visual information, though the precise mechanisms underlying the perception of causality remain unknown. Previous work on causality perception employing Michotte's launching paradigms (or variations thereof) implied a correlation between the physical plausibility of a stimulus and behavioral causality judgments, given an observer's sensory uncertainty around the objects' moving trajectories. To examine the relationship between sensory uncertainty, physical plausibility and perceptual causality, we rendered launching events with varying levels of physical realism and measured both observers' sensory uncertainty around moving trajectories and their respective causality judgments. Our experimental conditions included horizontal and oblique launch directions at fast and slow speeds to systematically manipulate uncertainty; in addition, we added physically implausible moving trajectories to the causality judgment task. We fitted psychometric functions to the responses and computed uncertainty measures from the slope of the psychometric fits. Overall, our results did not suggest a simple relationship between an individual's uncertainty around moving trajectory and their respective causality judgments. Instead, we found substantially higher uncertainty for causality judgments than for angle judgments with a stronger interaction of moving speed and launch direction, implying that causality perception was modulated to a greater extent by the temporal dynamics of the display than angle perception. That interpretation is consistent with previous research on the launching effect, indicating that timing is a prerequisite to visually integrate the individual objects' motion into the perception of one continuous motion. We speculate that observers' subjective causality reports were more related to the degree to which they could postdict a plausible sequence of events rather than a prediction about the physically plausible trajectories.

This work was supported by the cluster project "The Adaptive Mind" as part of the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art.

23.372 NEURAL REPRESENTATIONS THAT REVEAL A UNIFIED CONTINUUM BETWEEN PHYSICAL AND SOCIAL EVENTS

Sajjad Torabian[1] (torabias@uci.edu), John A. Pyles[2], Yujia Peng[3], Hongjing Lu[4], Emily D. Grossman[1]; [1]University of California, Irvine, [2]University of Washington, [3]Peking University, [4]University of California, Los Angeles

Introduction. Humans share strong intuitions about physical events and also about the behavior of social beings. These two capacities have often been treated as distinct. For example, evidence has linked intuitive physics and reasoning about social behavior to separate neural systems that include the intraparietal sulcus (IPS) and the superior temporal sulcus (STS), respectively. Recent computational work (Shu et al. 2021), however, has modeled the perception of violations of physical laws and impressions of goal-directed movement within a unified psychological space, namely the physical social forces model (PSF). Here, we test whether this computational model can explain neural responses to physical and social events. Methods. In a rapid event-related fMRI design, participants viewed short animations (3s each) of two moving shapes within a square box. The movement dynamics were systematically manipulated such that they elicited a wide range of violations of physics and degrees of intentionality. Results. A searchlight representational similarity analysis revealed cortical maps that were best explained by the PSF model, a categorical model of physical vs social events, and models based on low-level visual cues. The categorical model appeared as a strong predictor of the neural similarity around the pSTS and TPJ. The fine-grained estimates of violation of physics and intentionality of the PSF model also robustly predicted neural similarity around the IPS and early visual cortex, which variance partitioning attributes to perceptual features of average speed and average distance between the shapes. Importantly, after removing animations with strong physical interactions, the intentionality estimates also predicted responses around the pSTS, with variance jointly explained with a model based on speed variability in the animations. Conclusion. These results reveal brain activities that not only encode a categorical representation of physical and social events, but also capture a gradient in a unified psychological space.

Funding: NSF BCS-1658560 to EG and BCS-1658078 to JP and NSF BCS-2142269 to HL

SATURDAY MORNING POSTERS IN PAVILION

Saturday, May 18, 8:30 am – 12:30 pm, Pavilion

Visual Search: Attention, phenomena 1

23.401 TOP-DOWN INSTRUCTIONS INFLUENCE THE ATTENTIONAL WEIGHT ON COLOR AND SHAPE DIMENSIONS DURING REDUNDANT SEARCH

Zixu Gong[1] (zixug2@illinois.edu), Zoe (Jing) Xu[1], Alejandro Lleras[1], Simona Buetti[1]; [1]University of Illinois Urbana-Champaign

Previous research showed that targets defined redundantly along both color and shape dimensions are found faster than those defined by only color or only shape. In redundant situations, target-distractor distinctiveness is a sum of distinctiveness along the two individual dimensions (Buetti et al., 2019). Xu et al. (2023) computationally showed that this composition is influenced by top-down factors: when participants were instructed to search for a target color in redundant search trials, they allocated more attentional weight to the color dimension (and less weight to the shape dimension) compared to when instructed to search for a target shape, which then impact their redundant search efficiency. However, Xu et al. (2023) did not confirm if participants adhered to the given instructions throughout the experiment. It's plausible that over time, participants recognized the consistent redundancy across dimensions, leading them to utilize both features in their search. Our study introduces a more robust top-down manipulation to address this potential shortcoming. In Experiment 1A-C, participants were instructed to search for a target color, while the target differed from distractors in either color, or both color and shape, with the two types of trials intermixed. Experiment 2A-C replicated this setup, focusing on shape. This design ensures that only the instructed dimension consistently indicated the target from trial to trial. We compared participants' search efficiency on redundant trials (which were the same for Experiments 1 and 2) when they were instructed to rely on the color (Experiment1) or shape (Experiment2) dimensions. Results replicated previous findings, showing that people attended more the instructed dimension, which then determined the extent to which they utilized information from the given dimension in redundant search scenarios. Our study reaffirms the role of top-down influences in redundant search, highlighting how attentional weighting impacts the integration of distinctiveness signals from different visual dimensions.

23.402 WHERE YOU ATTEND IS WHERE YOU CLICK: ATTENTION GUIDES SELECTION ACTIONS IN VISUAL FORAGING WITH CONJUNCTION OBJECTS

Jan Tünnermann[1] (jan.tuennermann@uni-marburg.de), Anna Schubö[1]; [1]Philipps University Marburg

Visual foraging, where participants gather items through clicks or touches, gained popularity as a naturalistic paradigm in visual search and selective attention research. Unlike single-target search tasks with discrete trials and button responses, foraging involves multiple targets, continuous interactions with the search environment, and naturalistic goal-directed actions. Notably, the latter offers an opportunity to investigate the interplay between selective attention and actions, a characteristic of foraging not leveraged until now. In the present tablet-based study, we asked how selective attention influences the collecting actions in visual foraging. Earlier foraging experiments with feature conjunctions revealed same-type selection runs. For instance, with red circle and green square targets among green circle and red square distractors ("conjunction foraging"), foragers adhered to one target type (e.g., green squares) for extended runs to avoid costly switches. For the present experiment, we introduced a novel stimulus: "conjunction objects". Targets were composed, for instance, of red rectangular parts joined with green half-circles, and distractors had opposite arrangements. Anticipating the robust same-type run effect seen in conjunction foraging, we expected foragers to repeatedly attend to the same object parts, a bias that should be reflected in the selection actions. Indeed, we found such a bias, with an over-proportional number of stylus clicks landing on the same object parts. However, we also observed a bias toward the center of mass of the overall objects. Hence, the attended location might be fed into motor control, where it combines with motoric influences such as stable contact point selection. We argue that visual foraging is a fruitful tool for examining the relationship between attention and action. In high-paced foraging selection sequences, occasional errors bear minimal consequences, allowing organisms to program motor responses based on attentional selection without involving decision-level processes that might otherwise diminish the direct impact of attention on actions.

23.403 MANIPULATING PROCESSING BIASES TO MITIGATE THE LOW PREVALENCE EFFECT

Robin Celani[1] (rcelani@rollins.edu), Juan Guevara Pinto[1]; [1]Rollins College

When searching for rare items, targets are missed at a much higher rate than when they are more commonly encountered, this is known as the low prevalence effect (LPE; Wolfe et al., 2005). Previous research has been unsuccessful in eliminating this effect, even when explicitly warning participants about the effect (Wolfe et al., 2007). In this investigation, we attempted to implicitly manipulate participants' processing bias by using the Navon task (Navon, 1977). In the Navon task, participants are instructed to indicate the parity of either a larger or smaller number. As the LPE is characterized by short search times and fewer fixations (Godwin et al., 2015; Peltier & Becker, 2016), we hypothesized that a perceptual bias task may implicitly reduce it by increasing the number of fixations. We predicted that a local processing bias would increase fixations, increase search times, and therefore increase accuracy rates and mitigate LPE. Comparatively, this same task can also induce a global processing bias, which should exacerbate the LPE. Participants completed a Ts/Ls search with target prevalence manipulated between subjects (20% v 80% target-present). The search task was intermixed with Navon trials during each block to ensure participants could not predict whether a search or a Navon trial would appear next, inducing an implicit perceptual bias. The type of perceptual bias (local v global v control) was manipulated within subjects across three blocks of trials. The results replicate the traditional LPE regarding search hits, search times, and fixation counts, with no Prevalence x Type of Bias interaction across any measure. However, we observed a significant main effect of Type of Bias for fixation count, indicating that local processing bias did increase the number of fixations, relative to a global bias. This finding is promising as it suggests that perceptual biases can implicitly influence search behaviors.

Rollins College 2023 Student Faculty Collaborative Scholarship Program

23.404 THE TARGET PREVALENCE EFFECT IS MITIGATED BY LESS RESOURCE DEMANDING STIMULI

Brandon Eich[1] (beich1@lsu.edu), Juan Guevara Pinto[2], Melissa Beck[1]; [1]Louisiana State University, [2]Rollins College

Although the effect of target prevalence on visual search performance is well documented, the impact of attention use in different search environments on the target prevalence effect is not well understood. Therefore, it is important to better understand what search tasks and environments are appropriate for generalizing the typical target prevalence effect. Low target prevalence typically leads to earlier search termination and increased distractor decisions. This pattern of results could vary due to the amount of attention resources needed or how attention is allocated in different search environments. For example, nature scenes require fewer resources to process and lead to a broader allocation of attention than urban scenes and, therefore, could have a different target prevalence effect than urban scenes. To examine how search environments with different attention use and allocation impact the target prevalence effect, participants detected a rotated nature or urban scene in a rapid serial visual presentation (RSVP) stream of upright scenes from the same category. The rotated scene was either rarely (10%) or often (90%) present. Nature scenes produced fewer misses and a smaller target prevalence effect for the miss rate than urban scenes. This difference could be due to a lower need for attentional resources or a broader allocation of attention for nature scenes. However, on a subset of trials, participants also detected peripheral items, and there was no difference in peripheral item accuracy between nature and urban scenes, suggesting no difference in attention allocation. Therefore, it is more likely that the need for fewer attentional resources to process nature scenes led to better performance and a smaller target prevalence effect. This conclusion supports the idea that visual search environments that require fewer attentional resources lead to a smaller target prevalence effect.

23.405 LOCATION CUEING FROM COLOR DISTRIBUTIONS

Philippe Blondé[1] (phblonde@hi.is), Sabrina Hansmann-Roth[1], David Pascucci[2], Árni Kristjánsson[1]; [1]Icelandic Vision Lab, School of Health Sciences, University of Iceland, Reykjavík, Iceland, [2]Laboratory of Psychophysics, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

The visual system constantly extracts statistical regularities from the current environment to predict future visual input. Color is a prominent visual feature and a good candidate for extracting such regularities. Here we investigated whether a color ensemble can cue statistical regularities in target location. We assessed how associations between a target's location and the average and variance of different color distributions affected visual search. In our first experiment, 26 participants (61.54% female, 29 ± 10 years old) completed an odd-one-out task, looking for an oddly colored target in a 6 x 6 set of 36 colored diamonds. The targets and distractors came from two distributions with different color averages. Each distribution was associated with different target location probabilities, with an 80% chance that the target would appear on the left side of the display for one distribution, and on the right side for the other. Participants gradually became faster at detecting the target when it appeared on the high probability side of the display, showing how observers learned the association between color distribution and location probability. In a second experiment, 18 participants (44.4% female, 21 ± 2 years old) completed the same task but now the two distributions had the same mean color and different variance. This time, there was no evidence that observers learned the association between distributions and probable location. These results show that statistical properties of color distributions are not used in the same way to extract regularities, with color average being an informative property while variance is not. It is also possible that the variance differences between the two distributions of experiment 2 were too small to distinguish them. Additional experiments will be performed, using larger variance differences to increase their discriminability.

23.406 PROPORTION OF COLOR PROPORTIONALLY INFLUENCES SEARCH GUIDANCE FOR COLOR CONJUNCTIONS AND REAL-WORLD OBJECTS

Jessica N. Goetz[1] (jessica.goetz@ucf.edu), Mark B. Neider[1]; [1]University of Central Florida

Color is widely recognized as a key guiding feature during search for objects defined by one or two colors. However, real-world objects are oftentimes defined by many colors in varying proportions. In three experiments, we examined how color proportions directly affect guidance. Experiment 1 compared guidance for two colors utilizing color conjunction stimuli distributed as 70%/30% or 90%/10% with the predominant color denoted as primary and the second predominant color denoted as secondary. We presented three key distractor types: swapping primary and secondary (inverted), primary only (primary), and non-target colors (neutral). For the 70/30 conjunction, proportionally more first saccades were directed to inverted over primary distractors (p = .02). Conversely, for the 90/10 conjunction, more first saccades were directed to primary compared to inverted distractors (p < .001). Overall, there were significantly more first saccades to color-matching distractors than neutral distractors (all ps < .001). Experiment 1's results indicate that when the color proportion within an object is sufficient, both colors guide attention; when the proportion is insufficient, the predominant color guides attention. Experiments 2 and 3 examined the effect for real-world objects by using a color histogram to quantify the color proportions and exact color coordinates for each object. To generate color-matching distractors, we grouped distractors with targets by color via a nearest neighbor machine learning algorithm. Keeping lightness constant, we transformed the distractor colors to the target's primary or secondary color. To prevent guidance by shape, a recolored target object replaced the inverted distractor. In Experiment 2, we confirmed that search was guided by the primary color for real-world objects (p < .001). In Experiment 3, proportionally more first saccades were directed to primary over secondary over neutral distractors (all ps < .001). Overall, the data suggest that color proportions affect guidance in real-world objects.

Natural Sciences and Engineering Research Council of Canada

23.407 SEARCHING FOR THE ALERTING EFFECT: THE OPTIMAL SOA IS LONGER IN COMPOUND – THAN IN SIMPLE – SEARCH TASKS.

Nadja Jankovic[1] (njankovi@sfu.ca), Amanjot Grewal[1], Evan Caldbick[1], Vincent Di Lollo[1], Thomas M. Spalek[1]; [1]Simon Fraser University

Simple visual search involves the single step of finding a target (e.g., a red ring) among a set of distractors (e.g., green rings). In contrast, compound search involves two steps. For example: i) find the target ring in the display, and ii) identify the orientation of a line segment inside the target. Performance is known to be facilitated when the search display is preceded by an alerting stimulus, such as a brief brightening of the screen. Until recently, alerting was studied using only "simple" tasks. In these studies, the optimal stimulus-onset asynchrony (SOA) between the alerting stimulus and the search display was found to be about 100 ms. Recent work that employed a 100-ms SOA showed that while alerting does occur in simple search, it does not occur in compound search. A temporal-period model was proposed to account for these findings. In the present work, we varied the SOA to test predictions from that model. In Experiment 1, we used a compound task with two SOAs: 100 and 150 ms. The results confirmed the absence of alerting when the SOA was 100 ms and revealed significant alerting when the SOA was 150 ms. To examine the time course of the effect, Experiment 2 included four SOAs: 50, 100, 150, and 200 ms. An alerting effect was found when the SOA was 150 ms, as in Experiment 1, but not when it was 100 or 200 ms. When the SOA was 50 ms, the alerting stimulus led to worse performance than when the alerting stimulus was absent. The temporal-period model was revised to account for this pattern of results.

23.408 WHAT MAKES AN OPTIMAL VISUAL SEARCHER? INSIGHTS FROM POST-TRIAL MEMORY PROBES

Tianyu Zhang[1] (zhang.11476@osu.edu), Andrew B. Leber[1]; [1]The Ohio State University

People vary greatly in how strategically they guide their attention during visual search tasks. To understand why, recent research has explored the factors that distinguish optimal from suboptimal searchers, but this question remains largely unresolved. One possible explanation is that individuals focus on different aspects of the task stimuli, leading them to develop divergent strategies. We investigated this idea by using the Adaptive Choice Visual Search (ACVS) paradigm, which is designed to measure visual search strategy (Irons & Leber, 2016). In the ACVS, the optimal strategy is to search through the less numerous of two color subsets. We hypothesized that optimal participants who focused more on the ratio of color subsets would be more likely to exhibit optimal performance. Alternatively, both optimal and suboptimal participants are equally likely to process the ratios, but only the optimal participants exploit that information. To test this, we assessed participants' memories of several display properties. In Experiment 1, we utilized after-trial probe questions to examine which aspect of the stimuli that participants attended to during the search task. After 30% of search trials, participants were questioned about various target properties, including the color subset (red/blue), subset ratio (large/small), location, and digit. Results showed that only the accuracy for ratio probes strongly correlated with individuals' optimality in the ACVS (r = .55, p < .001). In Experiment 2, we explicitly informed participants which feature would be probed at the beginning of each block. Preliminary evidence suggests that participants are most optimal in blocks when they expect to be probed with ratio questions. These findings confirm that participants who attend display characteristics most relevant to the optimal strategy are more likely to use the optimal strategy. Furthermore, the results suggest that encouraging attention to such information makes participants more likely to adopt the optimal strategy.

This work is supported by NSF BCS-2021038 to ABL

23.409 TRAINING ON A DIFFICULT SEARCH TASK IMPROVES UNTRAINED EASIER SEARCH AS WELL

Johan Hulleman[1] (johan.hulleman@manchester.ac.uk), Jason Taylor[1], Gorana Pobric[1]; [1]University of Manchester

We conducted a large-scale (N=110) experiment into the effects of training on a difficult visual search task. On Day 1, participants performed a baseline session with easier search (T vs. L) and difficult search (square with small square in left-top corner vs. squares with small square in different corner). Display sizes were 8 and 16 items and target prevalence was 50% for both tasks. On Days 2 – 6 participants were trained on the difficult search task. Display sizes were unchanged, but target prevalence was 16.7%. On Day 7, participants performed a post-training session with the same search tasks as on Day 1. Eye movements were recorded for all sessions. We found substantial positive training effects: reaction time gains from Day 2 to Day 6 were ~500 ms and error rates on present trials more than halved. Despite higher target prevalence during baseline and post-training sessions, performance clearly improved on both the difficult search task (reaction times up to 1000 ms faster and error rates on present trials again more than halved) and the easier search task (reaction times fairly constant, but error rates on present trials halved). The eye tracking data revealed that training had increased the size of the functional visual field not only in the difficult search tasks, but even in the (untrained) T vs L task. Training also reduced target decision time (TDT), the period between a fixation landing on the target and the present response. For the difficult search tasks TDT was reduced by several 100 ms, whereas for T vs L the reduction was around 20 ms. The bigger improvement in difficult search reflects that recognizing the target is harder in this task. Consequently, the TDT is longer and the scope for improvement larger. Training gains on difficult search transfer to untrained easier search.

This work was funded by the Defence Science and Technology Laboratory (Dstl) and delivered through the Human Social Science Research Capability (HSSRC).

23.410 SEARCHING FOR INTERACTIVE PEOPLE: VISUAL SEARCH ASYMMETRY IN DYNAMIC DYADS

Jinglan Wu[1] (delta@zju.edu.cn), Qikai Zhang[1], Zaifeng Gao[1], Mowei Shen[1]; [1]Zhejiang University

Social interaction pervades our daily experiences and has been a focal point of numerous studies in recent years. A pivotal inquiry revolves around whether social interaction information confers a perceptual advantage in visual search tasks. Existing evidence has previously suggested the absence of such an advantage, albeit constrained by a notable limitation—the reliance on static images that may inadequately capture social interaction nuances. We introduced dynamic biological motions into the visual search asymmetry paradigm to address this issue. And by mismatching the interactive dyads, we created the non-interactive dyads that are also presented in a face-to-face configuration. The deliberate creation facilitated the isolation of the social nature aspect for an independent investigation into the searching advantage attributed to social interaction. In experiment 1, participants were presented with sets of 3 or 5 face-to-face dynamic dyads, with one dyad differing in interactive properties from the rest. They were required to search for an interactive dyad among several non-interactive dyads or vice versa. Analysis of reaction time data revealed that participants exhibited heightened efficiency in detecting interactive targets over non-interactive ones across both set size conditions. To ensure the robustness of our findings, we expanded the range of set sizes from 3 and 5 to include 3, 4, and 5 conditions, and replicated the experiment, consistently yielding congruent results. These findings support the hypothesis that social interaction information confers an advantage in visual search. By introducing dynamic elements and disentangling the essence of social interaction within the experimental setup, our study provides a nuanced understanding of the inherent advantages associated with perceiving social cues in visual contexts.

23.411 WHAT, WHERE, WHEN DID I FIND THIS? ASSOCIATIVE LEARNING IN HYBRID SEARCH.

Iris Wiegand[1] (iris.wiegand@ru.nl), Jeremy M. Wolfe[2], Joseph R. Maes[1], Roy P.C. Kessels[1]; [1]Donders Institute for Brain, Cognition and Behaviour, [2]Brigham and Women's Hospital, Harvard University

In "hybrid" visual and memory search, observers look for multiple, previously memorized target objects among distractors. Hybrid search is akin to many real-world searches, such as looking for items on your mental shopping list in the grocery store. Thus, hybrid searches occur in spatial and temporal contexts that we encounter repeatedly. In several experiments, we investigated if observers would incidentally learn and utilize spatial and temporal associations in hybrid search. Specifically, we examined learning of four different types of regularities: 1) target item sequences (e.g., the banana always follows the yoghurt), 2) target location sequences (e.g., a target in the lower left corner always follows a target in the upper right corner), 3) target item-location associations (the banana is always in the upper right corner), and 4) target item-location sequences (the banana in the upper right corner always follows the yoghurt in the lower left corner). Learning would be reflected in a decrease in search times. Our results show only weak incidental learning for the temporal sequences of target items or target locations alone, even after many repetitions of the sequence. By contrast, learning of target item-location associations was fast and effectively reduced search times. Furthermore, the experiments show a reliable effect of temporal sequence learning for target item-location associations. These findings suggest that spatiotemporal learning in hybrid search is hierarchical and conditional: Only if spatial and non-spatial target features are bound do temporal associations bias attention, pointing the observer to the task-relevant features expected to occur next.

23.412 WE NEED AN ONTOLOGY OF VISUAL SEARCH TASKS

Keith Racioppo[1] (kraciop1@gmail.com), Benjamin J. Tamber-Rosenau[1]; [1]University of Houston

Visual search tasks are widely used in the laboratory, yet also generalize to important real-world tasks. While some real world tasks (e.g., medical imaging, security screening) are the topic of much research, connecting the cognitive expectations of in-lab tasks to real world scenarios is hampered by the lack of consistent categories and terminology to describe important parameters of visual search. Thus, we here propose an ontology to categorize visual search based on conceptual differences across tasks, hopefully easing generalization of predictions between laboratory tasks and to real world behavior. This ontology initially stemmed from a literature review of research on attentional templates, which revealed that researchers categorize visual search tasks and the accompanying attentional templates using idiosyncratic categories that are often only implicitly defined. In addition to the aforementioned difficulties that stem from this problem, the lack of agreed-upon categories and terms leads to difficulty in comparing between existing studies or identifying gaps in existing knowledge because the mappings from task-specific parameters or conditions to abstract concepts can be unclear or inconsistent. We suggest that these problems might be ameliorated by establishing a systematic classification system to help better organize the existing visual search literature and new studies. These distinctions are based on: (1) the number of target-defining dimensions, (2) number of targets sought per dimension, and (3) disjunctive vs. conjunctive search. These distinctions will allow for more nuanced comparisons between visual search tasks because task-specific parameters will be mapped back to standardized conceptual categories. The proposed ontology will also help to identify potential areas in the visual search literature that are lacking and may require further investigation. Moreover, the systematicity of the ontology makes it readily expandable: for example, further divisions could be identified based upon spatial vs. temporal search.

23.413 ESTIMATING CAPACITY LIMITATIONS IN ENSEMBLE AVERAGING USING SET SIZE MANIPULATIONS AND INDIVIDUAL DIFFERENCES IN SEARCH PERFORMANCE

Laramie Starling[1] (lls@usf.edu), Chad Dubé[1]; [1]University of South Florida

When confronted with a group of stimuli which are similar along one or multiple feature dimensions, the ability to extract a statistical summary representation (SSR) of the group is thought to be one way of decreasing cognitive load. The broad set of processes by which this summarization occurs is known as ensemble coding. Predominant theories of ensemble coding state that the SSR is calculated before selective attention can take hold, thus the process utilizes each stimulus in the group. Other theories state that attention may select a subsample of stimuli from which to form the SSR. This subsample would be subject to the capacity limits of selective attention. To investigate these possibilities, we tested capacity limits by manipulating set size on two visual search tasks and an ensemble averaging task, using similar stimuli in each. This manipulation allowed us to view potential set size effects on ensemble averaging as well as compare individual differences in set size effects for the search tasks and the ensemble averaging task. We established evidence of a relationship between individual differences in search performance and ensemble averaging performance. These findings suggest that SSRs may not be calculated preattentively but are influenced by attentional selection of stimuli.

23.414 LANGUAGE EXPERIENCE MAY MODULATE ATTENTIONAL DISENGAGEMENT TO SCENE GRAMMAR INCONSISTENCIES DURING FREE-VIEWING

Naomi Vingron[1] (vingron@psych.uni-frankfurt.de), Melissa Vo[1]; [1]Goethe University Frankfurt, Department of Psychology, Scene Grammar Lab, Germany

Orienting oneself in real-world scenes relies on a set of rules (“scene grammar”), which support object localization and identification (e.g., knowing that a spatula goes neither in a toaster nor a bathroom). The acquisition of these rules might be linked to language, drawing on many of the same cognitive mechanisms. As such, scene grammar inconsistencies (e.g., a pot on the sofa) may impair object processing as viewers need to disengage attention and resolve ambiguity resulting from the unexpected element of the scene. Psycholinguistic research suggests that multilingual individuals face similar challenges as they experience concurrent activation of all known languages. This may increase resilience to distraction from task-irrelevant information in a domain-general manner. We conducted two experiments in which multilingual participants either viewed realistic indoor scenes freely for five seconds (Experiment 1; N = 27) or performed a word-cued visual search task (Experiment 2; N = 24) while their eye movements were recorded. Across both experiments scenes contained semantic, syntactic or no violations. Experiment 2 also involved a target present/absent decision. Additionally, participants responded to a language background questionnaire. During free-viewing semantic violations were fixated earlier than syntactic violations and for longer than syntactic violations and neutral areas in the scene. In the visual search task, this effect was attenuated when the target was absent, and participants showed no indication of distraction from either type of inconsistency when there was a target to be found. During free-viewing bilingual experience was associated with shorter dwell times on the inconsistent object. More specifically, individuals with higher second language proficiency and earlier age of acquisition were able to move past scene grammar inconsistencies more efficiently. These results suggest that while task demands guide attention, there may be some influence of individual differences in language experience on attentional disengagement from object-scene inconsistencies.

Acknowledgements: This work was supported by SFB/TRR 26 135 project C7 to Melissa L.-H. Võ and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project ‘The Adaptive Mind’) as well as the Humboldt postdoctoral research fellowship awarded to Naomi Vingron.

Saturday, May 18, 8:30 am – 12:30 pm, Pavilion

Attention: Capture

23.415 MEMORY-BASED ATTENTIONAL CAPTURE IS ONE-SHOT

Dirk Kerzel[1] (dirk.kerzel@unige.ch), Werner X. Schneider[2]; [1]University of Geneva, [2]University of Bielefeld

It was observed that search-irrelevant features stored in working memory guide attention. That is, stimuli matching the stored feature capture attention in a search task occurring between encoding of the irrelevant feature and memory test. We asked whether attentional capture during search occurred only to a single stimulus or whether it would occur continuously to several stimuli. Participants memorized the shape of a colored object before they searched a rotated T among randomly rotated Ls. The set size of the search display varied between 8, 12, and 18 items. The color of half the stimuli matched the irrelevant color of the memorized shape, whereas it was different for the other half. In one experiment, we observed that RTs increased by 59 ms when the target was among the stimuli in a non-matching color compared to when it was among the stimuli in a matching color. However, the search slopes were the same (16 ms/item). When we changed the target shape to make search more serial, the effect of color match was 141 ms and the search slope was larger when the target was among nonmatching than matching stimuli (48 vs. 60 ms/item). However, the difference in search slopes was limited to the large set sizes (12 and 18 items). Overall, an irrelevant feature in working memory guides attention to matching stimuli, but not continuously. Rather, memory-based capture is mostly one-shot.

23.416 ATTENTIONAL CONTROL THROUGH COLOUR-LOCATION ASSOCIATIONS INDUCES CONTINGENT CAPTURE

Samantha Joubran[1] (sjoubran@uoguelph.ca), Naseem Al-Aidroos[2]; [1]University of Guelph, [2]University of Guelph

Is attention automatically captured to the location of salient stimuli, or is capture under our control? The best evidence that capture can be controlled comes from contingent capture in attention cueing tasks: When looking for a visual target (e.g., an orange target), distracting stimuli only capture attention if they resemble the target (e.g., a non-predictive orange onset pre-cue). Put differently, an observer's attentional goals during the target display determine which types of features capture attention in the cue display. Our recent work assessed another level of control and found that target-related goals also determine where attention goes in response to the cue. In this work, participants completed a spatial-cueing task where on every trial a target was presented to the left and right of fixation, and a separate, coloured stimulus indicated which target the participant should report (e.g., orange meant report left target; green report right). Thus, the target display created an association between colours and shifting attention to the left or right. The non-predictive pre-cues captured attention to the location associated with their colour (e.g., orange cues captured attention to the left location) regardless of where the cues physically appeared in space. Does this type of attentional control through colour-location associations also induce contingent capture? To test this, we presented pre-cues in colours such as purple, which were completely task-irrelevant in that their colour had no association with space or target identification. While cues with location-associated colours continued to capture attention to their associated locations, no cues captured attention to their physical location, including task-irrelevant (purple) cues. Thus, attentional capture was contingent on cues matching one of the location-associated colours. These findings are consistent with a single attentional control process that determines both what types of stimuli capture spatial attention, and where attention goes in response to these stimuli.

Funded through NSERC

23.417 BEYOND DISTRACTION: EXPLORING THE CONDITIONS OF TASK-IRRELEVANT FEATURE REINSTATEMENT AND WORKING MEMORY-BASED CAPTURE DURING VISUAL SEARCH

Jessica Kespe[1] (jkespe@uoguelph.ca), Naseem Al-Aidroos[1], Rory Lippert[1]; [1]University of Guelph

Objects actively stored in visual working memory (VWM) for a recall task, produce VWM-based attentional capture. However, capture often only occurs for the object's task relevant features. For example, when shown a red circle and asked to later report its shape, in an intervening singleton-search task, circle distractors (task-relevant feature) preferentially capture attention, and red distractors (task-irrelevant feature) do not. Interestingly, for an object's features to be stored in VWM, they do not need to be shown, but can also be reinstated from long-term memory (LTM); for example, if the shape was previously associated with a colour. Unlike shown features, reinstated features in VWM appear to preferentially capture attention even when task irrelevant. To investigate this difference, we examined how capture by reinstated features is influenced by memory goal (remember vs. search), and search goal (singleton vs. object). Participants memorized a set of objects with specific colours, and then completed a dual task in which they were instructed to either remember or search for one of the objects' shapes in any colour. On each trial, one memorized object was brought into VWM, and participants had to complete a shape-singleton search task followed by either an object recall, or an object search task. We found that, regardless of whether participants were instructed to remember or search for the object shape, the irrelevant reinstated colour did not capture attention more than any other colour in the shape-singleton search task. However, in both object tasks, participants were quicker to locate and report the object when it appeared in the irrelevant memorized colour. These results suggest that under both remember and search conditions, the task-irrelevant feature was reinstated, actively maintained in VWM, and could bias attention. But merely actively maintaining the irrelevant reinstated feature did not mean it always captured attention.

NSERC

23.418 DO PHYSICAL EFFORT AND ELECTRICAL STIMULATIONS SIMILARLY AFFECT ATTENTIONAL CAPTURE?

Laurent Grégoire[1] (lgregoire1@tamu.edu), David Lee[1], Brian Anderson[1]; [1]Texas A&M University

Visual attention prioritizes stimuli previously associated with punishment, even if they are non-salient and task-irrelevant. A recent study demonstrated that punishment-related attentional capture could be suppressed when participants were strongly encouraged to resist attentional capture. The present study aimed to determine whether a similar suppression effect would be observed when physical effort is used as punishment instead of electrical stimulations. Participants first performed a training phase involving fixating a shape-defined target among distractors, one of which was uniquely colored. The color of the distractor predicted the possibility of exerting physical effort, which was required following thirty-three percent of trials in which the effort-associated distractor appeared. Participants could avert physical effort via fast and accurate performance (which would be facilitated by ignoring the effort-associated distractor). Results revealed no suppression of punishment-related attentional capture in the training phase as well as in a subsequent task (similar to training) in which no effort was exerted. These findings suggest that mechanisms underlying physical effort do not affect attention similarly as electrical stimulations.

23.419 FROM ATTENTION CAPTURE TO SOCIAL COORDINATION: "ABRUPT ONSET" RALLIES JOINT PLANNING

Shaozhe Cheng[1] (shaozhecheng@zju.edu.cn), Jingyin Zhu[1], Jifan Zhou[1], Mowei Shen[1], Tao Gao[2]; [1]Zhejiang University, [2]UCLA

The abrupt onset of a new object is known to automatically capture human attention. In this study, we explore whether this visual cue's impact extends beyond individual attention, aiding in social coordination and joint planning. A key challenge in social coordination is the presence of multiple equally desirable targets, with the need for the team to commit to just one. But which one? We tested the hypothesis that in such joint planning tasks, where cooperation is visually guided, all players can spontaneously use the abrupt onset as a rallying point to jointly pursue the newly present target, instead of the old target that has lost its saliency. We used a collaborative Pac-Man task, where pairs of participants could only 'eat' an object on a 2D map by jointly navigating to it. At any moment, there were always two objects: one left-over from the previous navigation, and a new one that appeared abruptly at the moment participants jointly ate a current object. Results showed that participants were biased toward reaching the new object, even when the distance to both the old and new objects was controlled. We then explored whether this effect is specific to establishing a rally point in joint planning or it impacts group behavior through individual planning. An individual Pac-Man task showed an opposite bias: participants tended to eat the old object. Furthermore, in a relay task where two participants took turns controlling a single PacMan, as opposed to simultaneously controlling two PacMans, they again demonstrated a bias toward the old object. Collectively, these findings suggest that the unique challenges of group coordination allow the effect of abrupt onset objects to transcend individual attention, making visual salience a critical factor in facilitating the establishment of joint attention. This provides insight into the visually grounded nature of human cooperation.

23.420 DOES CONTINGENT CAPTURE OCCUR IN DRIVING SCENES?

Rachel A. Eng[1] (engr@uoguelph.ca), Naseem Al-Aidroos[1], Lana M. Trick[1]; [1]University of Guelph

Contingent capture theory suggests that only stimuli consistent with the observer's internal goals will capture attention. For example, when looking for a red target, red stimuli should automatically capture observers' attention. It is unclear whether contingent capture occurs in complex real-world scenes. Arexis et al. (2017) investigated contingent capture using a search task where participants viewed photographed driving scenes in which a single red letter appeared at random locations. Their task was to decide whether the red target letter was a T or an L. A GPS navigation system image was also shown in the bottom right corner of the display. The GPS either had a blank screen, a red-coloured route (goal-relevant distractor colour) or a green-coloured route (goal-irrelevant distractor colour). The GPS appeared 1 second in advance of the driving scene and letter, acting as a pre-search display. If contingent capture occurs, participants should be slower to respond when the GPS showed a red route (goal-relevant distractor colour). However, goal-relevance had no effect, perhaps because the GPS pre-search display appeared so far in advance of the search display (1 s). In the present study, we manipulated the presentation duration of the pre-search display (0 ms, 100 ms, 1 s) and instructed participants to ignore the GPS. We predicted that goal-relevance would have an effect, but only when there was insufficient time to disengage attention from the GPS distractor before search display onset (the 0 and 100 ms conditions). Results call into question the contingent capture hypothesis within the context of real-world scenes.

We would like to thank NSERC for funding this research.

23.421 NEURAL EVIDENCE FOR ATTENTIONAL CAPTURE BY SALIENT DISTRACTORS

Rongqi Lin[1] (linrongqi.tenuosi@gmail.com), Xinyu Li[2], Ole Jensen[3], Jan Theeuwes[4], Benchi Wang[1]; [1]South China Normal University, [2]Zhejiang Normal University, [3]University of Birmingham, [4]Vrije Universiteit Amsterdam

Salient objects often capture our attention, serving as distractors and hindering our current goals. It remains unclear when and how salient distractors interact with our goals and our knowledge on the neural mechanisms responsible for attentional capture is limited to a few brain regions recorded from non-human primates. Here we conducted a multivariate analysis on human intracranial signals (18 patients) covering most brain regions, and successfully dissociated distractor-specific representations from target-arousal signals in the high-frequency (60-100 Hz) activity. We found that salient distractors were processed rapidly around 220 ms, while target-tuning attention was attenuated simultaneously, supporting initial capture by distractors. Notably, neuronal activity specific to the distractor representation was strongest in superior and middle temporal gyrus, amygdala, and anterior cingulate cortex, while there were smaller contributions from parietal and frontal cortices. These results provide neural evidence for attentional capture by salient distractors engaging a much larger network than previously appreciated.

23.422 ASSESSING CONNECTIVITY BETWEEN BRAIN REGIONS DURING OBJECT CATEGORY-TUNED ATTENTION AND SPATIAL DISTRACTION

Ehi Okojie[1,2] (okojie.6@osu.edu), Yong Min Choi[1], Blaire Dube[3], Julie D. Golomb[1]; [1]Department of Psychology, The Ohio State University, [2]Discovery PREP, The Ohio State University Medical Center, [3]Department of Psychology, Memorial University

To prioritize goal-relevant visual information, our brain is able to filter out less relevant visual input. Specific object categories, such as faces and scenes, are processed by specific regions in the brain. These brain areas can act as attentional filters biased toward processing object category information relevant to our goals. In our everyday life, there are distractions that can capture our spatial attention. What happens to these category-tuned attentional filters when we get distracted? A recent fMRI study by (Dube et al., 2022), investigated the effects that visual distraction has on category-tuned filters in the ventral visual cortex (Fusiform Face Area [FFA] and Parahippocampal Place Area [PPA]) and discovered a novel consequence of distraction on these filters which regulate category-specific object processing. Participants in this study viewed hybrid face/house images and were told to attend to either faces or houses. The presence of a salient distractor disrupts these filters, such that our brain incidentally processes the goal-irrelevant category more than during distractor-absent trials. However, it is still unclear which brain areas may be modulating this filter disruption. To learn more about the neural pathways implicated in this filter disruption, here we expand on the previous study by conducting functional connectivity analyses among particular brain regions during distractor-present versus distractor-absent trials. We compared connectivity between the category-tuned areas (FFA or PPA) and the early visual cortex (EVC) showing that in the absence of a distractor, task-relevant category-tuned areas have significantly higher connectivity with the EVC during their attend-preferred conditions (e.g. FFA-EVC during attend-face). In the presence of a distractor, we observe some differences in connectivity patterns between the category-tuned areas, EVC, and fronto-parietal attentional control regions. These results suggest that connectivity patterns with the early visual cortex may inform us about the mechanism underlying category-selective attentional filter disruption.

NSF 1848939 (JG), NIH R01-EY025648 (JG), NIH 5R25GM089571 (PREP), NSERC PDF (BD)

23.423 CUES IMPROVE VISUAL WORKING MEMORY BUT FAIL TO COUNTERACT THE EFFECTS OF SALIENCE

Martin Constant[1] (martin.constant@unige.ch), Dirk Kerzel[1]; [1]University of Geneva

It has recently been shown that salience can have a massive impact on visual working memory (Constant & Liesefeld, 2021; https://doi.org/gjk9jh). That is, in an array of vertical bars, the color of strongly tilted bars was recalled better than the color of less tilted bars. This effect is also hard to erase or overrule with manipulations of relevance, even at long encoding times (Constant & Liesefeld, 2023; https://doi.org/gr6xzr). Here, across three pre-registered experiments, we tried to determine whether feature pre-cues, spatial pre-cues and retro-cues can counteract the effect of salience. Pre-cues were presented about 500ms before onset of the memory display and retro-cues 750ms after its offset. In all experiments, the memory array consisted of three tilted bars (12°, 28° and 45°) among 33 vertical bars and was presented for 350ms. Participants recalled the color of one of the tilted bars using a colorwheel. In the experiment with feature pre-cues, the target orientation was pre-cued by a black tilted bar in the center. While the cue improved overall memory performance when comparing valid (66%) to invalid (17%) trials, the effect of salience remained in both conditions. In the experiment with spatial pre-cues, a 66%-valid black circle was briefly presented at the future location of one of the targets. Again, overall memory performance improved but the effect of salience remained. In the experiment with retro-cues, the cue was either neutral (vertical bar presented centrally) or 100% valid (tilted bar presented at the location of the target). Once again, overall memory performance improved, but the effect of salience remained virtually unaffected. These results suggest that the effects of salience on visual working memory may not be purely attentional. That is, lower-salience targets were probably found and attended when they were cued, yet performance remained worse for these targets.

23.424 DOES ANTICIPATED EFFORT MODULATE ATTENTIONAL BIASES FOR AVERSIVELY CONDITIONED STIMULI?

Molly McKinney[1] (mollyr.mckinney@tamu.edu), David Lee[1], Brian Anderson[1]; [1]Texas A&M University

Our attention prioritizes stimuli that are of importance to us, whether that be stimuli to approach or avoid, and experience with stimuli associated with the exertion of effort has been shown to be prioritized in the same value-driven manner. The context in which we interact with these stimuli can also play a role in whether they are prioritized. Previous work has shown modulatory effects of a threatening context on value-driven attention. In an ongoing study, we are seeking to better understand how independent value-driven sources interact to determine attentional priority, particularly whether attentional bias toward aversively conditioned stimuli would be modulated by contexts that are predictive of different subsequent workloads. In a training phase, participants fixate one of three color-square targets (red, blue, or green). One of the three colors is paired with a shock (CS+) immediately after fixating, while for the other two colors, no shock occurs (CS-). Then in the following test phase, participants see one of two context background images (rock, forest) and are told that one predicts the need to exert high effort (manipulated via force applied to a hand dynamometer), while the other predicts no effort requirement. Overlaid on the background image, prior to the exertion of any physical effort, is a search array containing a circle target and square distractor, which are rendered in the same colors used during training. Preliminary data shows a bias toward the CS+ color, however this effect is unmodulated by expected effort requirements. A subsequent experiment will reverse the role of effort and shock to address whether there is a threshold at which contexts signaling a particular value modulate an existing bias, shedding light on potential factors that can mitigate bias when our goals have changed.

NIH-R01-DA046410

23.425 EXPERIENCE-DEPENDENT DISTRACTOR REJECTION OCCURS RAPIDLY BUT IS DIFFICULT MEASURE

Isaac Savelson[1] (savelson.1@osu.edu), Christopher Hauck[2], Mei-Ching Lien[2], Eric Ruthruff[3], Andrew B. Leber[1]; [1]The Ohio State University, [2]Oregon State University, [3]The University of New Mexico

Humans' ability to resist attentional capture by distracting stimuli appears to rely on learning the specific features of a distractor over the course of repeated encounters with the item. A seminal study by Vatterott and Vecera (2012) found substantial attentional capture when first encountering salient distractor, followed by a complete attenuation of capture as experience with the color singleton increased. Critically, this pattern of capture-then-rejection was repeatedly observed throughout the experiment following changes in the distractor color. Despite the apparent robustness of this effect in the original study, we report three experiments in which we appeared to fail to replicate Vatterott and Vecera's findings. Importantly, these experiments were conducted independently by two groups of researchers using much larger sample sizes than the original study. Using the conventional split-block analysis, we found no evidence for experience-based distractor rejection. That is, we found no rebound of attentional capture when presenting a novel distractor and were unable to detect a significant decrease in capture across block halves. Critically, the apparent replication failure did not appear to be due to a lack of distractor rejection; instead, minimal attentional capture was observed overall, leaving little room for further reduction. However, a finer-grained trial-by-trial analysis revealed that capture did occur at the start of each block, but complete distractor rejection appeared after only ~2 encounters with the distractor – much faster than Vatterott and Vecera originally reported. Further, rejection seemed to be partially robust to changes in distractor color: the initial capture in blocks 2-4 was greatly reduced relative to the first block. These results necessitate an update to our understanding and study of learned distractor rejection, especially regarding the speed at which distractor rejection is learned. Overall, we conclude that learned distractor rejection is more robust than initially believed but can be difficult to measure.

Funding: NSF BCS-2021038 to ABL

23.426 OBJECT STABILITY, ATTENTION, AND TEMPORAL ORDER JUDGMENTS

Ece Yucer[1] (ece.yucer@mail.utoronto.ca), Andrew Clement[2], Jay Pratt[1]; [1]University of Toronto, [2]Millsaps College

Our visual lives are filled with stable and unstable objects, some relatively human-sized like furniture and boxes, others much larger like trees and buildings. Yet not much is known about how the stability of these objects affects the deployment of our attention across the visual field. To address this, we have investigated how perceived stability interacts with our attentional system with a series of experiments using a cueless temporal order judgment task. Participants were presented with stable and unstable objects, separated by small temporal intervals, and asked to report which object appeared first in the display. Each experiment employed a different pair of stable and unstable stimuli, including objects with stability as a feature at the global scale (e.g., large art pieces; Experiment 1), local scale (e.g., traffic cone; Experiment 2), and never-before-seen objects (e.g., NOUN Database; Experiment 3). The use of never-before-seen objects was aimed to provide a control for novelty, as unstable objects may be more novel due to context (i.e., a physically unstable traffic cone) and attended to more. Additionally, Experiment 4 compared an object, an ottoman, that maintained its stability while being presented upright vs. upside down to control for inversion effects. Participants' responses were fit to logistic regression models, and their point of subjective simultaneity was calculated using the fitted model. The results show that stability at the local scale captures attention, and this effect is present even in never-before-seen objects (Exp 3). Likewise, the lack of attentional capture in with the ottoman (Exp 4) suggests that these observed effects are due to the objects' stability rather than their inverted nature. Overall, the results of this study help inform us of the role stability plays in perceiving objects in our visual environments.

23.427 ON THE TIMING OF OVERT ATTENTION DEPLOYMENT: EYE-MOVEMENT EVIDENCE FOR THE PRIORITY ACCUMULATION FRAMEWORK

Mor Sasi[1] (mor.sasi1992@gmail.com), Daniel Toledano[1], Shlomit Yuval-Greenberg[1], Dominique Lamy[1,2]; [1]Tel-Aviv University, [2]Sagol School of Neuroscience, Tel Aviv University

The interpretation of many well-established findings rests upon the assumption that at any given moment attention is automatically allocated to the peak of a priority map, which is determined by physical salience, goals, and selection history. The Priority Accumulation Framework (PAF) challenges this assumption. It suggests that the priority weight at each location accumulates across successive events and that evidence for the presence of action-relevant information contributes to determine when attention is deployed to the location with the highest accumulated priority. Here, we tested these hypotheses for overt attention by recording eye saccades in a free-viewing spatial-cueing task. We manipulated search difficulty (Experiments 1 and 2) and cue salience (Experiment 2). Standard theories posit that when oculomotor capture by the cue occurs, it is initiated before the search display appears; therefore, these theories predict that the cue's impact on the distribution of first saccades should be independent of search difficulty but influenced by the cue's saliency. By contrast, PAF posits that the cue can bias competition later, after processing of the search display has already started, and therefore, predicts that such late impact should increase with both search difficulty and cue salience. The results supported PAF's predictions. Our findings call for a revision of visual search theories that have developed around the concept of a priority map without integrating the insights from research on temporal attention.

Support was provided by the Israel Science Foundation (ISF)

23.428 THE ROLE OF SPATIAL REFERENCE FRAME IN IMPLICIT DISTRACTOR LOCATION LEARNING

Litian Chen[1], Freek van Ede[1], Heleen Slagter[1]; [1]Vrije Universiteit Amsterdam

The ability to ignore distracting information plays an important role in many daily life situations. Recent research shows that when we become familiar with features of a distracting stimulus, such as its likely location, we can better ignore it, allegedly because the brain implicitly learns where in the environment distraction is likely to occur. In this study, we examined whether implicit distractor learning is environment-centered and/or depends on the viewer's perspective. To this end, participants (n=29) performed an additional-singleton visual-search task, displayed on a monitor lying flat on a table. In this task, unbeknownst to the participants, the singleton distractor appeared more often at one of the search locations (high-probable location) during a training phase, to induce distractor location learning. In a subsequent test phase, participants moved 90 degrees around the table, changing their viewpoint, and the distractor regularity was removed. Thus, in the test phase, the distractor could now appear on the previously likely distractor location in the environment, on the previously likely distractor location from their own new viewpoint, or at other previously low probability locations. Initial analyses showed a robust capture effect in singleton-present trials during the training phase, that was significantly reduced when the singleton distractor occurred at the high-probable location, demonstrating distractor-location learning. In the test phase, the effect of the previous distractor location showed trend-level significance, with numerically faster responses when the singleton distractor occurred at the viewer-centered location, but with no clear difference between the environment-centered and the other non-learned distractor locations. These first findings suggest that distractor-location learning may be predominantly viewer-centered, possibly entailing implicit learning of how one should not move their attention from an egocentric perspective.

23.429 TRACKING THE TEMPORAL DYNAMICS OF STATISTICALLY LEARNED DISTRACTOR SUPPRESSION USING SSVEP'S

Dock H. Duncan[1,2], Norman Forschack[3], Dirk van Moorselaar[1,2], Jan Theeuwes[1,2,4]; [1]Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, [2]Institute Brain and Behavior Amsterdam (iBBA), [3]Wilhelm Wundt Institute for Psychology, University of Leipzig, Leipzig, Germany, [4]William James Center for Research, ISPA-Instituto Universitario, Lisbon, Portugal

Visual selection is influenced by our past experience. More specifically, through a process known as statistical learning, participants learn to respond faster to targets when presented at predictable locations while distractors interfere less with visual search when regularly encountered at the same location. While a great deal of recent research has focused on the extent that prior experience influences visual attention, the brain mechanisms by which our past experiences affects attentional selection is still elusive. Recent work has demonstrated that our prior experience modulates pre-stimulus excitability of suppressed or enhanced locations in space, possibly via transient changes in synaptic connectivity in the visual system. However, the precise location within the visual system where these changes occur and how early in visual processing these shifts manifest remain unclear. To address these questions, the present study employed steady state visually evoked potentials (SSVEP's), a technique offering a straightforward measure of attentional processing with excellent temporal fidelity. Participants performed a modified version of the additional singleton task across two EEG sessions with counterbalanced distractor high probability locations. The search stimuli flickered on and off the screen at various high frequency rates, leading to frequency entrainment in the visual system and enabling the tracking of current attentional engagement. Preliminary results show that SSVEP's can be effectively used to track the timecourse of attentional capture; revealing how selection differs in suppressed spatial field as opposed to unsuppressed fields and how this learned suppression changes the temporal character of attentional capture. SSVEP's tied to neutral interstimulus placeholders additionally tracked how statistical learning alters pre-stimulus excitability, demonstrating how learned suppression exerts its influence proactively before stimulus onset.

Saturday, May 18, 8:30 am – 12:30 pm, Pavilion

Attention: Neural mechanisms

23.430 GUIDANCE BY VISUAL AND VERBAL REPRESENTATIONS DURING VISUAL SEARCH

Mikel Jimenez[1] (mikel.jimenez@durham.ac.uk), Jerilyn Harrison[1], Daisy McGonigal[1], Anna Grubert[1]; [1]Department of Psychology, Durham University, United Kingdom

Visual search is guided by target representations held in visual working memory (attentional templates). While several studies have explored how visual-perceptual representations of target defining features guide attention during visual search, much less is known about visual search guided by verbal representations. In this study, we used the contralateral delay activity (CDA) and the N2pc component of the event-related potential to measure template activation and target selection, respectively, during search guided by visual-perceptual versus verbal templates. Each trial started with the presentation of a cue display indicating one or two task-relevant colour(s) for the upcoming search. In different blocks, cues were either coloured squares (visual-perceptual cues) or the initial letters of the colour words (R for red; verbal cues). After a retention period of 1000ms, search displays with six differently coloured bars appeared and participants reported the orientation of the bar in (one of) the cued target colour(s). Target N2pcs were virtually identical in the visual-perceptual and verbal cueing tasks, both in one- and two-colour search. However, CDA components, measured during the retention period, were substantially increased in response to verbal as compared to visual-perceptual cues. These results suggests that the verbal cues were translated into visual-perceptual representations during the retention period (CDA), so that when the search display arrived, visual-perceptual colour representations were available to guide target selection equally efficiently in both search task (N2pc). We tested this hypothesis in a follow-up experiment in which we shortened the retention period to 300ms to reduce the time for the translation from verbal code to visual-perceptual representations. N2pc components were now substantially attenuated and delayed in the verbal versus visual-perceptual cueing task, demonstrating impaired target selection during visual search when verbal target representations cannot (sufficiently) be translated to a visual-perceptual representation before the arrival of the visual search display.

This work was funded by a research grant of the Leverhulme Trust (RPG-2020-319)

23.431 THE SENSITIVITY OF THE DOMINANT RHYTHM DESYNCHRONIZATION TO VISUAL STIMULI ACROSS INFANCY: A STUDY USING LONGITUDINAL AND CROSS-SECTIONAL DATA

Mina Elhamiasl[1] (melhamiasl@ufl.edu), Maeve R Boylan[1], Ryan Barry-Anwar[1], Andreas Keil[1], Lisa S Scott[1]; [1]University of Florida

The peak frequency of the infant dominant EEG rhythm (6-9 Hz) increases from 6 to 9 to 12 months across occipital, frontal, and central regions regardless of stimulus presentation (Elhamiasl et al., 2023). Desynchronization, referring to the attenuation of the dominant rhythm activity in response to visual processing is an additional neurodevelopmental marker of the dominant rhythm. However, the developmental trajectory of desynchronization is not well understood. To address this gap, longitudinal (n=19) and cross-sectional EEG data sets of 6-month-old (n=49), 9-month-old (n=32), and 12-month-old (n=30) infants were collected while infants were completing a sustained attention task in which novel computer-generated objects and female faces were floating down the screen. For the longitudinal analysis, there was significant desynchronization across ages and stimuli. However, in the larger cross-sectional data desynchronization of the dominant rhythm was present in infants as young as 6 months and desynchronization increased marginally from 6 to 9 months of age. For both the longitudinal and cross-sectional data, desynchronization was significantly greater in response to faces compared to objects, suggesting greater allocation of attention to faces throughout the first year of life. The findings, overall, demonstrated that unlike peak frequency, which was not found to be sensitive to condition differences, desynchronization of the dominant rhythm differentiates infant neural responses to faces and objects.

Funding for this research was provided to L. Scott and A. Keil from a National Science Foundation (BCS:1728133).

23.432 FROM PREPARATORY ATTENTION TO STIMULUS SELECTION: NEURAL MECHANISMS REVEALED BY MULTIVARIATE ANALYSIS OF FMRI DATA

Qiang Yang[1] (yang.qiang@ufl.edu), Sreenivasan Meyyappan[2], George R Mangun[2], Mingzhou Ding[1]; [1]Department of Biomedical Engineering, University of Florida, [2]Department of Psychology and Center for Mind and Brain, University of California Davis

Preparatory attention is often studied using the cueing paradigm. According to the prevailing theory, following an attention-directing cue, top-down signals from the frontoparietal attention control regions propagate to visual cortex to bias sensory neurons to enable stimulus selection. Despite years of research, the underlying neural mechanisms remain to be better elucidated. We recorded fMRI data from participants performing a cued visual spatial attention task. At the beginning of each trial, participants were asked to covertly deploy attention to one of the two visual fields. Following a random cue-target period, a stimulus appeared either at the attended location or at the unattended location, and participants discriminated the stimulus appearing at the attended location and ignored the stimulus appearing at the unattended location. Applying MVPA to fMRI data, we reported the following findings: (1) attend-left vs attend-right can be decoded from the cue-evoked neural activity in all visual areas, (2) stimulus-left vs stimulus-right can be decoded from the target-evoked activity in all visual areas, (3) classifiers built on the cue-evoked data can decode stimulus-evoked activity in all visual areas and vice versa, and (4) the higher the cross-decoding accuracy, the better the behavioral performance. These results suggest that top-down control signals form neural patterns in the cue-target period that resemble the neural patterns evoked by the stimulus and these "attentional templates" enable stimulus selection and improve behavior.

23.433 ALPHA TRAVELING WAVES INDEX SPATIAL ATTENTION

Camille Fakche[1] (c.fakche@bham.ac.uk), Laurie Galas[1], Kirtsen Petras[1], Laura Dugué[1,2]; [1]Université Paris Cité, CNRS, Integrative Neuroscience and Cognition Center, F-75006 Paris, France, [2]Institut Universitaire de France (IUF), F-75005 Paris, France

Voluntary, spatial attention has been associated with alpha brain oscillations (8-12Hz) resulting in periodic behavioral performance. It was recently proposed that considering the spatial dimension of brain oscillations could explain more of the variance in performance (Fakche and Dugué, 2023). Here, we tested the hypothesis that attentional allocation is related to alpha traveling waves (Alamia et al., 2023), i.e., the propagation of alpha oscillations with a monotonic phase shift across cortical locations. Healthy human participants (n=20) performed a detection task while their brain activity was recorded with electroencephalography (EEG). Attention was first manipulated using a cue instructing participants to direct their spatial attention to either the bottom right or left quadrant. After a 1500ms-delay, a checkerboard stimulus flickering at 10Hz (visual inducer) was presented for 4500ms. During this presentation, a low contrast target was flashed at the attended (valid) or unattended (invalid) location, and participants were instructed to press a key when detecting it. Behavioral responses showed higher detection performance for valid than invalid trials, and EEG responses in the cue-to-stimulus period showed alpha lateralization (higher alpha amplitude in the ipsilateral electrodes relative to the cued location). These control analyses confirmed that voluntary attention was successfully manipulated. Alpha traveling waves were then assessed using optical flow analysis, i.e., a technique for tracking the displacement of similarly valued pixels over a sequence of sensor array snapshots, to reveal spatio-temporal patterns of phase variations in the data. Our results showed that although the 10Hz-visual inducer produced posterior-to-anterior propagation of 10Hz oscillations, attentional orienting (cue-to-stimulus period) showed alternating anterior-to-posterior and posterior-to-anterior alpha traveling waves. We speculate that the alternation of top-down and bottom-up traveling waves reflects functional communication between sensory and higher-level brain areas.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 852139 – Laura Dugué).

23.434 SUSTAINING ATTENTION UNDER MONITORING CONDITIONS: WHAT CHANGES IN THE BRAIN WHEN ATTENTION LAPSES?

Anina Rich[1] (anina.rich@mq.edu.au), Benjamin Lowe[1], Hamid Karimi-Rouzbahani[2], Katie Hensley[1], Denise Moerel[3], Alexandra Woolgar[4]; [1]Macquarie University, Sydney, Australia, [2]Mater Hospital, South Brisbane, Australia, [3]University of Sydney, Sydney, Australia, [4]University of Cambridge, Cambridge, UK

The move to semi-automated systems has created many situations where humans are required to sustain attention over time with only infrequent responses (e.g., monitoring in train and aircraft network control, semi-automated vehicles). Unfortunately, when targets are rare, observers are more likely to miss them or be slow to react (e.g., Wolfe et al. 2005). Our Multiple Object Monitoring (MOM) task includes many key features of modern environments: it requires selection of targets from distractors, tracking of dynamic stimuli, and allows distinction of time-on-task effects from target frequency effects. In the MOM task, dots approach a central obstacle from different directions along predictable trajectories. The task is to avoid a collision between dots of the relevant colour and the central obstacle. On a variable proportion of trials, the dots are automatically deflected away from collision at a set point in the trajectory. In our original study, we found that in 'active' conditions, where participants manually deflect 50% of dots; they can successfully sustain performance across time. In monitoring conditions, where participants manually deflected 6% of dots, performance dropped sharply over time, showing a clear vigilance decrement. Here, we report task parameter manipulations that show vigilance decrements decrease with increasing target frequency, as well as using electroencephalography (EEG) combined with multivariate pattern analyses (MVPA) to examine patterns of neural activity during successful and lapsed attention. Our neural studies show that we can use the pattern of activity across the brain to identify information about task-relevant aspects of the task (e.g., target dot location) and use the neural difference between correct and miss trials to predict behavioural results. These findings contribute to our understanding of successful versus unsuccessful sustained attention and provide a foundation for predicting behavioural errors before they occur, based on neural patterns of activation.

DP220101067 to ANR and AW

23.435 THE NEURAL BASES OF FLUCTUATIONS IN SHIFT READINESS AND SUSTAINED ATTENTION

Anna B. Toledo[1,2] (annabctoledo@gmail.com), Yuxin Xie[1], Emily E. Oor[1], Anthony W. Sali[1]; [1]Wake Forest University, [2]Georgetown University

The capacity to anticipate environmental demands and adjust spatial attention accordingly is a crucial aspect of attentional control. As an individual's goals change, they must regulate their readiness to shift attention, such that the cost associated with shifting decreases as the likelihood of shifting increases (e.g., Sali et al., 2015). Moreover, the efficacy of sustained attention is subject to variability over time, with periods of high and low response time (RT) variability indicating unfocused and focused states, respectively (e.g., Esterman et al., 2013). In the current study, we measured the behavioral consequences and neural correlates of both forms of control fluctuations simultaneously. While inside an MRI scanner, participants completed a variant of the gradual continuous performance task (gradCPT) in which they detected frequent targets among distractors in one of two simultaneous visual streams, shifted or held attention in response to embedded visual cues, and made a digit categorization as a measure of shift readiness. We manipulated the shift likelihood across blocks of trials. As in previous studies, blocks with a high shift likelihood were associated with a reduction in response time shift costs relative to blocks with a low shift likelihood, reflecting learned shift readiness. We defined periods of high and low sustained attention efficacy, referred to as "in the zone" and "out of the zone," respectively, based on ongoing changes in continuous performance task RT variability and found that commission errors were significantly more frequent when participants were "out of the zone" than when "in the zone." Functional MRI analysis suggested that separate but interacting neural systems account for moment-by-moment changes in attentional control, with the frontoparietal cortex detecting and responding to violations of shift readiness and the default mode network regulating ongoing changes in sustained attention.

Funding provided through R15MH127491 awarded to A.S.

23.436 BOTH TARGET AND DISTRACTOR ARE SAMPLED RHYTHMICALLY IN A MOTION DETECTION TASK

Changhao Xiong[1] (xiong.changhao@ufl.edu), Ke Bo[2], Nathan Petro[3], Andreas Keil[1], Mingzhou Ding[1]; [1]UNIVERSITY OF FLORIDA, [2]DARTMOUTH COLLEGE, [3]BOYS TOWN NATIONAL RESEARCH HOSPITAL

It has been shown that the visual system samples the attended information rhythmically. Does rhythmic sampling also apply to distracting information? How do attended information and distracting information compete temporally for neural representations? We recorded electroencephalography from participants who detected instances of coherent motion in a random dot kinematogram (RDK) (primary task) overlayed on different categories (pleasant, unpleasant, and neutral) of affective images from the IAPS library (distractor). The RDK was flicked at 4.29 Hz whereas the IAPS pictures at 6 Hz. From the SSVEP time series, the time course of the power at 4.29 Hz was extracted in a moving window approach, and its fluctuation was taken to index the temporal dynamics of attended information processing. The time course of the power at 6Hz was similarly extracted and support vector machine (SVM) was applied to decode different categories of affective images with the resulting fluctuating decoding accuracy taken to index the temporal dynamics of distracting information processing. We found that (1) both the 4.29 Hz power time course and the 6 Hz decoding accuracy time course exhibited rhythmicity at 1 Hz and (2) the phase difference between the two rhythmic time courses predicted task performance, i.e., phase difference close to pi corresponded to a higher rate of coherent motion detection whereas phase difference close to 0 corresponded to a lower rate of coherence motion detection. These results suggest that (1) both attended and distracting information were sampled rhythmically and (2) alternating the sampling between target and distractor reduces the adverse impact of distractor.

23.437 THETA- AND ALPHA-BAND FREQUENCY ADVANTAGES FOR SENSORY GATING

Oakyoon Cha[1] (oakyoon.cha@sungshin.ac.kr); [1]Department of Psychology, Sungshin Women's University

Several studies propose that neural oscillations in the theta and alpha bands serve as mechanisms for sensory gating. According to the communication-through-coherence (CtC) theory (Fries, 2009), effective neural communication relies on aligning the phases of ongoing theta/alpha oscillations across neurons in different brain regions. Consequently, theta- and alpha-band neural oscillations modulate gamma-band neural activities to facilitate or suppress neural communication. The rhythmic theory of attention (Fiebelkorn & Kastner, 2019) posits that theta-band oscillations act as a clock signal for attentional sampling and shifting mechanisms. In this context, the visual system samples information during one phase of the theta band oscillation and shifts attention to a new location in the subsequent phase. To date, no study has investigated whether theta and alpha bands offer advantages over other frequencies in the context of gating mechanisms. The present study explored optimal frequency bands of neural oscillations for sensory gating by examining how oscillations can function as gating mechanisms in a modified one-choice drift-diffusion model (DDM). The modified DDM incorporates CtC theory by using a sine wave to modulate the drift rate, simulating oscillations in the population of signal-sending neurons. Simultaneously, the decision criterion is modulated with a cosine wave, simulating oscillations in the population of signal-receiving neurons. I found that lower-frequency oscillations in the drift rate left stronger traces in the response time distribution of the DDM, while lower-frequency oscillations in the decision criterion exhibited a less pronounced effect. Importantly, phase differences resulted in the most substantial modulation on the strength of oscillatory traces when the oscillation frequency embedded in the drift rate and decision criterion was within theta and alpha bands. In conclusion, the present study underscores that theta- and alpha-band oscillations exhibit optimal characteristics for gating mechanisms within the context of CtC theory.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2023-00211668).

Saturday, May 18, 8:30 am – 12:30 pm, Pavilion

Attention: Inattention, attentional blindness, suppression

23.438 SENSITIVITY TO HIGHLY SALIENT FEATURES IN DYNAMIC INATTENTIONAL BLINDNESS

Makaela Nartker[1] (makaela@jhu.edu), Chaz Firestone[1], Howard Egeth[1], Ian Phillips[1]; [1]Johns Hopkins University

In inattentional blindness (IB), subjects who fail to report unexpected stimuli are typically assumed not to have seen them. Recent work challenges this assumption by showing that inattentionally blind subjects can respond above-chance to stimuli they report not noticing (Nartker et al., 2022), suggesting that inattention may not completely abolish awareness. However, these results have been limited to briefly and peripherally presented static stimuli (e.g., a line appearing on the edge of a display for 200 ms). Does this pattern extend to long-lasting IB involving highly salient dynamic stimuli? Here we report data from a large-scale online study (N>10,000) addressing precisely this question. Subjects were shown a gray rectangular display containing moving white and black squares, and counted how often the white squares bounced off its perimeter (adapted from Wood & Simons, 2017). For some subjects, the third trial included an additional brightly colored and highly salient shape (a circle or triangle that was orange or green), which traversed the height of the display for five full seconds. After this critical trial, subjects were asked the standard IB question: “Did you notice anything unusual on the last trial that wasn’t there on previous trials?” (yes/no) followed by additional questions probing the extra object’s color, location and shape. By including absent trials in which no additional stimulus appeared, we found that subjects were biased to report not noticing (c=0.45, 95% CI=[0.41,0.49]), suggesting greater awareness than revealed by yes/no questioning. Consistent with this interpretation and our previous studies, inattentionally blind subjects could report the color of the unexpected object at above-chance levels (d'=0.12, 95% CI=[0.02,0.23]). Strikingly, these ‘non-noticing’ subjects were also above-chance in discriminating the objects’ shape (d'=0.23, 95% CI=[0.13,0.33]), raising the possibility that even mid- or high-level features survive inattention.

National Science Foundation BCS #2021053; JHU Catalyst Award

23.439 ASYMMETRY FOR SHADING DIRECTION IN VISUAL SEARCH PERSISTS IN INATTENTIONAL BLINDNESS

*ShuiEr Han[1,2] (han.shuier@gmail.com), Yang Shen[1,2], Mengmi Zhang[1,2]; [1]I2R and CFAR, Agency for Science, Technology and Research (A*STAR), [2]Nanyang Technological University (NTU), Singapore*

When human observers search among shaded visual stimuli, they find the targets in vertical shading much faster than the ones in horizontal shading. Here, we demonstrate that this asymmetry persists in an inattentional blindness (IAB) paradigm. In our study, subjects viewed naturalistic simulations of moving balls that were vertically or horizontally shaded. A portion of the trials contained an unexpected target, which had a reversed shading gradient, and was introduced to the simulation at random times. During each trial, subjects tracked a ball and counted the number of midline crossings made. They were also instructed to indicate when they noticed an unexpected target. Results showed that almost twice as many vertically shaded targets were detected compared to horizontally shaded targets, and this difference could not be attributed to differences in target visibility, false target detection rate, and average ball counting accuracy. To gain insight into the underlying mechanisms of these results, we propose a biologically inspired, computational IAB model based on predictive coding. The model undergoes unsupervised training to anticipate subsequent video frames by minimizing expected errors inherited from preceding predictions made during the structure analysis of naturalistic video sequences. Subsequently, the model is tested on the same videos used in human psychophysics experiments. Remarkably, this model exhibits a more pronounced variance in predictive errors, when the unexpected target is in horizontal shading. Together, our findings point to the emergence of IAB asymmetry through top-down expectation biases derived from the visual stimuli presented to both humans and the model.

23.440 ELUCIDATING FLUCTUATIONS OF VISUAL ATTENTION: REACTION TIME VARIABILITY AND MIND-WANDERING PROVIDE COMPLEMENTARY INSIGHTS

Matthieu Chidharom[1,2,3] (matthieuchidharom@gmail.com), Edward Vogel[1,2], Monica Rosenberg[1,3]; [1]University of Chicago, [2]Institute for Mind and Biology, [3]Neuroscience Institute

Sustained attention refers to the ability to maintain focus on a task over an extended period of time. However, sustaining attention is challenging, as there are inherent fluctuations between periods of good attention (stable and less error-prone) and poor attention (unstable and error-prone). Two main methodologies have been used to isolate these attentional fluctuations in visual attention tasks: objective and subjective approaches. The objective methodology analyzes intraindividual reaction time (RT) variability. For example, Esterman et al. (2013) revealed higher errors during periods of high RT variability (out-of-the-zone state) compared to periods of low variability (in-the-zone state). The subjective methodology uses thought probes to identify mind-wandering episodes (task-unrelated thoughts), that can be either intentional (i.e., deliberate) or unintentional (Seli, 2016). Although previous studies revealed higher error rates during out-of-the-zone and mind-wandering states, it is unclear whether those two methodologies isolate the same or different type of attentional fluctuations. This study compared the two approaches in a single Go/NoGo sustained attention task (N=38), objectively measuring attention via RT variability, and subjectively via intermittent thought probes administered every 30 trials. The entire task lasted approximately 40 minutes. If both methods isolate similar fluctuations, we hypothesize that the time spent out-of-the-zone should be higher during mind-wandering than during on-task periods. Our results revealed significantly higher time out-of-the-zone during mind-wandering (52.1%) compared to on-task periods (47.8%), suggesting the methods isolate, to some extent, similar aspects of fluctuations. However, time out-of-the-zone was significantly higher during intentional (57.7%) versus unintentional mind-wandering (49.7%), indicating higher overlap between the objective and subjective methods for deliberate mind-wandering rather that spontaneous. These findings suggest the two methods capture complementary information about attentional fluctuations and highlights the utility of combining objective and subjective methods to gain a comprehensive understanding of sustained attention and its lapses.

23.441 INATTENTIONAL BLINDNESS FOR A SALIENT TARGET IN VISUAL SEARCH: FINDING A SURPRISINGLY EASY TARGET CAN BE SURPRISINGLY HARD

Daniel Ernst[1] (daniel.ernst@uni-bielefeld.de), Gernot Horstmann[1], Johan Hulleman[2], Jeremy Wolfe[3]; [1]Bielefeld University, [2]The University of Manchester, [3]Brigham and Women’s Hospital/Harvard Medical School

In natural search settings, observers often do not know how difficult it will be to find the next target. The next security threat or the next potential cancer in a lung CT might be quite salient or very subtle. In the lab, visual search experiments typically involve multiple successive trials of constant difficulty. This allows participants to anticipate the salience of the next target. How will observers respond if the current target salience suddenly deviates from the target salience of all preceding trials? It would be unremarkable to find that it is harder to detect a low salience target when the observer expects a high salience one. More interestingly, here we report that observers are impaired when they are surprised with a high salience target after a series of difficult searches. In Experiment 1, observers searched for a hard-to-detect O target (always present, compound search) among C distractors with small gaps for 32 trials. On the 33rd trial, the gaps of the C distractors were large, making the O target much more salient. Yet, search efficiency was considerably more inefficient on this surprise trial (147 msec/item) than on subsequent trials with the same high target-distractor dissimilarity (66 msec/item). In Experiment 2, observers reported the presence or absence of an O in a display with short presentation duration. After multiple hard trials, observers frequently showed inattentional blindness towards an unexpectedly salient target that was reported almost perfectly when the identical target was presented on repeated trials. Gaze data suggests that observers adopted a tightly focused attentional window during the initial, hard search. This strategy made them surprisingly 'blind' to targets that were unexpectedly highly detectable. The wrong attentional set may be one explanation for situations where we "look but fail to see" obvious stimuli.

Daniel Ernst was supported by grant ER 962/2-1; Johan Hulleman was supported by UKRI grant ES/X000443/1; Jeremy Wolfe was supported by grants NEI EY017001, NSF 2146617, NCI CA207490

23.443 SPATIAL ATTENTION IS CAPTURED AND SUPPRESSED BY EMOTIONAL PICTURES IN HIGH TRAIT ANXIETY

Minwoo JB Kim[1] (cogminu@gmail.com); [1]Ulsan National Institute of Science and Technology

Physically salient objects capture spatial attention in a stimulus-driven way but this can be suppressed (Wang & Theeuwes, 2018). In a recent study, color singletons associated with emotional value captured attention but was resistant to suppression at locations where distractors occured frequently (Kim & Anderson, 2021). Meanwhile, emotional capture and suppression of spatial attention may depend on the individual emotional traits such as anxiety. The current study rigorously tested this using the mouse device in subclinical high and low trait anxiety individuals. 96 participants recruited on-line were instructed to detect a circle among rectangles containing horizontal or vertical bars, and report the bar orientation in it. They reported by making mouse movements to the corresponding response box at the circle's location presented peripherally and clicking it with the right or left mouse button. When distractors were present, one of the rectangles was a picture containing emotional or neutral context. Distractors appeared more frequently on one location (suppressed) than the other three. Mouse trajectory deviations from the straight line from the display center to the response box was used as the dependent measures. STAI-Y2 was collected for measuring state anxiety and groups were divided into high and low (moderate, no-low). Both negative and neutral distractors were attracted towards distractors at low frequency locations. Here, deviation was greater for negative distractors for high anxiety only. When distractors were at high frequency locations, trajectories were generally pushed away from the distractor. Emotional differences in trajectory were not observed at high frequency locations in any group. First, this suggests that emotional capture of spatial attention is modulated by individual anxiety traits. Also results suggest that even in high anxiety groups emotional distractors can be suppressed and this can be observed when using a sensitive measure of spatial attention such as mouse-tracking.

23.444 DISTRACTOR LOCATION FREQUENCIES BETTER ACCOUNT FOR THE INSTANTIATION OF LEARNED DISTRACTOR SUPPRESSION THAN DO REINFORCEMENT LEARNING PREDICTION ERRORS

Anthony W. Sali[1] (saliaw@wfu.edu), Catherine W. Seitz[1]; [1]Wake Forest University

Stimulus-driven attentional capture is reduced when a salient distractor regularly appears at a predictable spatial location (e.g., Wang & Theeuwes, 2018). This phenomenon is consistent with a growing body of work that suggests that selection history plays a powerful role in shaping the instantiation of attentional priority. However, the underlying mechanisms of distractor suppression learning remain poorly understood. In the current study, we fitted behavioral response time (RT) data from a variant of the additional singleton paradigm with a series of computational models to test how individuals harness previous experiences to guide attentional deployment. As in previous studies, we observed robust evidence of learned distractor suppression such that RTs were shorter when the distractor appeared at a high probability location than when it appeared at a low probability location. Furthermore, RTs were also longer when the target appeared at the high probability location relative to the low probability location. Next, we adjudicated whether distractor suppression was best explained by (a) the tracking of distractor location frequencies or (b) a reinforcement-learning (RL) prediction error mechanism. Under the location frequency account, individuals decrease the priority afforded to a particular location with each successive presentation of a distractor at that location. However, while participants did not receive explicit rewards, accurate performance is intrinsically rewarding. The RL account assumes that individuals attempt to maximize performance by increasing or decreasing the priority for a particular location depending on the size and direction of the trial-by-trial difference between expected and observed distractor location likelihoods. We used Hierarchical Bayesian Inference to simultaneously fit and compare models, finding that the distractor location frequency model best accounted for the data. Together, these results suggest that a simple frequency model outperforms models that nudge predictions up and down based on trial-by-trial outcomes.

23.445 DISTRACTOR SUPPRESSION OPERATES IN RETINOTOPIC COORDINATES

Yayla Ilksoy[1] (y.a.ilksoy@vu.nl), Dirk van Moorselaar[1], Benchi Wang[2], Sander Los[1], Jan Theeuwes[1]; [1]Vrije Universiteit Amsterdam, [2]South China Normal University

Our attention is influenced by past experiences, and recent studies have shown that people learn to extract statistical regularities in the environment resulting in attentional suppression of locations that are likely to contain a distractor. Here, we asked whether this suppression effect due to statistical learning operates in retinotopic (relative to the eyes) or spatiotopic (relative to the world) coordinates. In the current study, two circular displays were presented side by side. While fixating the center of one display, observers performed the additional singleton search task in which one location was more likely to contain a distractor. After searching for several trials in one display participants moved their eyes to the center of the other display and kept performing the additional singleton search task. Because of the saccadic eye movement to the other display, the suppressed location that in the previous display was in retinotopic coordinates became the spatiotopic location in the new display while the retinotopic location moved along with the eye movement in space. The results showed that, following the eye movement, the suppression remained in retinotopic coordinates, while there was no transfer of suppression to the spatiotopic coordinates. It is important to note that in the experiment conducted there were no environmental landmarks as the search displays were presented on a blank empty background. Also, the entire display shifted from side to side, making the whole visual field move along with the saccades. To address these limitations, a second experiment was conducted with a grid and placeholders in the display to enhance more environmental stability. The results however showed that even in a rich environment, attentional suppression still operated in retinotopic coordinates only. Given that suppression is only found in retinotopic coordinates, it is speculated that suppression maybe resolved by changing synaptic weights in early visual areas.

This project was supported by an open competition grant of NWO

23.446 ELECTROPHYSIOLOGICAL EVIDENCE FOR LEARNED FEATURE SUPPRESSION

Aylin A. Hanne[1] (aylin.hanne@uni-marburg.de), Sizhu Han[1], Anna Schubö[1]; [1]Philipps University Marburg

Humans are able to learn statistical regularities from their visual environment to reduce interference from non-relevant, salient distractors. Previous studies have identified statistical regularities as spatial or feature-based, but the neural mechanisms of feature-based distractor suppression are still unclear. Also, the effect of distractor feature learning on visual working memory (VWM) performance has not been explored yet. To examine the impact of distractor feature learning on attentional selection and VWM performance, we implemented a variant of the additional-singleton task in which the distractor appeared more likely in one specific color (high-probability color) than other colors (low-probability colors). During this learning task, we simultaneously recorded the EEG. Before and after the learning task, participants performed a change detection task in which the high- and low-probability distractor colors were used. The behavioral results of the learning task showed a decrease in response times when the distractor appeared in a high-probability color compared to the low-probability colors, indicating learning of the distractor feature regularities. In line with the behavioral pattern, the neural measures revealed a larger target N2pc and a smaller distractor PD for the high-probability color, suggesting more efficient attentional selection when the distractor appeared in the more likely color. Interestingly, our data revealed a modification of the distractor PD over time: in high-probability trials, we found a decrease in the late PD while the early PD increased from the first to the second half of the experiment, indicating a temporal shift from reactive to more proactive distractor suppression. VWM performance, in contrast, was not affected by distractor feature learning. In summary, our results suggest that learned distractor feature regularities are proactively used to reduce distractor interference before the first shift of attention allocation without noticeably affecting VWM performance.

This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation – project number 290878970-GRK 2271, project 9, and SFB/TRR 135, project number 222641018, TP B3).

23.447 THE COORDINATE SYSTEM OF SUPPRESSION IN DYNAMIC CONTEXTS

Seah Chang[1] (chang.2127@osu.edu), Julie D. Golomb[1]; [1]The Ohio State University

It has been shown that observers can learn to suppress a spatial location that frequently contains a salient singleton distractor, which is referred to as learned spatial suppression. Our previous work showed that the default coordinate system of learned spatial suppression is retinotopic: Spatial suppression is naturally learned and transferred across gaze position in retinotopic (eye-centered) coordinates (Chang & Golomb, 2023). However, in the real world, spatiotopic (world-centered) coordinates are generally more beneficial to guide behavior because retinotopic representations are often unstable due to frequent eye movements. Can spatial suppression be learned in spatiotopic coordinates when the context is more conducive to spatiotopic learning? In the current study, participants performed a visual search task while the fixation point randomly changed between two locations (gaze positions 1 and 2) across trials. Participants searched four items for a shape oddball target (e.g., a diamond among circles) while ignoring a salient color singleton distractor that appeared on two-thirds of trials. Critically, the salient distractor appeared more frequently at a spatiotopic high-probability location, regardless of gaze position, creating a spatiotopically-weighted context. In one experiment, the target appeared equally often at each location, providing only distractor-related probability information (distractor regularities only; pure distractor suppression). In another experiment, in addition to the distractor probability manipulation, the target never appeared in the spatiotopic high-probability location, providing both target- and distractor-related probability information for suppression (both target and distractor regularities). Strikingly, the results showed that spatial suppression could be learned in spatiotopic coordinates when driven solely by distractor-related probability information (pure distractor suppression) but was learned in retinotopic coordinates when driven by both target- and distractor-related probability information. These results raise an intriguing possibility that pure distractor suppression is susceptible to spatiotopic-based learning, but target-related processing may be more hardwired for retinotopic representations.

NIH R01-EY025648 (JG), NSF 1848939 (JG)

23.448 LEARNING TO SUPPRESS COLOR SINGLETONS VIA FEATURE-BASED REGULARITIES

Brad T. Stilwell[1] (brad.t.stilwell@gmail.com), Brian A. Anderson[1]; [1]Texas A&M University

Do salient items, such as uniquely colored stimuli, automatically capture our attention? This question has been heatedly debated for decades, traditionally spurred on by two competing accounts: stimulus-driven and goal-driven. According to stimulus-driven accounts, salient items, such as color singleton distractors, have the power to automatically capture attention. In contrast, goal-driven accounts posit that only items matching the observers top-down attentional set will capture attention. This failed dichotomy misses another important source of attentional control: an observer's previous experience or selection history. One such example of selection history is the finding that color singletons appearing in a frequent color result in less capture than singletons presented in less frequent colors. But what is the nature of this selection history? The current study varied exposure to feature-based statistical regularities to assess how they influence the control of attention. High- and low-frequency singleton distractors were presented either within blocks, to test "short-lived" regularities, such as intertrial priming, or between blocks, to test "longer-lived" regularities, such as attentional control settings. We found that capture by the high probable color was suppressed relative to capture by the low probable colors, and this learned suppression was evident for both the within- and between-block manipulations. These findings suggest that learning to suppress a distractor feature based on statistical regularities is robust and longer-lived.

23.449 EVIDENCE AGAINST THE LOW-SALIENCE CRITICISM OF SIGNAL SUPPRESSION

Nicholas Gaspelin[1], Brad T. Stilwell[2], Howard E. Egeth[3]; [1]University of Missouri, [2]Texas A&M University, [3]Johns Hopkins University

There has been a longstanding debate about whether salient stimuli, such as uniquely colored objects, have the automatic power to capture attention. The signal suppression hypothesis proposed a resolution whereby salient stimuli have inherent power to capture attention, but that capture can be prevented by inhibiting such stimuli. This hypothesis has recently been criticized on the grounds that initial studies supporting it may have used singletons that were only weakly salient (the low-salience criticism). According to this argument, salient stimuli have the automatic power to capture attention; but only when made sufficiently salient. This criticism may not be well-founded, however, as studies on the topic have not used an objective measure of salience (or, indeed, any measure at all) to evaluate their stimuli. We have recently developed a new psychophysical technique to measure salience. The current study used this technique to compare salience of color singletons from an original study of signal suppression (Gaspelin et al., 2015) and a study purporting to increase salience (Wang & Theeuwes, 2020). We found that, if anything, color singletons were more salient in the original studies supporting signal suppression than in the studies purporting to improve salience. Ultimately, these findings suggest that the low-salience criticism of the signal suppression hypothesis is unfounded.

This project was supported by the National Science Foundation Grant BCS-2345898 to N.G.

23.450 DECREASED PERCEPTUAL SENSITIVITY DURING ATTENTION SHIFTS

Zixiao Zhang[1,2] (zhangzixiao16@mails.ucas.ac.cn), Sheng He[1,2], Jiedong Zhang[1,2]; [1]Institute of Biophysics, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China

Both saccades and attention shifts are cognitive mechanisms for relocating the focus of information processing to different spatial locations. While it is known that visual sensitivity decreases during saccades (i.e., saccadic suppression) to support stable visual experience, here we asked whether attention shifts would also cause reduced perceptual sensitivity. In Experiment 1, we measured phases of the rhythmic spontaneous shift of attention between two task-relevant locations. To estimate the perceptual sensitivity at different attention states, a task-irrelevant visual probe was presented in the background at varying phases of attention oscillation, and the amplitudes of pupil constrictions induced by the probe were measured. Results showed a reduction in pupil constriction during attention shifts, indicating a decrease in perceptual sensitivity. Additional tests with a single-location attention task revealed an absence of modulation of pupil constriction, confirming that the reduced sensitivity was due to attention shifts and not to rhythmic attention sampling. In Experiment 2, instead of relying on pupil constriction, MEG responses to the probe stimulus were used to estimate the perceptual sensitivity during attention shifts. The experimental design was similar to that used in Experiment 1, and the results showed a decrease in MEG responses in the occipital region during attention shifts. Further analysis identified that the modulation of MEG response peaked around 150-200 ms after the probe onset, indicating that the modulation occurs at a late stage compared to that in the saccadic suppression. Additionally, a whole-brain analysis revealed that the decreased response was primarily located in the parietal region. Taken together, our findings suggest that attention shifts over space are accompanied by decreases in visual sensitivity, likely due to suppressive top-down modulation from the parietal attention network to the early visual cortex.

SATURDAY, MAY 18, 8:30 AM – 12:30 PM, PAVILION

Decision Making: Decision making and actions

23.451 ACTIVE OR PASSIVE INFERENCE? EFFECTS OF GOAL-DIRECTED ACTIONS ON PERCEPTUAL DECISIONS

Jet Lageman[1] (j.lageman@vu.nl), Johannes J. Fahrenfort[1], Heleen A. Slagter[1]; [1]Vrije Universiteit Amsterdam

Perceptual decisions can be regarded as the result of a Bayesian inference process, combining prior beliefs with sensory observations to form posterior beliefs about hidden states of the world. Depending on whether sensory observations are passively sampled or actively generated, prior beliefs guiding the inference process can be informed by probabilistic sensory cues, or by knowledge about action-outcome relationships. However, it remains unclear exactly how goal-directed actions impact belief updating and subsequent perceptual decision-making. Here, we compared the outcomes of inference for perceptual judgments or goal-directed actions during a probabilistic reversal learning task, in which we manipulated the uncertainty of sensory observations (Experiment 1) or the volatility of the environment (Experiment 2). Participants were either asked to infer a hidden state from computer-sampled observations, or to generate specific observations determined by a hidden state, while keeping track of sudden reversals in the hidden state. Critically, participants received the same amount of evidence for the current hidden state under each instruction. Results indicate that active inference may slow reversal learning by reducing responsiveness to conflicting evidence. In addition, using Bayesian computational modelling, we investigated trial-by-trial belief trajectories and response models, aiming to disentangle perception and learning from decision noise or response bias and to study how goal-directed actions may impact the way we perceive and form beliefs about the world in noisy and volatile environments.

23.452 GENERATING SACCADES FOR REDUCING UNCERTAINTY: COGNITIVE AND SENSORIMOTOR TRADE-OFFS

Matthew Jiwa[1] (mj3159@columbia.edu), Constantin Rothkopf[2], Jacqueline Gottlieb[1]; [1]Columbia University, [2]Technical University of Darmstadt

In uncertain environments, eye movements inspect relevant stimuli that reduce uncertainty, but the mechanisms of this process are not well understood. How do we trade off the benefits of reducing uncertainty against the costs of planning and executing saccades? To examine this question, we designed a task in which participants attempted to estimate the relative probabilities of two possible world-states, and used saccades to inspect a set of visual stimuli that had different predictive accuracy (diagnosticity) regarding those states. Diagnosticity was explained to participants in advance and indicated by stimulus color. The time available for inspection was limited, so that participants could only inspect a subset of the available stimuli. Preliminary results together with computational modelling suggest that participants traded off stimulus diagnosticity against eccentricity (saccade size) and planned saccade sequences rather than individual saccades. Future task versions and computational models will examine how these trade-offs depend on uncertainty, shedding light on the sensorimotor mechanisms that the brain uses to sample visual stimuli to reduce uncertainty.

23.453 GLOBAL ROUTE SELECTION USING LOCAL VISUAL INFORMATION

Cassandra Engstrom[1] (cassandra_engstrom@brown.edu), William H Warren[1]; [1]Brown University

When the structure of the environment is unknown, humans must navigate using local visual information. One strategy involves minimizing the angular deviation of one's heading from a distal goal (θ). Others include minimizing the local distance (d) or turning angle (γ) to available routes. We investigated whether these variables interact to influence navigational decisions, as previously observed by Baxter & Warren (2020) for routes around a barrier. Participants walked to a goal pole through a virtual environment (32'x32') viewed in a Quest Pro VR headset. The environment contained 3 parallel walls ("layers"), each with two doorways, yielding three binary choices per trial. Door placement was randomized to produce 64 novel configurations (half mirror-reversed), each visited once. In Experiment 1 (N=17), the goal was always visible above the walls. Experiment 2 (N=17) was identical, except that the goal disappeared before walking began. Logistic regression analyses revealed that subjects used all three local variables, minimizing deviation angle (θ), distance (d) and turn angle (γ) when selecting a doorway in each layer (all $p <.01$). The influence of d and θ increased with goal proximity, with θ dominating in the middle layer and d at the end. Although the goal's disappearance weakened the θ strategy ($p < .05$), presumably due to spatial updating error, the other variables were constant across experiments (ns). To estimate the global consequences of local strategies, we measured the energetic cost of humans walking all possible routes and compared the performance of simulated agents following different strategies. We found that the agent that minimized θ alone selected energetically optimal routes roughly as often as the regression model, while both performed better than the d and γ agents. Our results suggest that humans navigate using a flexible local strategy that incorporates multiple variables and yields efficient global routes.

NIH R01EY029745

23.454 SACCADIC DECISION MAKING BASED ON UNCERTAIN AUDITORY CUES IN MONKEY SUPERIOR COLLICULUS

Christopher Conroy[1] (cconroy@sunyopt.edu), Robert M. McPeek[1]; [1]SUNY College of Optometry, Department of Biological and Vision Sciences

Saccadic decisions are often based on uncertain auditory cues, which must be evaluated by the auditory or saccadic system and transformed into a motor plan. Relatively little is known about the neural mechanisms that support such saccadic decisions, at least as compared to saccadic decisions based on uncertain visual cues. The superior colliculus (SC) is involved in visual-saccadic decision making and receives auditory-sensory input from multiple brain regions. It is possible, therefore, that the SC is involved in auditory-saccadic decision making in addition to being involved in the evaluation of visual evidence. To examine this possibility, we recorded the activity of SC neurons in a rhesus monkey that was making decisions about where to look based on uncertain auditory cues. On each trial, the monkey fixated a visual stimulus and a sequence of brief, auditory-white-noise bursts was presented in the frontal-horizontal plane. The horizontal position of each burst was drawn from a Gaussian distribution, the mean of which determined the strength of auditory evidence. The monkey's task was to decide if the mean of the distribution was to the left or right of the frontal midline and to report that decision with a saccade to a visual target. The monkey's behavior varied systematically with evidence strength: sensitivity increased as the distribution's mean moved away from the frontal midline. A subset of SC neurons exhibited decision-related activity, meaning that there was a modulation of their activity patterns based on auditory-evidence strength. This suggests that the SC plays a role in the evaluation of uncertain auditory cues when the evaluation of such cues is required for saccade planning. It also suggests a more general role for the SC, and perhaps other oculomotor structures, in the formation of saccadic decisions based on nonvisual sensory input.

NIH NEI Grant R01-EY030669

23.455 STIMULUS-RESPONSE MAPPING SHAPES UNINTENDED RESPONSE DURATION MODULATION IN DURATION JUDGMENT TASKS

Juhyeon Yeo[1] (jhyeo@unist.ac.kr), Oh-Sang Kwon[1], Yun Kyoung Shin[2]; [1]Ulsan National Institute of Science and Technology, [2]University of Ulsan

The ideomotor theory, as originally posited by William James in 1890, suggests that the mere mental representation of a movement has the potential to trigger the actual movement. The precise nature of the representational format—often colloquially referred to as 'an idea'—that serves as a direct conduit for initiating movement remains unclear. Shin et al. (2023) addressed the issue by introducing an experimental paradigm, where participants were required to categorize six audiovisual stimuli durations as 'short' or 'long' by pressing a left or right key. Notably, response durations reflected the nature of the decision category, despite the absence of explicit instructions on how long they were to press the key, thereby demonstrating a genuine ideomotor effect. The current study further examined how categorizing stimuli as long or short influenced motor execution, contingent on the task's difficulty level. With stimuli durations set at 85, 100, 115, 135, 150, and 165 ms for the 'difficult' categorization task, and 85, 100, 115, 165, 180, and 195 ms for the 'easy' task, the results showed a robust ideomotor effect, more pronounced during the 'easy' task. Intriguingly, the preassigned mapping of the hand to the decision category modulated the ideomotor effect differently, particularly in the 'difficult' task; the ideomotor effect was negligible when the left hand was assigned to the 'short' decision category, whereas the right hand to short decision assignment resulted in a strong ideomotor response, $F(5, 1) = 24.61$, $p = <.001$, $\eta p^2 = 0.552$. Such an interaction was not found in the 'easy' task, $F(5, 1) = 0.445$, $p = 0.816$, $\eta p^2 = 0.022$. Possible mechanisms underlying the observed pattern may include the cognitive mechanism similar to the spatial-numerical association of response codes (SNARC), or an asymmetric hemisphere control within the context of the ideomotor effect.

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2022S1A5A8054398).

23.456 VISUAL FACTORS THAT DETERMINE UNCERTAINTY IN RAPID INTERCEPTIVE MOVEMENTS

Abibat A. Akande[1] (akande02@student.ubc.ca), Philipp Kreyenmeier[1], Miriam Spering[1]; [1]University of British Columbia

The ability to rapidly assess visual motion information is critical for daily tasks that require rapid interceptions such as catching a falling object. When visual information is sparse, interception decisions have to be made under uncertainty. To investigate this in the laboratory, we can constrain the availability of visual information or the time over which information has to be extrapolated. We recorded human observers' (n=10) eye and hand movements while they viewed the launch of a simulated flyball on a screen. The ball was occluded shortly after launch, and observers had to manually intercept it along its predicted trajectory within a hit zone. We measured interception accuracy in eye and hand movements for different trajectory shapes yielded by ball speed variations. In two sessions, we manipulated uncertainty by either varying ball presentation duration (100-500ms) at constant occlusion duration (500ms), or by varying occlusion duration (100-500ms) at constant presentation duration (500ms). Reducing uncertainty by increasing presentation duration rapidly improved eye and hand interception accuracy and reduced a bias toward the center of the trajectory space that was observed at the shortest presentation duration. These improvements occurred within the first 300ms of ball presentation. By contrast, long occlusion durations yielded a much weaker center-bias. Reducing uncertainty by decreasing occlusion duration from 500-100ms linearly improved interception accuracy until eye and hand interception errors were minimal. The availability of visual information and time available to extrapolate information both determine uncertainty in rapid interceptive control, but they do so at different rates. Effects of presentation duration plateaued early, indicating that 300ms are sufficient to accurately read out visual trajectory information. Occlusion duration improved performance at a later time, as interception switched from predictive to visually-guided control.

23.457 DOES SOCIAL CONTEXT INFLUENCE INTENTION, PREDICTION AND MOTOR BEHAVIOR DURING A SIMPLE IN-PERSON CARD GAME?

Bethany B. Jantz[1,2], Madison L. Fankhanel[3], Dana A. Hayward[2,3], Craig S. Chapman[1,2]; [1]University of Alberta, Faculty of Kinesiology Sport, and Recreation, [2]University of Alberta, Neuroscience and Mental Health Institute, [3]University of Alberta, Department of Psychology

Recently we showed that 2 players in a simple online card game were better able to predict each other's intentions when cooperating than competing (Ma et al, 2023), which also generated spontaneous differences in mouse movement. However, online interactions using cursor icons have limitations in representing physical presence and conveying social cues. Thus, the current study aims to explore the influence of social contexts (competition and cooperation) on decision-making and intention prediction when two players engage in the same card game in person. Participant dyads played a card game six times under two social contexts: competing or cooperating (3 games each). When competing, only the player with the highest score earns points. When cooperating, participants split points evenly if they reach a combined threshold. Each game consisted of 8 turns, and each turn participants could obtain points by collecting goal-aligned cards and/or correctly guessing the other person's goal. We tracked card and guess point performance, as well as recorded gaze and hand movements. We predict in-person play will match online play: cooperating dyads will have higher guess scores and move more confidently (less time and hand distance traveled). Preliminary data revealed 1) As designed, card scores remain consistent across games and condition; 2) Guess scores improve across turns as information is acquired; 3) Guess scores are impacted by social context but not in the same way as the online study. Further analysis will reveal if in-person gameplay is genuinely altering the strategies players use to help and interfere with the communication of intention and will consider how gaze and hand movement contribute to this communication. Overall, we believe that cooperation facilitates the communication of intention during gameplay, but may be richer and more nuanced when people are playing in-person than when they played the same game online.

NSERC; AGRI

Saturday, May 18, 8:30 am – 12:30 pm, Pavilion

Decision Making: Perceptual decision making 1

23.458 DISSOCIATING SENSORY, DECISIONAL, AND METACOGNITIVE NOISE IN PERCEPTUAL DECISION MAKING

yunxuan zheng[1], Medha Shekhar[1], Kai Xue[1], Dobromir Rahnev[1]; [1]Georgia Tech

Perceptual decisions are subject to sensory, decisional, and metacognitive noise. However, dissociating between these different types of noise has proven challenging within conventional paradigms where the different types of noise can mimic each other. Here, we isolated each of these types of noise using an external noise paradigm where the same stimulus value could be generated from two categories. Subjects judged whether the number of dots presented on the screen was generated from a distribution with a higher or lower mean. These judgments are corrupted with a combination of sensory and decisional noise. In addition, subjects rated their decision confidence, with these judgments being corrupted by a combination of sensory and metacognitive noise. Subjects also performed a separate 2-alternative forced choice (2AFC) task where they identified which of two squares had more dots. 2AFC tasks involve minimal decisional noise and therefore allow one to estimate an upper bound of the sensory noise induced by our stimuli for each subject. Having estimated the level of sensory noise, we built a computational model that calculates the level of decisional and metacognitive noise in the external noise task. We found evidence for substantial decisional and metacognitive noise that, in some cases, exceeded the sensory noise. Moreover, metacognitive noise was higher than decisional noise, suggesting that confidence ratings may largely inherit the decisional noise of the initial perceptual judgment. Importantly, metacognitive noise became larger for confidence criteria further away from the decision criterion. This finding supports the recent lognormal meta-noise model of metacognition, which postulates that metacognitive noise is signal-dependent, such that it increases for more extreme evidence values (Shekhar & Rahnev, 2021). Overall, our study successfully dissociated sensory, decisional, and metacognitive noise, enabling quantification of each factor's impact on the corruption of perceptual decisions.

National Institute of Health (award: R01MH119189) and the Office of Naval Research (award: N00014-20-1-2622).

23.459 ESTIMATING AND INTEGRATING THE UNCERTAINTY OF NATURALISTIC STIMULI

Corey Plate[1,2] (corey.plate@nih.gov), Zoe Boundy-Singer[3], Corey Ziemba[1,2,3]; [1]Laboratory of Neuropsychology, Division of Intramural Research Program, National Institute of Mental Health, [2]Laboratory of Brain and Cognition, Division of Intramural Research Program, National Institute of Mental Health, [3]Center for Perceptual Systems, University of Texas at Austin

Perceptual decisions are accompanied by a sense of confidence. To usefully guide behavior, confidence should integrate all sources of decision-relevant uncertainty into a single decision reliability estimate. However, this metacognitive ability is not perfect, and can vary across individuals and contexts. Confidence is often studied in experiments where decision reliability is only varied through changing stimulus strength (e.g. only varying orientation in an orientation discrimination task), or not varied at all. Here, we sought to test the limits of the ability to appropriately integrate the uncertainty of complex, naturalistic stimuli into their perceptual confidence reports. We created a set of synthetic texture stimuli matched to the statistics of natural images and asked participants to judge in which direction their dominant orientation was rotated from vertical. Subjects reported their decision and binary confidence level with a single button press. Accuracy and confidence reports were incentivized by rewarding a large number of points for high confidence correct responses, but a substantial loss of points for high confidence errors. To control decision reliability, we varied both the mean and variance of orientation energy in the stimulus. We additionally adjusted the presence of higher-order, naturalistic pixel correlations, whose strength determines neural signatures of uncertainty in visual cortical regions downstream of V1 without affecting the orientation content. We fit the responses of subjects with the CASANDRE model for confidence, allowing us to assess whether different sources of uncertainty were appropriately integrated into a single confidence value. Our results indicate that subjects generally demonstrate the ability to represent and integrate naturalistic uncertainty into their perceptual confidence reports, but with some subjects exhibiting striking and idiosyncratic failures.

Grant: ZIA MH002928, Grant: NSF GRFP, Grant: NIH Grant EY032102

23.460 IMPROVING META-COGNITION WITH PRACTICE

Akash Raj G P[1] (akashrajgp@utexas.edu), Zoe Boundy-Singer[1], Corey Ziemba[1], Robbe Goris[1]; [1]University of Texas at Austin

Observers are aware of the fallibility of perception. When we feel confident in a perceptual interpretation, it is more likely to be correct. However, such metacognitive judgements are not flawless. In general, performance in difficult perceptual, cognitive and behavioural tasks improves with practice. Might metacognitive abilities improve with practice as well? Previous investigations yielded inconclusive results, potentially because the confidence judgments were not difficult enough, leaving little room for improvement in metacognitive ability. To test this idea, we leveraged insights offered by CASANDRE, a process model of perceptual confidence in which confidence reflects a subject's noisy estimate of the reliability of their perceptual decisions. The quality of this estimate is limited by the subject's uncertainty about the variable that informs their decision ('meta-uncertainty'). This meta-uncertainty can be manipulated experimentally by increasing the number of levels of stimulus reliability within a single experiment or by making the stimuli stochastic. We conducted a series of psychophysical experiments in which 38 subjects judged ambiguous visual stimuli and additionally reported their confidence. Each experiment consisted of 3000 trials and was completed over 2 sessions. We fit CASANDRE to each subject's data and studied the temporal evolution of meta-uncertainty. For the majority of the subjects (27 out of 38), meta-uncertainty decreased over the course of the experiment. The median fractional change in meta-uncertainty was a decrease by 83% across all subjects (Wilcoxon signed-rank test, $p < 0.001$). Meta-cognitive learning appears to be a general phenomenon: it was evident in both orientation- and texture discrimination tasks. As hypothesized, this learning was most prominent in experiments that involved many levels of stimulus reliability and/or stimulus stochasticity. Together, these results demonstrate that metacognitive ability can improve with practice, provided that the confidence judgments are difficult enough.

23.461 MULTIMODAL METAPERCEPTION: INSIGHTS FROM MULTISENSORY INTEGRATION

Nicola Domenici[1], Pascal Mamassian[1]; [1]École Normale Supérieure, PSL University, CNRS, Paris, France

Trusting our perception is crucial to interacting with the external world. Although most of the research on the topic focused on one single sense at a time, in our daily lives the brain is stormed by different sensory stimulations which are often fused into unitary percepts. To date, however, how the brain evaluates the trustworthiness of multimodal representations is still unclear. To investigate this, we developed an ensemble of visual, auditory, and audiovisual temporal bisection tasks. Stimuli consisted of three sequential events and participants had to report if the second event was closer to the first or last. Stimuli could be purely visual (100ms flash), purely auditory (100ms pink noise), or a combination of the two. When both sensory cues were presented, they were either synchronous or with an asynchrony of 100ms. Introducing such bimodal stimulations was essential to increasing sensitivity and creating consistent perceptual biases. Four stimulus difficulties were chosen, leading to 0.15, 0.35, 0.65, and 0.85 probability of responding 'closer-to-the-last-event'. These four difficulty levels were then placed in a confidence-forced-choice design, asking 15 participants to perform two consecutive perceptual decisions and report which one they felt was most likely to be correct. Notably, participants were instructed to base their confidence on visual information only, which made it possible to investigate how bimodal conflicts were mirrored at the confidence level. To better represent the dynamics of multimodal metaperception, we then used a confidence generative model to compare different predictions. Our results indicated that confidence evidence was generated from individual unisensory cues and subsequently combined with their corresponding sensory reliability. Surprisingly, participants computed confidence from their unimodal representations even after fusing them into an integrated percept. This suggests that confidence does not necessarily develop subsequently to multisensory integration, but that it still has access to the original unisensory evidence.

ANR grant no. ANR-18-CE28-0015

23.462 CONFIDENCE RESPONSES IN GLOBAL MOTION DISCRIMINATION TASK ARE WELL PREDICTED BY VISUAL RELIABILITY

Can Oluk[1] (coluk@utexas.edu), Martin Szinte[2], Guillaume S. Masson[2], Pascal Mamassian[1]; [1]Laboratoire des Systèmes Perceptifs, École Normale Supérieure, Université PSL, CNRS, Paris, France, [2]Institut de Neurosciences de la Timone (UMR 7289), Centre National de la Recherche Scientifique and Aix-Marseille Université, Marseille, France

Recent studies on perceptual confidence judgments have shown that human observers can be very accurate in estimating their own performance. However, these studies primarily focused on simple perceptual decisions about a single sensory feature. Here, we aim to understand the mechanisms of confidence formation in a global motion direction discrimination task where motion information must be integrated across a large number of local motions. Stimuli were composed of 468 drifting Gabors of identical size and spatial frequency, but random orientations (Amano et al., 2009, JOV). Motion coherence was manipulated as the fraction of Gabors whose velocity is consistent with a common global motion. In an 8-AFC task, participants reported the global motion direction of the multiple Gabors pattern. After two successive decisions, they judged which of them was more likely to be correct in a confidence forced-choice paradigm. We found that perceptual discrimination performance increases with motion coherence, and participants are consistently accurate in judging their own performance. We also found that when they make errors, participants often choose the direction directly opposite to true motion direction rather than random directions. Interestingly, participants are also more confident about these opposite direction reports. To model these results, we extended a classical motion integration model by incorporating strongly tuned inhibition that accounted for the opposite-direction effect. We also included a gain-like mechanism, ensuring the model's total response to stimulus (summed across all directions) increases with motion coherence. We next showed that our model can account for the observed variance in both perceptual and confidence judgments. We concluded that confidence responses are strongly tied to perceptual reliability and predominantly inherit characteristics from the associated visual mechanisms. These findings can serve as a foundation for understanding the mechanisms of confidence formation in global motion perception.

ANR grant no. ANR-19-NEUC-0003

23.463 REPULSION BIASES IN MOTION PERCEPTION ARE ATTENUATED BY WAITING

C. Stella Qian[1] (cheng.qian.2@warwick.ac.uk), Jian-Qiao Zhu[2], Nick Chater[3], Adam Sanborn[1]; [1]Department of Psychology, University of Warwick, [2]Department of Computer Science, Princeton University, [3]Warwick Business School, University of Warwick

Perceptual decision making does not terminate when the stimuli are off. Instead, previous work has found that a discrimination task could bias the later perceptual judgement task away from the reference in former task, which is termed the repulsion effect (Zamboni et al., 2016). The repulsion effect has been separately explained as an effect of self-consistent inference (Luu & Stocker, 2018), or as the result of optionally stopping an internal sampling process when the discrimination is the clearest (Zhu et al., 2023). Here, we explored the impact of more time between the two tasks on the strength of the repulsion effect, which has consequences for both theoretical positions. Participants were presented with a random dot motion stimuli and asked to do two tasks sequentially: a motion direction discrimination task relative to a probe stimulus presented after the motion stimuli and then a motion direction report task with the mouse cursor. We found better accuracy and faster reaction time as the difference between motion direction and probe stimulus became larger. The motion direction report task also had good accuracy and precision (absolute deviation M = 13.77 °, SD = 36.23°). The repulsion effect was replicated: response deviation conformed to a bimodal distribution when the discrimination judgement task was at chance level; the reported direction relative to the probe also conformed to a bimodal distribution. We tested the effect of adding a pause (1.5s and 2s) after the discrimination task results while keeping the probe onscreen. The repulsion effect decreased with a pause but did not disappear. The reference repulsion effect decreased with a delay between the discrimination task and the judgement task. We speculate that either self-consistency becomes less important with a delay, or the internal sampling process of the motion signal continues after the discrimination is made.

This work is supported by a European Research Council consolidator grant (No. 817492-SAMPLING).

23.464 FEEDBACK REDUCES BUT DOES NOT ELIMINATE CONFIRMATION BIAS

Marvin R. Maechler[1], Tobias H. Donner[2], Alan A. Stocker[1]; [1]University of Pennsylvania, [2]University Medical Center Hamburg-Eppendorf

Previous studies have demonstrated that a categorical choice can bias the subsequent integration and evaluation of sensory evidence: evidence that is congruent with the choice is given more weight in a subsequent perceptual judgment than incongruent evidence. However, participants in these studies did not receive feedback about the accuracy of their perceptual judgments and therefore produced these biases unknowingly. Thus it is unclear whether confirmation bias is the result of a suboptimal or uncalibrated inference strategy, and thus might be malleable by proper feedback. Or alternatively, whether confirmation bias is resistant to feedback and possibly the result of a normative strategy. To resolve this, we conducted a psychophysical experiment in which participants (N=8) had to estimate the unknown mean angular position from a stimulus sequence consisting of eight normal distributed samples of this mean (fixed variance; 11 different mean positions, angular positions on a circle equidistant [5dva] from fixation). Every trial required two responses from the participants. First, participants had to report whether the mean was clockwise or counterclockwise of a reference position after either seeing four or all eight samples (prospective and retrospective condition, respectively; randomly interleaved). Then they were asked to provide an estimate of the mean either directly after their categorical choice or after seeing the remaining samples of the sequence (prospective condition). Crucially, participants were shown the true mean position immediately after reporting their estimate, and also received monetary reward for estimates close to the true mean. We found that the incentivized feedback substantially reduced confirmation bias but without eliminating it. This shows that previously reported confirmation biases are, in part, caused by suboptimal decision strategies due to the missing feedback. Our results suggest that the remaining confirmation bias is the result of a deliberate decision process that considers other objectives in addition to estimation accuracy.

This work was supported by the NSF CRCNS grant IIS-1912232 to A.A.S., and Federal Ministry of Education and Research (BMBF) Project Nr. 01GQ1907 to T.H.D.

23.465 MATCHING IDENTICAL STIMULI WITH THE METHOD OF ADJUSTMENT PRODUCES COUNTER-INTUITIVE BIASES

Jiang Mao[1] (jiangmao@sas.upenn.edu), Alan Stocker[1]; [1]University of Pennsylvania

The method of adjustment is a widely used experimental procedure for measuring perceptual biases: subjects adjust a probe stimulus until it perceptually matches the test stimulus. Any resulting difference in feature value between probe and test is thought to reflect their relative perceptual difference (i.e. bias). Here, we challenge this view by empirically testing a fundamental prediction of this perceptual matching interpretation: if test and probe are identical then their match should be unbiased because perceptual distortions cancel out between identical stimuli. We run a psychophysical orientation matching experiment using the method of adjustment. Subjects (N=10) were presented with two orientation-filtered noise stimuli (0.5 deg radius Gaussian window, broadband spatial frequency) equidistant (4.5 deg) on either side of a central fixation mark with response-terminated display. Test and probe stimuli either had low or high noise (orientation) randomly assigned, which resulted in two same-noise and two cross-noise conditions. Test orientations were uniformly sampled from the full range (0 to 180 deg) in steps of 10 deg. Counter-intuitively, we found robust and characteristic repulsive biases away from cardinal orientations in both same-noise conditions. These biases cannot be perceptual because test and probe are identical. In addition, in the cross-noise conditions biases were not inverted when switching test and probe, which replicates a previous result. This is also incompatible with the view that the matches are purely perceptual. Finally, we demonstrate that a holistic matching model where perceptual inference and the matching process both simultaneously operate on all levels of the representational hierarchy, can accurately explains subjects' matching data from our experiment as well as the data from multiple previous studies. Together, our results demonstrate that biases measured with the method of adjustment are not purely perceptual but rather reflect optimal matching behavior according to a holistic objective.

23.466 SEQUENTIAL EVIDENCE ACCUMULATION IS A RESOURCE-RATIONAL PROCESS

Mengting Fang[1] (mtfang@sas.upenn.edu), Jiang Mao[1], Tobias H. Donner[2], Alan A. Stocker[1]; [1]University of Pennsylvania, [2]University Medical Center Hamburg-Eppendorf

Temporal integration of sensory information is an important aspect in many perceptual decision tasks. Our prior study found that a temporal break in the middle of an otherwise regular sequence of stimulus samples resulted in an increased weighting of the first post-break sample when performing a subsequent perceptual judgment. To further probe the origin of this phenomenon, we conducted a visual estimation experiment with varying durations and positions of the break in the sample sequence. Subjects estimated the angular position of an unknown generative mean based on 8 normally distributed stimulus samples of that mean (fixed variance). The samples were presented in rapid sequence. Subjects reported their estimates using a joystick and were given immediate feedback by showing the true generative mean. Subjects performed the task under three conditions. First, stimulus samples were presented without interruption. Second, in the "break" condition, sample sequences were interrupted after the 4th sample for varying break durations. Finally, in the "break" condition with different break positions, interruptions occurred randomly after either the 2nd, 4th, or 6th sample (1.75s duration). We calculated the relative contribution of each sample to the final estimate. Results showed that the weight of the first sample after the break monotonically increased with break duration, implying a continuous underlying process rather than a threshold mechanism. Furthermore, over-weighting persisted for other break positions in the sequence. We developed a normative model that assumes an active management of the trade-off between estimation performance on one hand and stimulus encoding and memory maintenance efforts on the other. This model effectively captures the temporal dynamics of the integration process from all our experimental data in conditions with or without breaks, which other models can not (e.g., bump attractor networks). Our results suggest that perceptual evidence accumulation is an actively controlled, resource-rational process.

This work was supported by the NSF CRCNS grant IIS-1912232 to A.A.S and Federal Ministry of Education and Research (BMBF) Project-nr: 01GQ1907 to T.H.D.

23.467 OPPOSING EFFECTS OF ONGOING ALPHA-BAND ACTIVITY ON MAGNO- AND PARVO-MEDIATED DETECTION

April Pilipenko[1] (apilipen@ucsc.edu), Vrishab Nukala[1], Jessica De La Torre[2], Montana Wilson[1], Maxwell Volkan[1], Marcella Williams[1], Jason Samaha[1]; [1]UC Santa Cruz, [2]UC San Diego

Spontaneous occipital alpha-band activity (8-12 Hz) has been shown to influence perceptual variability, leading one to report seeing a stimulus more often during states of weak alpha power, likely due to a shift in detection criterion. However, prior work has paid little attention to the specific stimulus properties mediating detection. In early vision, different stimulus properties are preferentially processed along the magnocellular (MC) and parvocellular (PC) pathways, which vary in their preference for spatial and temporal frequency and chromatic information. The goal of this study was to understand how spontaneous alpha power effects the detection of stimuli which are preferentially processed by either the MC or PC pathway. To achieve this, we used the "Steady/Pulsed Paradigm" which presented a brief, near-threshold stimulus in two conditions intended to bias processing to one or the other pathway. The pulsed condition presents the target stimulus atop a luminance pedestal, whose transient onset is believed to saturate MC firing and bias detection to the PC pathway. In the steady condition, which more closely resembles canonical detection paradigms, the luminance pedestals are present throughout the entire trial which is thought to evoke a sustained response from the PC pathway, biasing detection towards the MC pathway. Our results showed an interaction effect of alpha power on detection between the two conditions. While weak alpha power was predictive of seeing the stimulus in the steady condition (MC-biased), the opposite was found in the case of the pulsed condition (PC-biased). This interaction was driven by opposing alpha-related criterion shifts between the two tasks, suggesting that alpha oscillations may deferentially regulate excitability in the MC and PC pathways.

SATURDAY AFTERNOON POSTERS IN BANYAN BREEZEWAY

Saturday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

Motion: Optic flow

26.301 CAN PEOPLE LEARN THEIR UNIQUE RETINAL MOTION STATISTICS?

Jiaming Xu[1] (jiaming.xu@utexas.edu), Karl Muller[1], Kate Bonnen[2], Robbe Goris[1], Mary Hayhoe[1]; [1]UT Austin, [2]Indiana University

Sensory representations are adapted to the statistical regularities of the environment. In the case of retinal motion generated by self-motion, Matthis et al (2022) demonstrated that retinal flow results from the way the body moves during the gait cycle. This is because gaze is held stable in the environment during fixations and the eyes counter-rotate, transported by the body as it moves forward and sways during a step. This body motion is determined by its passive dynamics, and therefore differs between individuals. Do retinal motion statistics also differ between individuals? We examined the data of Muller et al. (2023) who tracked eye and body movements during natural locomotion. This allowed the calculation of the retina-centered motion patterns across a 90° region of the visual field for 7 subjects. Motion speed varies systematically across the visual field. Highest speeds were in the lower visual field where the mean of the distributions was 28.4° with a SD between subjects of 9.8°. Similarly, in the upper visual field, the mean speed was 13.1°, with a SD across subjects of 8.5°. Similar variability between subjects was found in the left and right visual fields. The average retinal motion directions, measured as a function of polar angles of the visual field, displayed a bimodal distribution with an over-representation of upward and downward motion. The first mode had a mean of 86.9° with a SD of 43.0°, and the second mode had a mean of 267.5° with a SD of 39.0°. Because the variability is substantial, and reflects motion statistics individuals are exposed to throughout experience, it seems likely that subjects learn their own motion statistics. Thus, individuals may have unique internal models of their own motion statistics, allowing prediction of their time-varying motion patterns through the gait cycle.

Supported by NIH grants EY05729 and EY032999

26.302 A COMPUTATIONAL MODEL FOR THE CONCURRENT RETRIEVAL OF OBJECT AND SELF-MOTION INFORMATION FROM OPTIC FLOW

Malte Scherff[1] (malte.scherff@uni-muenster.de), Markus Lappe[1]; [1]University of Münster

In a scene where self-motion and observer independent movement is present, the optic flow is complex because the global flow pattern resulting from ego-motion is locally confounded. Important information can still be obtained, although various biases were reported. The estimation of self-motion direction is affected by the direction of object movement and the perception of the object's trajectory by the ego-motion. While the underlying processes and their interactions remain largely unknown, some research proposes a sequential procedure. First, an initial estimation of the heading is made, followed by segmentation of areas of flow that do not match the estimation. The heading estimation is then refined, excluding these areas. The estimation could then aid in accounting for self-motion and disentangle the combined flow at the objects' retinal location for the retrieval of object properties. However, other research contests the need for prior heading estimation to estimate an object's trajectory. We present a computational model that computes retinotopic maps displaying the likelihood of heading directions given local flow. The likelihood distribution serves as a reliable indicator for the presence of independent object motion. Omitting certain parts of the flow, the objects' influence on heading estimation can be reduced. Furthermore, details about the objects' retinal position and movement can be extracted from the distributions. In summary, the model offers a concurrent estimation of both object properties and heading without either process relying on the outcome of the other. The model replicates various aspects of human performance, including the initial rise in heading estimation error with an increase in object speed, followed by a reduction in error due to enhanced object detection

performance. The perception of an object's trajectory is biased by its heading direction, in line with prior research. This bias occurs without requiring the completion of the heading estimation process beforehand.

26.303 NEURAL PROCESSING OF SCENE-RELATIVE OBJECT MOVEMENT DURING SELF-MOVEMENT

Xuechun Shen[1,2] (xs2154@nyu.edu), ZhouKuiDong Shan[3,2], Simon Rushton[4], ShuGuang Kuai[1,2], Li Li[3,2]; [1]East China Normal University, Shanghai, China, [2]NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, Shanghai, China, [3]New York University Shanghai, Shanghai, China, [4]Cardiff University, Cardiff, United Kingdom

Much research has examined how the visual system identifies scene-relative object movement during self-movement. Here we examined the related neural processing by identifying brain regions involved in this task. In a Siemens Magnetom Prisma Fit 3T MRI scanner, participants viewed through prism glasses a stereo display (9.5°Hx19°V) that simulated lateral self-movement (speed: 0.032 m/s) through a 3D volume composed of 63 randomly positioned red wireframe objects (depth: 0.55-1.05 m) with counter rotation of gaze. In the non-moving target condition, a yellow target object was positioned at 1/4 (near) or 3/4 (far) of the scene's depth range. In the moving target condition, the target at the near distance was given its retinal speed at the far distance and vice versa, causing the target to appear moving in the scene. The target movement was thus not defined by higher or lower speeds than the rest of the scene objects, and the moving and non-moving target conditions were equated for all retinal information. A control condition without simulated self-movement was also tested in which the scene remained static on the screen. During scanning, on each 2-s trial, participants were asked to report when the scene objects underwent a luminance contrast change to control attention (irrelevant to object movement identification). We identified known visual and optic flow areas as regions of interest (ROI) using standard localizers and performed multiple-voxel-pattern-analysis on the most active 300 voxels for each ROI. Across 20 participants, the decoding accuracy of scene-relative object movement versus no object movement was significantly higher than chance in higher-level dorsal visual areas V7 and MT+. Furthermore, these areas could successfully differentiate scene-relative object movement with and without simulated self-movement. Using well-designed visual stimuli, the current study reveals that areas V7 and MT+ play a crucial role in processing the scene-relative object movement during self-movement.

Supported by research grants from the National Natural Science Foundation of China (32071041, 32161133009, 32022031), China Ministry of Education (ECNU 111 Project, Base B1601), the major grant seed fund and the boost fund from NYU Shanghai, and UK Economic and Social Research Council (ES/S015272/1)

26.304 BIASES IN PERCEIVED OBJECT SPEED IN DEPTH DURING VISUAL SELF-MOTION

Anita Pandey[1] (anitapandeymusic@gmail.com), Björn Jörges[1], Laurence R. Harris[1]; [1]York University

During sideways movement, optic flow parsing - in which an object's speed is extracted from a scene viewed by a moving observer - has been shown to be incomplete, resulting in an overestimation of object speed when target and observer move in opposite directions (Jörges & Harris, 2022, AP&P 84: 25-46). Here we assess the efficiency of optic flow parsing for an object moving in depth while the observer is also moving towards or away from the object. Participants were immersed in a 3D virtual environment and asked to compare the speed of a sphere moving towards or away from them at 2, 3, and 10m/s relative to a ball moving sideways either while they were stationary or during visually simulated self-motion (either forwards or backwards at 6 or 10 m/s) evoking a range of retinal speeds. The speed of the sideways-moving ball was adjusted using an adaptive staircase to match the perceived speed of the sphere. Overall, flow parsing was incomplete. When the observer and sphere moved in opposite directions, the perceived speed of the sphere was greater than when the observer was static. Results were mixed when the observer and the sphere were moving in the same direction. The perceived direction of the sphere's movement depended on its retinal motion. We conclude that movement-in-depth flow parsing is incomplete. Our results are relevant to perceptual processing in various real-world settings, such as driving or crossing the road.

26.305 EFFECTS OF VISUAL CUES ON FLOW PARSING AND SIMULTANEOUS HEADING PERCEPTION

Zhoukuidong Shan[1,2] (sz68@nyu.edu), Yinghua Yang[3,2], Li Li[1,2]; [1]New York University Shanghai, Shanghai, China, [2]NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, Shanghai, China, [3]East China Normal University, Shanghai, China

Flow parsing, the ability to identify scene-relative object motion during self-motion, and heading perception are both essential components of locomotion. The causal inference hypothesis proposes that when observers cannot accurately identify object motion, they pool it with the background optic flow for heading estimation, leading to increased errors. Conversely, when they can correctly identify object motion, they segregate it from optic flow for heading estimation, leading to decreased errors. Here we tested this hypothesis and examined the effects of different visual cues on flow parsing and simultaneous heading estimation. On each 1-s trial, the display (90°×90°) simulated an observer translating toward a 3D cloud (0.56-5 m in depth) composed of 162 randomly positioned dots with an opaque window (16°×16°) containing 5 object dots at 8° or 16° offset from the simulated heading. These object dots moved laterally (0-0.8 m/s, lateral condition) at a fixed distance from or simultaneously approached (approaching condition) the observer. At the end of the trial, participants first reported whether the object dots moved left or right in the scene and then their perceived heading direction. Reducing simulated self-motion speed, increasing object lateral speed or dot density, decreasing the object position offset from the simulated heading, and incorporating binocular disparity cues all improved flow parsing performance (both accuracy and precision) but did not necessarily improve heading estimation. A significant correlation between the highest uncertainty in flow parsing and the largest heading errors supporting the causal inference hypothesis was observed only when the object dots underwent lateral motion while simultaneously approaching the observer. The results show differential effects of visual cues on flow parsing and heading perception, suggesting separate processing mechanisms for the two tasks. The causal inference hypothesis for flow parsing and heading

perception can explain the data for some but not all types of object motion.

This study was supported by research grants from the National Natural Science Foundation of China (32071041, 32161133009), China Ministry of Education (ECNU 111 Project, Base B1601), and by the major grant seed fund and the boost fund from NYU Shanghai.

26.306 HEAD AND EYE DYNAMICS ACROSS DIFFERENT NAVIGATIONAL GOALS

Andrés H. Méndez[1], Cristina de la Malla[1], Joan López-Moliner[1]; [1]Vision and Control of Action Group, Department of Cognition, Development, and Psychology of Education, Institute of Neurosciences, Universitat de Barcelona, Barcelona, Catalonia, Spain

The pattern of visual motion that we experience as we move (i.e., the optic flow) has been proposed as the substrate from which visual self motion is estimated. Until recently, this was studied in controlled lab settings assessing participant's abilities to detect the focus of expansion or heading direction from 2D optic flow patterns. Subsequent work has concentrated on studying the statistics and dynamics of the visual input in natural settings (e.g. Durant & Zanker,2020; Matthis et al 2022; Müller et al, 2023). Results suggest that flow patterns do not frequently resemble the expansive symmetric structure used in psychophysical studies. The characteristics of the visual input, however, are not independent from behavior, as organisms control the generation of visual information which will vary depending on the position of the eyes in the head, and the head in the world. Here we study head and eye dynamics of participants using head-mounted eye-trackers and inertial measurement units across three different navigational tasks: free locomotion, recreating a previous path and following someone else (n=4 for each condition). Results show that head stabilization does not vary significantly across conditions. Fixations towards the ground were less frequent than reported in previous studies in locomotion across different terrains. Across all conditions, fixations were close to the centre of the image but horizontal variability was larger in free locomotion. These findings suggest that self-motion may vary across different navigational goals in real-world scenarios leading to distinct retinal inputs and in ways that are relevant to the ongoing task. Knowledge of these dynamics can contribute to advancing computational models of visual processing and navigation.

This work is funded by grants PID2020-116400GA-I00 and PID2020-114713GB-I00 funded by MCIN/AEI/10.13039/501100011033 to CM and JLM, respectively. AM is supported by the Maria de Maeztu grant PRE2021-097688 assigned to the Institute of Neurosciences of the University of Barcelona (MDM-2017-0729-21-1)

26.307 THE INFLUENCE OF ANTICIPATION ON HUMAN HEADING PERCEPTION

Roselind Si[1] (xsi25@colby.edu), Oliver Layton[1]; [1]Colby College

Self-motion relative to a stationary environment generates a pattern of motion on the eye known as optic flow. Straight self-motion without eye movements or other sources of rotation creates an optic flow pattern where the motion radiates with respect to a singularity point that specifies the direction of travel (heading). Decades of research have characterized the accuracy with which humans are capable of perceiving heading from optic flow. In the typical experiment, subjects judge their heading after viewing optic flow simulating self-motion along a constant direction. A key assumption is that the heading judgment reflects human perception of the optic flow during the current trial. Recent work, however, has shown that the heading on previous trials (Sun, Zhang, Alais, & Li, 2020) and changes in heading hundreds of milliseconds before the judgment period (Ali, Decker, & Layton, 2023) influence the heading judgment. In the present work, we investigate how the ability to anticipate the heading in the current trial based on the preceding trials on heading judgments. We used a modified version of the paradigm of Ali et al. (2023) wherein subjects viewed straight-forward self-motion and the heading would switch leftward or rightward during the trial. When the direction of the heading switch predictably alternated sides across successive trials, human heading judgments exhibited bias toward the side of the heading switch on the preceding trial. We found no such bias when the side of the heading switch was randomized. When we created an imbalance in the switch directions (e.g. left switches more likely than right switches), judgments exhibited bias toward the side with fewer heading switches. Our findings suggest that expectations about heading based on the predictability and distribution of recently experienced headings may influence human heading perception.

26.308 MODELING OPTIC FLOW TUNING IN MSTD WITH CONVOLUTIONAL NEURAL NETWORKS

Oliver Layton[1], Scott Steinmetz[2]; [1]Colby College, [2]Sandia National Laboratories

Performance-optimized convolutional neural networks (CNNs) have emerged as highly effective models at predicting neural responses in brain areas along the primate ventral stream, but it is largely unknown whether they effectively model neurons in the complementary primate dorsal stream. We explored how well CNNs model the optic flow tuning properties of neurons in dorsal area MSTd and we compared our results with the Non-Negative Matrix Factorization (NNMF) model proposed by Beyeler, Dutt, & Krichmar (2016), which successfully models many tuning properties of MSTd neurons. To better understand the role of computational properties in the NNMF model that give rise to MSTd-like optic flow tuning, we created additional CNN model variants that implement key NNMF constraints — non-negative weights and sparse coding of optic flow. While the CNNs and NNMF models both accurately estimate the observer's self-motion from purely translational or rotational optic flow, NNMF and the CNNs with nonnegative weights yield substantially less accurate estimates than the other CNNs when tested on more complex optic flow that combines observer translation and rotation. Despite their poor accuracy, however, neurons in the networks with the nonnegativity constraint give rise to tuning properties that align more closely with those observed in primate MSTd. Interestingly, the addition of the sparsity constraint has a negligible effect on the accuracy of self-motion estimates and model tuning properties. Across all models, we consistently observe the 90-degree offset in the preferred translation and rotation directions found in MSTd neurons, which suggests that this property could emerge through a range of potential computational mechanisms. This work offers a step towards a deeper understanding of the computational properties and constraints that describe optic flow tuning primate area MSTd.

SNL is managed and operated by NTESS under DOE NNSA contract DE-NA0003525

26.309 TRACKING VISUAL TARGETS DURING SIMULATED SELF-MOTION

Matt D. Anderson[1] (matt.anderson@berkeley.edu), Jorge Otero-Millan, Emily A. Cooper; [1]University of California, Berkeley, [2]University of California, Berkeley, [3]University of California, Berkeley

Movements of the head and body create visual motion across the retina, often called optic flow. When we try to fixate on a visual target embedded in optic flow, eye movements are required to compensate for the target's retinal motion and stabilize its image on the fovea. Previous research (Niemann et al., 1998; Lappe et al., 1998), however, has demonstrated that passive tracking of targets embedded in optic flow can be undercompensatory (gain < 1): stimulus speed exceeds eye speed and the foveal image is not stable. Undercompensation was observed in experiments simulating forward self-motion over a rendered ground plane. Features on the ground plane were rendered at a fixed scale and distributed uniformly in world-space, such that perspective projection created a density gradient on the retina. In this scenario, the average of perifoveal motion signals is biased towards slower speeds because feature density is maximized towards the horizon where speed equals zero. If perifoveal features influence tracking speed, this could explain the low gain. However, the contrast energy of natural images is, on average, scale-invariant (i.e., fractal). Thus, natural scenes, should not contain the same gradient in visual feature density. We thus predicted that passive visual tracking during forward translation over naturalistic ground textures should show a gain closer to unity. We tested this prediction by asking observers to passively view forward-translation optic flow with different ground density patterns. We measured eye movements with a video-based eye tracker. Ground-plane textures were of uniform feature density in either world or retinal coordinates. Our results suggest that tracking gain was higher for stimuli with uniform retinal density. If feature tracking during self-motion is sensitive to the spatial distribution of motion signals, then investigating the spatial properties of optic flow in natural scenes is an important element of modelling oculomotor behavior.

This was work partially supported by Alcon.

26.310 EXPLORING TOP-DOWN INFLUENCES ON ILLUSORY SELF-MOTION PERCEPTION (VECTION) IN YOUNGER AND OLDER ADULTS

Brandy Murovec[1,2] (bmurovec@torontomu.ca), Julia Spaniol[1], Behrang Keshavarz[1,2]; [1]Toronto Metropolitan University, [2]KITE-Toronto Rehabilitation Institute

Vection, defined as the sensation of self-motion in the absence of physical movement, is a critical component for an immersive Virtual Reality experience. Research in this domain has predominantly focused on strategies to modulate vection by manipulating physical properties of the visual stimulus (bottom-up factors) with younger-to-middle aged adults. That said, little research has been done investigating the role of cognitive (top-down) factors on vection, and further how these manipulations may differentially influence participants of different age groups. In the current study, we investigated the role of expectation and stimulus realism as cognitive factors for vection in a sample of 48 younger and 46 older adults. Through manipulation of the study instructions, participants were led to believe that they were either likely or unlikely to experience vection before they were exposed to a rotating visual stimulus aimed to induce circular vection. Stimulus realism was manipulated by disrupting the global consistency of a naturalistic visual stimulus (360° photograph), resulting in two image conditions (intact, scrambled). The speed of the stimulus was varied (faster, slower) to further investigate the influence of top-down effects under stronger and weaker vection intensities, respectively. During each trial, three measures of vection were recorded: onset time, duration, and intensity. Results indicated that image realism and speed affected vection: intact and faster-moving stimuli substantially increased vection relative to the other stimulus conditions for all measures. An interaction between speed and age showed reduced vection onset times in younger adults compared to older adults, but only for faster moving stimuli. No other age-related differences and no meaningful role of expectation was found. Taken together, our study demonstrates that certain bottom-up and top-down factors can influence vection independently as well as additively, whereas age does not seem to be a major factor for vection.

Saturday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

Temporal Processing: Neural mechanisms, models

26.311 COMPARING NEURAL RESPONSES TO NATURAL VISUAL STIMULI AND ELECTRICAL STIMULATION OF VISUAL CORTEX

Denise Oswalt[1] (denise.oswalt@pennmedicine.upenn.edu), Michael Beauchamp[1], Daniel Yoshor[1]; [1]University of Pennsylvania

It has long been known that electrical stimulation of early visual areas evokes the perception of a small flash of light, called a phosphene. Across the past several decades, researches have been working to translate this phenomena into a device that can restore limited visual function in the blind. Recent technological developments have allowed for a new generation of implantable devices. Despite these hardware advances, little has been done to develop electrical stimulation paradigms capable of recreating the experience of natural vision. To bridge this gap, we fist need to better understand the differences in neural activity driven by natural visual stimuli verses that driven by electrical stimulation. We presented small visual stimuli in the receptive field of a stereo-EEG electrode implanted in V1 of a patient undergoing invasive monitoring for epilepsy. The stimuli included small dots, a checkerboard, and natural images. We also delivered bursts of electrical stimulation to the same electrode at currents sufficient evoke phosphene perception. We compared the amplitude, spread and shape of evoked responses following stimulation by these different modalities. Both visual and electrical stimuli resulted in perception in the same area of visual space, but the characteristics of the electrophysiology varied widely between the two forms of input. The shape of the neural response evoked by natural visual inputs followed a classic progression of a brief upward deflection, which peaked at 135±26uV, 102±10ms after stimulus onset, followed by a return to baseline. Comparatively, responses evoked by electrical stimulation had an early sharp positive deflection (458±21uV, 11±2ms), followed by a strong negative component, then a slower positive deflection that peaked at 138±16uV and 82±18 ms. Both visual and electrical stimuli increased power in the high gamma range (80-150Hz). Correlation between signals collected on neighboring contacts increased more following electrical stimulation than visual stimulation.

26.312 THE EFFECT OF FAST FLICKER ADAPTATION ON CONTRAST DISCRIMINATION

Jaeseon Song[1] (jaeseon.song@uga.edu), Bruno Breitmeyer[2], James Brown[3]; [1]University of Georgia, [2]University of Houston, [3]University of Georgia

For many decades, aspects of spatiotemporal vision were studied via visual-masking and spatiotemporal-flicker measures. More recent studies used fast flicker adaptation (FFAd) to explore the relation between spatial and temporal vision. FFAd is thought to diminish the magnocellular (M) pathway's sensitivity, but only at low spatial frequencies ≤ 2 cpd, as suggested Kaneko et al. (2015) (and supported by Arnold et al., 2016). We compared FFAd effects of four types of flicker conditions: baseline (no flicker), uniform-field flicker (UFF), and pattern flicker (PF) at flicker contrasts of 10% and 50%. We measured FFAd effects by measuring contrast discrimination thresholds (ΔCs) for test-Gabor pedestals at each of four contrasts: 0, 10, 30, and 50%. The spatial frequencies (SFs) of test Gabors were a low, 0.5 and a high, 5 cpd. As expected, we found 1) that for both test-Gabor SFs, ΔCs increased progressively as the contrast of the Gabors increased from 0 to 50%; and 2) that compared to baseline (no flicker), UFF and both PF contrasts significantly increased ΔCs only with the 0.5 cpd test Gabor; with the 5 cpd test Gabor, all FFAs showed minimal differences from the baseline, except at a 10% test contrast. Interestingly, for the 0.5 cpd Gabor, progressively enhanced FFAd effects (larger ΔCs) were obtained as the PF contrast increased from 0% (UFF), through 30%, to 50%. Related to this, another interesting finding was a strong correlation among the obtained FFAd effects across all test-Gabor contrasts when the 0.5 cpd test Gabor but not the 5 cpd test Gabor was used. Overall, these results indicate that contrast-discrimination thresholds obtained with suprathreshold contrast pedestals may provide a more sensitive measure of FFAd effects than do measures of contrast-detection thresholds used in previous studies.

26.313 DELAYED NORMALIZATION ACCOUNTS FOR TEMPORAL DYNAMICS IN VISUAL AND SOMATOSENSORY CORTICES

Luhe Li[1] (ll3981@nyu.edu), Ilona M. Bloem[1,3], Stephanie Badde[4], Wouter Schellekens[5], Natalia Petridou[5], Michael S. Landy[1,2], Jonathan Winawer[1,2]; [1]Department of Psychology, New York University, [2]Center for Neural Science, New York University, [3]Netherlands Institute for Neuroscience, Amsterdam, the Netherlands, [4]Department of Psychology, Tufts University, [5]Radiology Department, Center for Image Sciences, UMC Utrecht, the Netherlands

Neural responses to sensory stimuli exhibit complex temporal dynamics. Recent studies demonstrated that temporal dynamics of visual neural responses, including sub-additive temporal summation and response reduction with repeated or sustained stimuli (adaptation), are well characterized by a delayed-normalization model. Do similar principles of temporal dynamics apply more generally to sensory coding? Here, we apply similar methods and modeling to the tactile domain. We used fMRI to measure responses to tactile stimuli within somatosensory cortex. Participants were presented with vibrotactile stimuli (110 Hz) simultaneously to all five finger pads of the non-dominant hand while they visually fixated a dot on the screen. We used an event-related design, in which on a single trial the vibrotactile stimuli either varied in duration (for a single stimulus) or in inter-stimulus interval (for pair stimuli), comparable to the visual studies. The single pulse durations and paired pulse intervals were 50, 100, 200, 400, 800, or 1200 ms. We estimated the underlying population neural response time courses from the fMRI BOLD response using deconvolution and computed the area under the response curve to estimate the total response. The results showed clear sub-additive temporal summation, comparable to responses in visual cortex. We modeled the neural time courses for all single pulse durations and paired pulse intervals with the delayed-normalization model and found that this model outperformed a linear prediction. Importantly, our results reveal similar temporal dynamics for visual and tactile neural responses; both are best explained by the delayed-normalization model. These findings suggest that delayed normalization constitutes a canonical neural computation across modalities.

EY08266 (MSL), R01MH111417(NP & JW)

26.314 NEURAL EVIDENCE FOR A TWO-STAGE MODEL OF CONSCIOUS PERCEPTION

Maëlan Q. Menétrey[1] (maelan.menetrey@epfl.ch), Michael H. Herzog[1], David Pascucci[1]; [1]Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Are we conscious at any moment? Postdictive effects, where later events influence the perception of earlier ones, suggest a two-stage model of conscious perception: extended periods of unconscious information processing (stage 1) occur before a discrete conscious percept emerges (stage 2). These effects are evident in the Sequential Metacontrast Paradigm (SQM), where a sequence of diverging lines is displayed. When one line is spatially offset (i.e., vernier), the offset appears to propagate along the entire stream, even if the vernier is actually presented later. Furthermore, when two verniers with opposite offsets are presented in the stream, they cancel each other and cannot be individually reported. Here, we analyzed electroencephalography (EEG) data during the SQM, presenting either one or two opposite verniers at different locations within the stream, and a control condition with only straight lines. We sought to disentangle the neural correlates of stage 1 —when individual verniers are unconsciously processed— and stage 2 —when an integrated conscious percept emerges. Using temporal decoding, we found that the presence of vernier in the stream (one or two verniers vs. none) can be decoded from occipital EEG activity patterns. Moreover, the earlier the vernier is presented, the earlier it can be decoded. Since this holds true even for conditions with two verniers, where neither is consciously accessible, we interpret this as neural correlates of unconscious processing (stage 1). We then showed that behavioral reports can be instead decoded from parietal EEG activity patterns, emerging later. Since this decoding relates to the reported percepts, we interpret this as neural correlates of conscious perception (stage 2).

This work was supported by the Swiss National Science Foundation (grant number 325130_204898, "Basics of visual processing: the first half second"; grant numbers PZ00P1_179988 and PZ00P1_179988/2)

26.315 THE (LACK OF) CORRELATION BETWEEN EVOKED AND SPONTANEOUS BRAIN OSCILLATIONS: AN INDIVIDUAL DIFFERENCE APPROACH

Sihan xiong[1] (sihanxiong7@outlook.com), Songlin Qiao[1], Matteo Valsecchi[2], Jing Chen[1]; [1]shanghai university of sport, [2]Universitá di Bologna

Rhythmic visual stimuli can evoke brain oscillations (i.e., the steady-state visual evoked potentials), while the brain also spontaneously generates endogenous oscillations. Previous research has suggested using rhythmic visual stimuli to entrain these spontaneous oscillations (Herrmann et al., 2016; Notbohm et al., 2016). However, it is unclear if brain oscillations induced by rhythmic stimulation are identical to spontaneous brain oscillations. This study adopted an individual difference approach to investigate whether evoked and spontaneous oscillations originate from the same neural mechanisms. If they do, a correlation across observers would be expected. We collected a large EEG dataset (N = 30), recording both evoked and spontaneous brain oscillations. For evoked oscillations, EEG activities were recorded while observers viewed periodic flickering stimuli at specific frequencies (i.e., 2Hz, 4Hz, 6Hz, …, or 30Hz). Each trial lasted 9 seconds, with 6 trials tested for each frequency. For spontaneous oscillations, resting-state EEGs were recorded for 4 minutes while observers fixated on a spot in the center of the screen. We analyzed cross-observer correlations within and between conditions at each frequency band (e.g., delta, theta, alpha, and beta). For evoked oscillations, we spit the trials into odd and even group, finding a high correlation (r=0.873, p<.01, averaged across all frequency bands). Similarly, for spontaneous oscillations, we observed a high correlation across two recording blocks (r=0.768, p<.01). These results indicate that our measurements for both evoked and spontaneous oscillations are highly reliable across individuals. However, the correlation between evoked and spontaneous oscillations was minimal (r=0.072, 95% confidence interval = [-0.30 0.45]). This lack of correlation suggests that evoked and spontaneous oscillations represent distinct forms of brain activity.

26.316 DISTINCT MECHANISMS ACCOUNT FOR PERCEPTUAL SUPPRESSION FORWARDS AND BACKWARDS IN TIME

Michael Epstein[1] (mlepstein88@gmail.com), Sophie Kovacevich[1], Rachel Denison[1]; [1]Boston University

Background: Temporal interactions between successive stimuli occur across time ranges of hundreds of milliseconds. We previously found decreased perceptual sensitivity to targets that were preceded or followed by high-contrast vs. low-contrast non-targets, consistent with a principle of normalization across time. However, the mechanisms underlying these forward and backward temporal interactions are unknown. Goal: Here we used an orientation reproduction task to test whether perceptual suppression by temporal competitors is due to reduced orientation precision or a total loss of orientation information. Methods: Participants viewed a sequence of two Gabors (T1 and T2), each randomly oriented from 0-180 degrees. At the end of each trial, a response cue indicated the target, instructing participants to reproduce the orientation of either T1 or T2 using an adjustable probe. Stimuli were presented with 50 ms durations and a 250 ms stimulus onset asynchrony, with no temporal uncertainty, in the lower right screen quadrant. Contrasts were independently manipulated so that each target was either high (64%) or low (16%) contrast. Participants also reported if they missed (i.e., had no awareness of) one of the stimuli; missed trials were excluded from further analysis. Mixed models combining Gaussian and uniform error distributions were fit to the orientation estimation data using MemToolbox to determine how much of the perceptual suppression was explained by reduced precision vs. increased guessing. Results: Orientation estimates for both targets had greater error when paired with higher vs. lower contrast non-targets. Critically, the mixed modeling showed that these impairments arose by different mechanisms for each target. Higher non-target contrast predominantly reduced precision for T1 targets but predominantly increased guess rate for T2 targets. Conclusions: The findings suggest two distinct mechanisms for competition across time between successive stimuli, with implications for phenomena such as temporal crowding and models of dynamic perception.

Funding for this project was provided by a National Institute of Health National Eye Institute fellowship to M.E. (1F32EY033625), Boston University Undergraduate Research Opportunity Program funding to S.K. and Boston University startup funding to R.D.

26.317 A DYNAMIC NORMALIZATION MODEL WITH TEMPORAL RECEPTIVE FIELDS CAPTURES PERCEPTUAL SUPPRESSION BY PAST AND FUTURE STIMULI

Angus Chapman[1], Michael Epstein[1], Rachel Denison[1]; [1]Boston University

Background: Perception can be influenced by the temporal context from both past and future stimuli. Recent work has shown that the perception of a target stimulus is affected by the contrast of not only earlier but also later stimuli, demonstrating bidirectional temporal suppression over hundreds of milliseconds. Temporal normalization is the idea that neural responses are divisively suppressed by their past activity, and has been used to explain adaptation. However, models incorporating temporal normalization have mostly focused on predicting neural time courses, rather than perception, and cannot capture the perceptual suppression of earlier stimuli by later ones. Goal: Here we tested whether temporal receptive fields could be used as a mechanism of temporal normalization, and whether this mechanism could capture bidirectional temporal suppression. Methods: We incorporated temporal receptive fields into the Normalization Model of Dynamic Attention (Denison, Carrasco, & Heeger, 2021), by allowing for excitatory and suppressive drives in a population of modeled sensory neurons to be affected not only by the current stimulus input, but also by recent stimulus history. We simulated the model's response to a sequence of two target orientations that independently varied in contrast. Across simulations, we manipulated the durations of both the excitatory and suppressive temporal receptive fields via their exponential time constants. Results: A model without extended temporal receptive fields failed to capture contrast-dependent suppression. Adding suppressive temporal receptive fields to the model yielded forward suppression, such that a higher contrast first stimulus reduced responses to the second stimulus. Lastly, adding both excitatory and suppressive receptive fields yielded bidirectional suppression, with higher contrasts of each stimulus suppressing model responses to the other. Conclusion: Integrating temporal receptive fields into a dynamic normalization model captured contrast-dependent suppression both forwards and backwards in time, furthering the goal of developing real-time process models of dynamic perception.

Funding for this project was provided by a National Institute of Health National Eye Institute fellowship to M.E. (1F32EY033625), and Boston University startup funding to R.D.

26.318 NEURAL DYNAMICS OF GROUPING EXPLAIN PROPERTIES OF SERIAL DEPENDENCE IN ORIENTATION

Gregory Francis[1] (gfrancis@purdue.edu), Mauro Manassi[2]; [1]Purdue University, [2]University of Aberdeen

In serial dependence, our percepts (such as orientation) are biased towards previously presented stimuli. Serial dependence was proposed as a mechanism to facilitate perceptual stability and improve efficiency by compensating for variability in visual input. Despite a large amount of research on the topic, the field lacks specific underlying mechanisms to explain how this attractive bias occurs. Here, we show that serial dependence in orientation perception is naturally predicted with a biologically plausible model of visual perception that identifies Gabor stimulus orientation by observing the interplay between perceptual filling-in of surfaces and grouping of orientation-sensitive contours. At its core, the model hypothesizes a serial process that searches for the grouping orientation that maximizes the difference between dark and light filled-in surfaces. A reasonable starting point for this search process is the orientation in recent visual experience. With this assumption, we show model simulations that account for many basic characteristics of serial dependence. Serial dependence gradually decays over time for trials further in the past (up to 3 trials back; temporal tuning). Serial dependence strength peaks at around 15 degrees of relative orientation difference between the current and previous trials and it decreases as the relative difference differs (feature tuning). Consistent with recent proposals of serial dependence as an optimizing strategy, the model also predicts that orientation discrimination performance improves when serial dependence occurs. The model can also successfully explain attentional tuning and spatial tuning properties as being due to changes in the probability of using the grouping orientation from the previous trial as the start of the search process for the current trial. Together, our model provides extensive evidence for a biologically plausible grouping mechanism that explains the basic characteristics of serial dependence effects.

26.319 POOLING AND SEGREGATION ACROSS DIFFERENT TIME SCALES

Ilanit Hochmitz[1] (ilanit57@gmail.com), Amit Yashar[1], Ahmad Abu-Akel[1], Yaffa Yeshurun[1]; [1]University of Haifa

Identifying a single item presented within a sequence of items requires segregating its signal from that of the preceding and/or succeeding task-irrelevant items. In contrast, representing a global aspect of the sequence, in terms of its average orientation for example, requires information pooling across time. Together, these two opposing processes (segregation and pooling) enable us to balance the need to mind individual events while maintaining a general sense of our surroundings. Thus, understanding the relations between the two and how they unfold across time has important theoretical and practical implications for theories of human visual perception. Typically, pooling and segregation were studied separately. Here, we used the same stimuli and orientation estimation procedure to examine both processes and to determine whether and how they vary across different temporal scales. In five experiments, the participants viewed a sequence of three oriented items. Depending on the experiment, they had to either reproduce the orientation of the second item in the sequence (segregation) or report the average orientation of all three items (pooling). The SOAs were either short (≤130ms; short temporal scale) or long (150-475ms; long temporal scale). Mixture modeling analyses of the error distribution revealed a dissociation between pooling and segregation and their temporal scales. With the long temporal scale, the SOA affected encoding precision in both segregation and pooling tasks, but in a very different manner. With the short temporal scale, the SOA also had distinctive effects on the two tasks. While the SOA affected the guessing rate without affecting precision in the segregation task, it substantially affected precision without affecting the guessing rate in the pooling task. These results suggest that temporal segregation and pooling are mediated by different processes and that each of these processes further reflects different mechanisms when considered over a short vs. long temporal scale.

26.320 ANALYSIS OF THE ERG OFF-RESPONSE

Christopher Tyler[1] (cwt@ski.org); [1]Smith-Kettlewell Eye Research Institute

Introduction. The electroretinogram (ERG) is a powerful non-invasive assay of the functional integrity of the human retina, providing measures of the retinal receptor potential and the bipolar cell function, together with signals attributable to the inner plexiform layer of the amacrine retinal ganglion cells. An accurate model of ERG dynamics is important for understanding the multifold processes of light transduction to ecologically useful signals by the retina. Methods. A neuroanalytic approach to modeling the human rod ERG is based on the general principle of a dynamic serial/parallel model of the kinetics of each component contributing to the ERG, such as the receptor potential, the bipolar response, the ganglion-cell response, etc. This approach is now extended from the brief flash response to account for the light-adapted On/Off step response. A direct model with a linear-waveform Off-response is compared with a model incorporating separate half-wave rectifying generators deriving the On and Off bipolar responses. Both models incorporate adaptive gain control of the Off-response amplitude. Results. This new model provides a substantially better match than previous models of rod responses in six different waveform features of the canonical ERG flash intensity series, together with the compound nonlinearities of the On/Off ERG step responses to white light. The comparative analysis shows that the half-wave rectifying model generates an Off-response waveform that is too slow, when combined with the Off-step of the receptor potential. Only the direct form with adaptive gain control can match the dynamic properties of the physiological recordings. Conclusion. The provision of an accurate model of the On and Off pathways of the retina provides a significant step towards more accurate quantification of retinal processing deficits than is available from the standard peak statistics of the ERG flash responses.

Supported by NEI grant EY30155

Eye Movements: Learning, expertise, context and faces

26.321 SACCADE TARGET STATUS INFLUENCES THE REFERENCE FRAME OF OBJECT-LOCATION BINDING

Tzu-Yao Chiu[1] (chiu.315@osu.edu), Julie D. Golomb[1]; [1]The Ohio State University

Despite constant retinal image shifts caused by saccadic eye movements, humans perceive external environments as stable. To achieve perceptual stability across saccades, the visual system must update spatial information (remapping) and keep track of non-spatial information bound to each location (object-location binding). Here, we investigated whether saccade target status affects the reference frame of object-location binding across saccades. Previous studies examining the reference frame of object-location binding showed that objects that appear at peripheral, non-saccade target locations are naturally bound to retinotopic, not spatiotopic, coordinates. But real-world saccades are generally directed towards objects of interest: Might trans-saccadic object-location binding occur in more ecologically relevant spatiotopic coordinates for objects with saccade target status? We adopted a modified spatial congruency bias paradigm, in which participants were asked to judge if two objects presented sequentially have the same or different identity. In Experiment 1, we inserted a saccade between object presentation, directed either towards the first object's location (saccade-target condition) or to another location (saccade-elsewhere condition). We found purely retinotopic object-location binding in the saccade-elsewhere condition (replicating previous reports), but co-existing spatiotopic and retinotopic binding in the saccade-target condition, indicating that saccade target objects can be additionally bound to spatiotopic coordinates across saccades. In Experiment 2, we added a saccade-away condition, in which participants saw the first object at fovea and launched a saccade away from that location. We again found significant spatiotopic and retinotopic object-location binding in the saccade-target condition, but only retinotopic binding in the saccade-away condition, suggesting that it is the saccade target status specifically that allows an object to be additionally bound to spatiotopic coordinates. Overall, the results highlight the important role of saccade targets in maintaining stability across saccades, allowing for object-location binding to be encoded and/or remapped in more ecologically relevant spatiotopic coordinates.

This study was funded by NIH R01-EY025648 (JG).

26.322 CONTEXTUAL SACCADIC ADAPTATION : YOU CAN SEE IT BUT YOU CAN'T LEARN FROM IT

Maxime Martel[1] (maxime.martel@univ-lille.fr), Laurent Madelain[1]; [1]University of Lille - SCALab, CNRS, UMR 9193

Contextual saccadic adaptation is investigated using a variant of the double-step paradigm, in which two directions of intra-saccadic steps are signaled by two different contexts. This allows to simultaneously inducing two distinct saccadic adaptations. Surprisingly, effective contextual adaptation is found when using spatially relevant cues but not when the target color and shape serve as context. Here we aim at further exploring this phenomenon to better describe the mechanisms underlying contextual motor learning. We tested eight different contextual cues to signal the intra-saccadic steps in a contextual double-step paradigm (n=80): the duration of a visual stimulus, the lateralization of a sound, various statistical regularities across trials, symbolic cues, as well as the amplitude of the first step or the target color and shape. We found robust systematic contextual learning under the amplitude condition, while no learning occurred with any of the other cues. This absence of contextual learning further confirms that predicting the intra-saccadic step strongly depends on the nature of the context. In two additional experiments replicating the previous conditions of target color and shape (n=10) and symbolic cues (n=10), participants were periodically prompted (approximately every ten trials) to explicitly report the contextual cue they had just experienced. Again, no systematic contextual adaptation was observed despite participants achieving a 91% correct report. This dissociation between perceptual reports and motor tasks involving the same visual information aligns with previous results on the constraints for contextual motor learning. It has been proposed that motor cues, as opposed to purely perceptual cues, are necessary for contextual arm or eye movement adaptation, underscoring the necessity of having contexts associated with different motor states. However, it remains unclear whether it is the planning or the execution component of the saccade that enables contextual saccade adaptation.

26.323 EFFECTIVE AND NON-EFFECTIVE CUES FOR PROBABILISTIC CONTEXTUAL VISUOMOTOR ADAPTATION

Anna Montagnini[1] (anna.montagnini@univ-amu.fr), Laurent Madelain[2], Guillaume S. Masson[1], Vanessa Carneiro Morita[1]; [1]Institut de Neurosciences de la Timone, CNRS & Aix-Marseille Université, [2]ScaLab, CNRS & Université de Lille, France

Saccades and smooth pursuit eye movements adapt efficiently to environmental regularities. In the presence of a probabilistic bias favoring a particular visual motion direction, a robust linear correlation between human anticipatory smooth eye movements (aSEM) and direction probability has been previously reported. In real life, regularities are not immutable, and relevant behavioral contingences vary depending on the context. Importantly, sensorimotor learning depends on the nature of the specific cues signaling the context (Azadi and Harwood, 2014; Howard et al. 2013). Here we ask whether aSEM adapt efficiently to a probabilistic direction bias signaled by either a color cue or a motion-direction cue. Participants had to accurately track a small visual target moving either rightward or leftward (11°/s). In two experiments, unbeknownst to the participant, the probability of target motion direction (say P(Right)) was paired with the cue type and was manipulated across blocks. In practice, in experiment 1, when the fixation stimulus (color cue) was green the probability of rightward motion was P(Right|green)=0.5,0.25 or 0.75, while with the red cue the probability was the complementary one, P(Right|red)=1-P(Right|green). In a second experiment, the probability of target rightward motion was paired with the target vertical motion direction during the previous motion epoch: P(Right|target UP)=0.5,0.25 or 0.75; P(Right|target DOWN)=1- P(Right|target UP). Our data revealed two radically different oculomotor behaviors. aSEM did adapt to the contextual probabilistic contingency schedule signaled by the motion-direction cue but not when using the color cue. These results are well predicted by alternative probability-learning models based on different learnt probabilistic functions. Importantly, the initiation of visually-guided horizontal smooth pursuit was also modulated differently by the cue-conditional probability in the two experiments. Overall, these results highlight the complexity of contextual visuomotor adaptation and call for a new theoretical framework encompassing both probabilistic learning and cue efficacy for visuomotor learning.

This work was supported by grants from the French Agence Nationale de la Recherche (ACES:ANR-21-CE28-0013-02; Vision-3E:ANR-21-CE37-0018-01)

26.324 IN-BUILT AND LEARNT PRIORS FOR MOTION DIRECTION PERCEPTUAL DECISION-MAKING

Liubov Ardasheva[1] (liubov.ardasheva@univ-amu.fr), Anna Montagnini[1]; [1]Institut de Neurosciences de la Timone, Aix-Marseille University, CNRS

In the face of sensory uncertainty several factors can impact our perceptual choices. This study aims at exploring the influence upon human motion direction estimation, of the repeated presentation of specific visual motion features, leading to the formation of experience-based priors. We test the hypothesis that these learnt priors should influence subsequent perception and sensorimotor behavior in interaction with longer lasting, built-in beliefs. Recent studies suggest that learnt priors exert a sort of "magnetic force," drawing perception toward or pushing it away from the expected direction. The focal point of our research is to gain insight into this "repulsion" phenomenon. In our experiment, participants had to estimate the global motion direction of a Random Dot Kinematogram (RDK), with 16 possible directions of movement and with three coherence levels: 0.4 (easy), 0.15 (medium), and 0.05 (hard). Additionally, we conducted three sessions: one unbiased, fair for each angle, and two biased sessions with 60% of trials with a given direction (Right, 0°, or Up-Left, 135°) potentially inducing a non-uniform direction prior. Our results do not validate a repulsion effect of direction estimate, neither in the unbiased session nor in biased sessions. Instead, they reveal a persistent prevalence of preferences for cardinal directions in erroneous trials across all sessions (43%) and especially evident for the low coherence RDK, with a distinct choice-bias for horizontal right, with the notable exception for the session with the over-represented rightward motion, and the upper hemifield. In contrast, oculomotor anticipation displays adaptability to the experimentally-induced bias in two sessions, offering minor but noteworthy evidence supporting a dissociation between anticipatory eye-movements and perceptual decisions. Overall our results suggest that our internal beliefs carry more weight than situational biases, emphasising the importance of dissecting the different sources and the specific dynamics of internal priors in motion estimation and visuomotor tracking.

This work is partly funded by ANR21-CE37-0018-01, Vision-3E Grant

26.325 MODULATING COLOR CUE EFFECTIVENESS: THE ROLE OF ACTIVE SELECTION IN VISUOMOTOR ADAPTATION

Hamza OUELD KADDOUR EL HALLAOUI[1] (hamza.oueld-kaddour-el-hallaoui@univ-amu.fr), Vanessa CARNEIRO-MORITA[1], Emmanuel DAUCÉ[1,2], Anna MONTAGNINI[1]; [1]Institut de Neurosciences de la Timone, CNRS & Aix-Marseille Université, Marseille, [2]Centrale Méditerranée

Motor control research highlights the varying efficacy of cues in conveying behavioral context information (Howard et al., 2013). Notably, studies on visuomotor adaptation reveal the ineffectiveness of color cues in cue-dependent contextual saccadic adaptation (Azadi and Hardwood, 2014) and anticipatory smooth pursuit eye movements (aSPEM) (Carneiro-Morita et al., 2021), though the reasons remain elusive. This study explores how participants' active cue selection influences effectiveness in a visual motion tracking task. Participants tracked a colored dot (Green or Red) at the screen center, which disappeared for 300 ms, reappeared, and moved horizontally (15°/s) right or left. The probability of rightward motion was color-dependent (e.g., P(right|Green) = 0.75, P(right|Red) = 0.25). aSPEM was analyzed in two conditions: In the first condition, dot color choice was predetermined by a Bernoulli distribution Ber(0.5), no participant action needed before oculomotor tracking. In the second condition, participants actively chose colors, alternating fairly between Green and Red. Results show active cue selection significantly influences cue valence: In the first condition, no effect on anticipatory velocity ($P > 0.05$) across direction-bias blocks. In the second condition , a significant difference emerged ($P < 0.001$), with higher anticipatory eye velocity for greater motion probability in a specific direction, depending on cue color. Findings demonstrate efficient integration of color cue-conditional probability into oculomotor anticipation when actively selected. In conclusion, our study reveals the unexplored role of active selection in modulating informative color cue effectiveness for aSPEM in direction-biased contexts. Further work is required to discern whether this phenomenon reflects a general attentional mechanism or is specifically tied to motor agency in cue selection.

This work was supported by grants from the French Agence Nationale de la Recherche (ACES:ANR-21-CE28-0013-02; Vision-3E:ANR-21-CE37-0018-01)

26.326 EYE TRACKING IN EXPERTISE ASSESSMENT CASE STUDIES

Evgenii Shelepin[1] (xensey@gmail.com); [1]Neuroiconica, [2]Pavlov Institute of Physiology

This research investigates the cognitive mechanisms behind a person's analysis of source code (Python programming language) images and error detection, focusing specifically on the impact of professional visual error search skills in Python code on eye movement control. Method: The study employed eye-tracking technology using the "Neurobureau" system for psychophysiological research. Programmers, with 1-13 years of experience, were tasked to 1) explain and 2) find errors in 10 Python code stimuli, featuring syntax highlighting. The stimuli were normalized for length and complexity. The tasks were not time-bound. Results: The study discovered that with increasing professional skill, programmers develop efficient eye movement strategies, characterized by fragmenting the code into analytically significant units. More experienced programmers displayed fewer fixations, shorter scanning paths, and larger saccade amplitudes. Notably, there was an increase in the speed of saccades, especially in large searching movements, correlating with professional skill in code explanation tasks. Error detection in visual searches was found to be primarily influenced by recognizing text details, relying on the semantics and grammar of the programming language. This differs from the processing of natural scenes or texts in natural languages. Professional experience was observed to reduce the effort required in such activities. Conclusion: The study highlights how professional experience shapes eye movement strategies in source code analysis, differentiating it from natural language reading. These findings can inform the development of neuromorphic algorithms for code generation and correction; for an automated assessment of professional skill. No conflicts of interest were identified in this research.

26.327 EYE MOVEMENT DYNAMICS AS A MEASURE OF EXPERTISE IN THE VIDEO GAME EXCITEBIKE

Emily Levin[1,3] (eml179@pitt.edu), Jasmine Horton[2], J. Patrick Mayo[3]; [1]University of Pittsburgh School of Medicine, [2]Hillman Academy, [3]University of Pittsburgh

Eye movements help us navigate the world successfully. Saccades allow us to rapidly reorient our gaze to an object of interest, and smooth pursuit allows us to track a moving object. Understanding how these eye movements interact is important because saccadic and pursuit abnormalities are a common presentation in movement disorders. Although many studies have demonstrated the interplay between saccades and fixation in laboratory settings, far fewer have examined the interaction between saccades, fixation, and pursuit in naturalistic settings such as when driving. The goal of this experiment was to examine how eye movement dynamics vary with behavioral performance when participants played the Nintendo 64 racing videogame ExciteBike. We hypothesized that participants who performed better at ExciteBike, as measured by faster lap times and fewer crashes, would spend more time tracking the bike they were controlling and less time saccading to irrelevant obstacles. Therefore, we predicted that better performers would exhibit fewer saccades and more pursuit. To test this idea, we recorded participants' eye movements using Tobii Pro Glasses 3 and defined individual eye movement events (saccades, pursuit, and fixation) using an automated classification algorithm. We next correlated time around the circular track with number of saccades and pursuits. We found that there was a significant correlation between track time and number of saccades, with participants that had faster track times making fewer saccades. Additionally, there was a nonsignificant correlation between track time and pursuit events, but this trended opposite to the direction we predicted: top performers instead had fewer pursuits than poorer performers. There was also a significant negative correlation between performance and fixation events, with better performers tending to have fewer fixations. These results suggest that eye movement interactions could be used as a measure of behavioral performance in motion-rich, naturalistic settings.

26.328 INVESTIGATING THE LINKS BETWEEN SUSTAINED ATTENTION, GAZE FIXATION PATTERNS, AND FACE IDENTITY DISCRIMINATION PERFORMANCE

Jesse Michaels[1], Catherine Francoeur[1], Claudine Habak[2], Armando Bertone[1]; [1]Perceptual Neuroscience Lab (PNLab) for Autism and Development, McGill University, [2]Cognitive Neuroimaging Unit, Emirates College for Advanced Education

Background. Sustained attention ability is often only considered during cognitively demanding task completion; it is rarely considered when tasks are less demanding or primarily perceptual. Should variability in sustained attention be shown to impact performance on such tasks, researchers may benefit from considering this in future experiments. Methods. 19 adult participants (19-25 years) completed a two-alternative forced-choice face identity discrimination task with eye tracking. Face pairs with three different viewing conditions were used to manipulate access to local information cues: (i) both forward-facing (F-F), (ii) forward- vs. side-facing (F-S), and (iii), both side-facing (S-S). In each trial, one face identity was either identical to the other face or changed by either 4% or 8% toward another identity. Sustained attention was defined via a novel metric derived from the Connors CPT. Finally, performance was compared across two identical sequentially presented blocks of trials. Results. A logistic mixed model revealed no attention-related ($p=.26$) or block-related ($p=.12$) differences in identity discrimination. A mixed model demonstrated a trend for the novel attention measure to significantly predict the number of fixations made during trials; participants with lower sustained attention made fewer fixations before responding ($p=.09$). All participants further reduced their number of fixations in the second testing block ($p < .001$). Additionally, they simultaneously made more errors ($p < .001$) and made fewer fixations ($p = .01$) in the F-S condition with restricted access to local information. Conclusions. Manipulating access to local information cues affected the number of fixations made and increased error rate during task completion. While lower sustained attention ability may result in fewer fixations per trial, this behaviour did not negatively impact discrimination accuracy. Results are interpreted as more liberal response bias in participants with lower sustained attention ability. Additional analyses will examine visual areas of interest fixated during the task.

CIHR operating grant awarded to AB

26.329 THE EYES MOVE TOWARDS FEARFUL FACES HUNDREDS OF MILLISECONDS BEFORE THEY REACH AWARENESS

Junchao Hu[1] (junchao.hu@unifr.ch), Stephanie Badde[2], Petra Vetter[1]; [1]University of Fribourg, Switzerland, [2]Tufts University

Do emotional faces have prioritized access to visual awareness? In the absence of awareness, emotional faces guide oculomotor responses contingent on their emotional expression, suggesting emotion processing in the absence of awareness as well as a dissociation of eye movements and visual awareness (Vetter, Badde, Phelps & Carrasco, 2019). However, it is unclear how early the eyes see emotional faces that reach awareness. Using continuous flash suppression, we rendered fearful and neutral faces invisible from observers' awareness. The contrast of the face images slowly increased, and participants were instructed to press a key corresponding to the position of the face as soon as they started seeing something. In addition to the position of the face, participants reported its emotional expression, and the image's visibility at the time they pressed the button. Meanwhile, we tracked observer's eye movements. Our behavioral results show that fearful faces broke into awareness more often and earlier than neutral faces (in line with previous studies, e.g., Yang et al., 2007; Gray et al., 2013). Eye-tracking results show that the eyes moved several hundreds of milliseconds earlier towards suppressed fearful than towards suppressed neutral faces. Once participants' gaze was centred on the face image, manual reaction times were identical for fearful and neutral faces. When the faces were superimposed on the flashing mask, and thus not suppressed from awareness, neither manual reaction times nor oculomotor responses differed between fearful and neutral faces. These novel results show that fearful faces have prioritized access to awareness while avoiding the potential confounds of decision criteria and response processes associated with classical breakthrough paradigms. We suggest that fearful faces' advantage to guiding oculomotor responses in the absence of awareness might be the mechanism facilitating their perceptual detection.

This work was supported by a PRIMA grant (PR00P1_185918/1) from the Swiss National Science Foundation to Petra Vetter.

26.330 SPECIFIED FUNCTIONS OF THE FIRST TWO FIXATIONS IN FACE RECOGNITION: SAMPLING THE GENERAL-TO-SPECIFIC FACIAL INFORMATION.

Meng liu[1,2], Jiayu Zhan[3], Lihui Wang[1,2,4]; [1]Institute of Psychology and Behavioral Science, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200030, China, [2]Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China, [3]School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China, [4]Shanghai Center for Brain Science and Brain-Inspired Intelligence Technology, Shanghai 201602, China

Although it has been well known that the first two fixations are crucial and can even be sufficient for face recognition, the functions of the individual fixations remain to be specified. Here we show the distinct and collaborative roles of the first two fixations (Fix I and Fix II) during face perception across two different tasks. We recorded eye-movement data of 28 participants during face viewing in an identity recognition task and an emotion recognition task. Using graphical analysis of the fixation locations, we showed that Fix I was clustered along the nose bridge, whereas Fix II diverged to the key regions such as the eyes, nostril and lips. A subsequent distance analysis uncovered that, regardless of task type, the location of Fix I had comparable distances to all of the four key face ROIs (eyes, nose, mouth) to cover the broad facial information, whereas the location of Fix II was biased for specific ROIs to get local information. Moreover, the patterns of Fix II correlated more with differentiating information between faces. Meanwhile, in the emotion task where the recognition performance was suboptimal, the correct and incorrect recognition can be differentiated by the distribution patterns of Fix II, but not Fix I. Yet, the combined patterns of the two fixations yield better differentiation than Fix II. Together, the fixation sequence reflects an optimized information sampling mechanism where Fix I gets the general information to build up initial hypothesis and Fix II gets the specific information to confirm the hypothesis for face recognition.

26.331 EYE MOVEMENT MODULATES THE FACE INVERSION EFFECT IN EMOTION RECOGNITION

Angeline Yang[1] (angelineyang@berkeley.edu), Yueyuan Zheng[2], Janet Hsiao[2], Susana Chung[1]; [1]UC Berkeley, [2]University of Hong Kong

The Face Inversion Effect (FIE), characterized by a greater reduction in recognition performance for inverted versus upright faces than objects, suggests that unlike objects, faces are processed holistically – a process disrupted with inversion. Despite many studies investigating FIE in identity recognition, its effect in emotion recognition has shown mixed results. We aim to clarify the effect of FIE on the recognition of different emotions and uncover mechanisms of holistic processing by linking behavioral performance with eye movement. Participants (n=40, White, 30 females, M=21.45 years) completed an expression recognition task of anger, fear, happiness, and sadness for upright and inverted faces with eye-tracking (400 trials per participant). The same face stimuli were presented in a random order with gender, emotion, and orientation counter-balanced across blocks. As expected, participants performed worse in identifying emotions of inverted faces in general. However, while recognition of sadness and anger was worse, recognition of happiness and fear remained unaffected by inversion. Using a data-driven machine-learning-based approach, Eye Movement analysis with Hidden Markov Models (EMHMM), we discovered two representative eye movement patterns adopted by participants during the emotion recognition tasks – eyes-focused and nose-focused patterns. Consistent with literature on diagnostic face regions for identification of different emotions, for upright faces, participants' eye movement patterns were more eyes-focused for anger, fear, and sadness recognition and more nose-focused for happiness recognition. Interestingly, for inverted faces, participants' eye movement patterns were only more eyes-focused for fear recognition, while anger, happiness and sadness recognition were more nose-focused. We thus show that the face scanning patterns for different emotions were influenced by the orientation of the face (upright versus inverted), and that FIE and disruption of holistic processing in emotion recognition are modulated by adherence to scanning patterns that obtain the most diagnostic information to identify the expression.

26.332 EFFECTS OF CONTEXTUAL INFORMATION ON EYE MOVEMENTS AND RECALL PERFORMANCE IN FACE LEARNING

Martina J Speck[1], Maya J Egerton-Graham[1], Charles Collin[1], Isabelle Boutet[1]; [1]University of Ottawa

The presence of person information can make learning new faces more efficient (e.g., Schwartz & Yovel, 2018; Wiese & Schweinberger, 2015). We sought to replicate this finding and extend it to face-name associations. We also examined how visual scanning patterns change when participants learn new faces and face-name associations. Based on past literature, we predicted that there would be fewer fixations to a face as it becomes familiar, and that a higher proportion of these fixations would be to the eyes. We also predicted that person information would enhance recollection performance. Participants (N=26) took part in a learning phase, where each face was associated with a name. For half of the faces, additional person information (a hobby) was provided. Lighting conditions were varied across repeated presentations to avoid image-based recognition. Recognition and naming of the learned faces were tested immediately after the learning phase, and after 1-week and 2-week intervals. Faces were presented in a different viewing angle during recall. For scanning patterns, we focused specifically on the number of fixations on different interest areas (eyes, nose, mouth, face). Preliminary results suggest that the total number of fixations to a face is reduced as it becomes more familiar. In addition, the eyes were generally viewed more often than any other area of interest region, regardless of the number of exposures. However, unlike previous research, we did not find that the proportion of total fixations to the eyes increased with familiarity. The presence of person information did not influence on scanning patterns

during learning, nor recognition performance at testing. These findings advance our understanding of face recognition, particularly regarding how eye movements can be used as an index of familiarity, and under which conditions person information aids (or does not aid) in forming face representations and face-name associations.

This research is funded by a Natural Sciences and Engineering Research Council of Canada Discovery grant (2022-03998) to IB

SATURDAY, MAY 18, 2:45 – 6:45 PM, BANYAN BREEZEWAY

Eye Movements: Saccades

26.333 MOTION MASKING AT SACCADIC SPEED IS LARGELY INVARIANT TO MOTION AMPLITUDE

Wiebke Nörenberg[1,2,3] (wiebke.noerenberg@hu-berlin.de), Thomas Symank[1], Richard Schweitzer[1,3], Martin Rolfs[1,2,3]; [1]Humboldt-Universität zu Berlin, Berlin, Germany, [2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Germany, [3]Cluster of Excellence 'Science of Intelligence', Technische Universität Berlin, Germany

The rapid retinal image shifts imposed by saccadic eye movements routinely escape conscious perception—a phenomenon called saccadic omission. A potential cause of saccadic omission is motion masking, a notable reduction in perceived motion amplitude resulting from brief stationary presentation (for a certain masking duration) of an image before and after its motion. While a prior study (Duyck, Wexler, Castet & Collins, 2018) explored the contribution of masking duration and image content, we investigated the impact of the kinematics of the motion itself. In our simulated saccade paradigm, observers fixated the center of a gray screen. A pink-noise background, with the same average luminance, swiftly appeared, rapidly shifted across the screen, and vanished, remaining stationary for the masking duration before and after the shift. Each noise image was repetitive, ensuring it was identical before and after its motion to remove any static cues to the movement's amplitude. Masking duration varied from 0 to 320 ms. Observers adjusted an on-screen arrow to indicate perceived amplitude and direction of the image shift. We manipulated motion profile (constant vs. saccade-like), amplitude (6, 12, and 18 dva) and duration (39, 55, or 72 ms, corresponding to the expected durations of saccades of the tested amplitudes). Observers reported shorter amplitudes for saccade-like compared to constant motion profiles. Interestingly, reported motion amplitudes remained largely consistent across the wide range of motion amplitudes and durations tested, even when motion was unmasked. Although motion amplitudes were considerably underestimated at all masking durations, discrimination of motion direction remained intact. Despite unmasked motion leading to larger reported amplitudes, motion direction discrimination was poorest in this condition. This study compellingly demonstrates that motion masking systematically reduces the perceived amplitude of a movement, even for large amplitudes. Furthermore, saccadic velocity profiles enhance this effect possibly contributing to the phenomenon of saccadic omission.

European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [865715 – VIS-A-VIS]) and Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy – EXC 2002/1 "Science of Intelligence" – project number 390523135.

26.334 CHARACTERISTICS OF HEAD-EYE SACCADES IN NATURAL TASKS

Paul Jolly[1,2] (pjolly3@ur.rochester.edu), Michele A. Cox[1,2], Ashley M. Clark[1,2], Sanjana Kapisthalam[1,2], Yuanhao H. Li[1,2], Ruitao Lin[1,2], T. Scott Murdison[3], Alina Neverodska[1,2], Jie Z. Wang[1,2], Bin Yang[1,2], Zhetuo Zhao[1,2], Michele Rucci[1,2]; [1]University of Rochester, [2]Center for Visual Science, University of Rochester, NY, USA, [3]Reality Labs, USA

It is well established that humans tune their saccades according to visual demands. Recent research has shown that while centering objects of interest onto the high-acuity fovea, saccades also yield luminance transients that facilitate neural encoding (Mostofi et al, 2020) and enhance visual sensitivity (Boi et al, 2018) in a low range of spatial frequencies. These previous studies were conducted with the head of the observer strictly immobilized. Under natural viewing conditions, most saccadic gaze shifts involve coordinated movements of the eye and head, and it remains unclear whether the luminance transients resulting on the retina from these joint movements differ from those measured under head immobilization. In this study, we simultaneously measured head and eye movements using a custom apparatus that enables arcminute-level resolution. This device is composed of a motion capture system (Optitrack) integrated with the oscillating field monitor (Eibenberger et al, 2016), a magnetic-induction eye-tracker. Subjects (N=17) wore scleral eye coils and a tightly-fitting helmet while performing four real-world tasks with varying acuity demands: reading a Snellen eye-chart, threading a needle, sorting beads, and searching targets in natural scenes. As expected, saccade amplitude distributions varied considerably across tasks. However, all distributions peaked for very small saccades, ranging from approximately 10' in Snellen to 1.5 deg in sorting. On the retina, head and eye movements yielded traces highly similar to those given by saccades alone under head immobilization, so that the power spectra of the resulting luminance modulations were also similar to those previously reported for head-fixed saccades. Because of the differences in amplitude distributions, these modulations emphasized high spatial frequencies in the Snellen task and extended to lower spatial frequencies in the other tasks. These results suggest that the luminance transients from head-eye saccades contribute to task-relevant spatial representations, suggesting an additional function of head-eye coordination.

Research supported by Reality Labs and NIH grants EY018363 and P30 EY001319

26.335 THE EFFECT OF VISUAL COMPETITION ON SACCADIC BEHAVIOR IN A STOP SIGNAL TASK

Zachariah A. Weir[1] (zachariah.weir@ucdenver.edu), Ryan V. Ringer[1], Carly J. Leonard[1]; [1]University of Colorado, Denver

Visual inhibition paradigms, such as the stop-signal task, often assume a unitary model of inhibition. While useful for gathering quantitative data on disorders like ADHD and schizophrenia, this oversimplifies inhibitory processes, which previous studies have shown to be distinct between and within go/no-go, Stroop, stop-signal and stop-change tasks (Boecker 2013 et al., Raud et al. 2019, Verbruggen et al. 2004). Our study further investigates the possibility of distinct yet interacting mechanisms of inhibition during a saccadic stop-signal task by examining the effects of visual competition between a stop-signal and target. If stop signal reaction time (SSRT) measures the influences of not only "high-level" inhibition but also low-level dynamics in the oculomotor system, a visual stop signal at fixation will cause greater interference when a saccade target appears at a small eccentricity compared to a larger eccentricity due to competition. This greater interference was predicted to aid the stop-process and produce shorter SSRT values. Failed stop trials were also expected to have smaller saccade amplitudes and slower velocities than successful go trials due to this stop signal interference. To test this, two target eccentricity conditions (2.5 and 7.5 degrees) were intermixed in a stop-signal task, with a central visual stop signal matched in luminance to the target. The results did not show the predicted differences in SSRTs between eccentricity conditions; however, significant differences in the characteristics of saccades in terms of reduced amplitude and slower velocity on go versus failed-stop trials were aligned with predictions. These findings suggest that saccades remain susceptible to competition's influence, even when visual competition does not significantly alter the winner of the race between stop and go processes. These findings have important implications for how the specific neural subsystems involved in generating and inhibiting saccades interact.

R15EY035056 to CJL

26.336 OCULOMOTOR FREEZING REVEALS PERCEPTUAL PRIORITY DURING FREE-VIEWING

Nina Hanning[1,2] (hanning.nina@gmail.com), Martin Rolfs[1]; [1]Department Psychologie, Humboldt-Universität zu Berlin, [2]Department of Psychology & Center for Neural Science, New York University

Perceptual dynamics around saccadic eye movements are typically probed via local discrimination targets presented at specific locations around the time of a saccade. This approach requires the observer's explicit judgement, usually via a key press at the end of a trial, which introduces an artificial perceptual task and necessarily interrupts the natural flow of eye movements. Building on the well-established finding that the mere appearance of a stimulus rapidly and automatically inhibits saccades when it reaches awareness (termed "oculomotor freezing"; White & Rolfs, 2016), we present an approach to infer perceptual priority across the visual field without disrupting active visual orienting by a secondary perceptual task. Using their gaze, observers rapidly scanned a field composed of 19 concentrically arranged tiles (r=~14°) to find a hidden search target. Certain tile colors were associated with higher or lower target likelihood. During their exploration, we repeatedly (~1Hz, jittered) flashed brief oculomotor probes (8.3ms) and selectively manipulated their contrast and distance relative to observers' current gaze position. Flash presentation reliably induced oculomotor freezing during free gaze exploration, as reflected by a depression in saccadic frequencies ~100ms after flash presentation. Inhibition strength increased with flash contrast and decreased with flash distance from the current gaze position. Crucially, freezing was more pronounced for flashes presented at behaviorally relevant locations, such as the upcoming fixation location (saccade goal) or likely search target tiles. This indicates that implicit variations in saccade characteristics elicited by spatiotemporally specific probes directly reflect perceptual priority throughout the scene. By overcoming the need to ask participants for explicit perceptual judgements, our approach provides a versatile tool for testing continuous, uninterrupted perception during natural vision. This should prove particularly beneficial for testing populations that may struggle to learn or adhere to explicit task instructions, including children, patients, or non-human primates.

This research was supported by a Marie Skłodowska-Curie individual fellowship by the European Commission (898520).

26.337 COLOR DISTRIBUTION LEARNING MODULATES SACCADE ENDPOINTS: A STUDY OF THE GLOBAL EFFECT

Léa Entzmann[1] (leaentzmann@hi.is), Árni Kristjánsson[1], Jónatan Baldvinsson[1], Árni Gunnar Ásgeirsson[2]; [1]Icelandic Vision Lab, University of Iceland, [2]Icelandic Vision Lab, University of Akureyri

In visual search, attention is affected by the features of targets and distractors encountered in preceding trials. Furthermore, the visual system can encode surprisingly detailed information about the distribution of these features. Notably, such feature distribution learning has been demonstrated in visual search tasks where observers learn the distribution that distractors from consecutive trials are drawn from. Finding a target on test trials depends on the difference in feature space between current target features and distractor distributions on preceding trials. Response times were slowed if the target feature was drawn from the previous distractor distribution and reflected the shape of the distribution. The present study assessed the impact of feature distribution learning on saccade endpoints. Participants performed a visual search task with distractors drawn from either a Gaussian or uniform color distribution. On a subsequent test trial, they had to make a saccade to a target with a distractor presented in close spatial proximity. We expected the endpoint of the first saccade to land in between the target and the distractor, a phenomenon known as the global effect. The deviation of the saccade endpoints as a function of the difference between the current target color and the distractor distribution in preceding trials was analyzed. Our findings reveal that endpoints deviated more toward the distractor when the target color was within the previous distractor's color distribution. The precise shape of the previous distractor's color distribution (uniform or Gaussian) had no impact on the endpoint deviation. Overall, our results suggest that previous distractor characteristics affect saccade endpoints, showing how feature distribution learning impacts motor behavior and visual selection.

Supported by grant #228366-051 from the Icelandic Research Fund.

26.338 SPATIOTEMPORAL COMPETITION RESOLUTION DURING ANTI-SACCADES

Julie Ouerfelli-Ethier[1] (j.o.ethier@hotmail.com), Laure Pisella[2], Aarlenne Khan[1]; [1]University of Montreal, [2]University Claude-Bernard - Lyon 1

Anti-saccade are eye movements directed in the opposite direction of a visual target. Because the visual target and the saccade goal are

decoupled, it has been suggested that the competition between those two locations need to be resolved for the execution of a correct anti-saccade. However, it remains unclear how this competition is resolved temporally and spatially. To examine how anti-saccade metrics reflect competition across time and space, we tested 14 participants on a pro-saccade paradigm as well as three different spatial configurations of anti-saccades: 90° away across hemifields, 90° away within the same hemifield, and 180° away with different target eccentricities (4, 5.5 and 7 degrees). We measured anti-saccade metrics including error rates, amplitudes and saccade endpoints and reaction times. We observed a dynamic reduction of error rates as a function of saccade reaction times. Moreover, compared to pro-saccade endpoints, we found a spatial bias toward the visual target in the endpoints of anti-saccades across all anti-saccade paradigms. This bias was specific to the vector inversion direction of each anti-saccade paradigm. Further, the magnitude of the bias was modulated by the different eccentricities tested. Taken together, these results show evidence for competition between the visual target and saccade goal location that varies dynamically across time and space during anti-saccades.

26.339 REMAPPING IN LIP TAKES TIME

Yelda Alkan[1], Alanna Morris[1], James W. Bisley[1]; [1]University of California, Los Angeles

Despite constantly shifting our gaze, our perception remains stable. It has been proposed that receptive field (RF) remapping in the lateral intraparietal area (LIP) may play a role in maintaining spatial stability. To improve our understanding of the mechanisms underlying remapping, we investigated how the remapped response relates to saccade length. If remapping occurs by a one-step process, then we expect the remapped response to better align by saccade onset. Alternatively, if remapping occurs by a moving shift over time, then the remapped response will better align by saccade offset as longer saccades will require a greater shift, which will take more time. We recorded LIP activity in animals performing a saccade task. We confirmed neurons were in LIP using a memory guided saccade task. In most trials, animals fixated a point for 800-1050 ms and then made a visually guided saccade to a target 7, 14, or 21 degrees away. 300 ms after fixation onset, a task-irrelevant probe appeared in the post-saccadic RF and remained on throughout the trial. The saccade target appeared 500-700 ms after the probe appeared. 25% of trials were no-saccade trials, in which the animal maintained fixation and the probe appeared and remained in the neuron's RF. 20% of trials were saccade-only trials, in which the animal made a visually guided saccade, but no probe was presented. We only include LIP neurons that exhibited a remapping response that was not driven by a motor response in the saccade-only task. In single neurons and in the population, we found that the remapped responses were better aligned by saccade offset, with the onset of the remapped response occurring later for longer saccades. This result supports the idea of a moving shift of RF during remapping and should help constrain models of remapping in parietal cortex.

26.340 DIFFERENTIAL SACCADE RELATED MODULATIONS IN MARMOSET V1 ACROSS CELL LAYERS AND TYPES

Neya More[1] (nmore@u.rochester.edu), Yonatan Abrham*[1], Amy Bucklaew[1,2], Jude Mitchell[1,2]; [1]Brain and Cognitive Sciences, University of Rochester, [2]Neuroscience Graduate Program, University of Rochester*

Marmoset monkeys have high acuity foveal vision and use saccadic eye movements to rapidly sample across visual locations. Each eye movement produces a wave of neural activity in primary visual cortex that includes saccadic suppression starting from the saccade onset followed by a positive post-saccadic rebound. Across the population the amplitude of suppression and rebound varied considerably, as well as latency of these components (Parker et al., 2023). Here, we investigated how saccadic-related neural responses were related to different cell types in V1 focusing on distinctions by cortical layer (superficial, input, and deep) and by spike waveform shape (narrow and broad). We used linear silicon arrays to record well isolated single units across laminar depths in one monkey that freely viewed either natural images or a blank screen. We found that narrow and broad spiking waveforms were distributed bi-modally as found in previous studies, but this was only the case in superficial and deep layers. By contrast, the input layer showed a unimodal distribution of durations which could reflect underlying differences in its anatomy as the V1 input layer is distinguished by stellate neurons and calbindin expressing inhibitory neurons (Bourne et al, 2007). We found the saccade modulation among narrow spiking neurons in all layers showed little suppression and had earlier latency positive responses. The differences between narrow and broad spiking responses were weaker but significant in the input layer, and grew the strongest in deeper layers, where narrow spiking cells showed the earliest positive responses. These early positive responses among narrow spiking cells could play a role in mediating suppression to the other cell classes.

Funding: NIH EY030998, AB from NIH T32EY007125

26.341 SACCADE RESPONSE MODULATION IN AREAS MT/MTC ACROSS CELL TYPES AND LAYERS

Amy Bucklaew[1] (amy_bucklaew@urmc.rochester.edu), Shanna Coop[2], Jude Mitchell[2]; [1]University of Rochester, Brain and Cognitive Sciences, [2]University of Rochester, Neuroscience

To maintain a stable percept of the world around us, the information during saccades is suppressed to avoid retinal blur, termed saccadic suppression. Extra-retinal motor signals controlling saccadic suppression could originate from oculomotor areas such as area FEF and the Superior Colliculus (Sommer and Wurtz 2008), but less is known about how the signals are integrated locally. Previously we have found that area MT and MTC neurons vary considerably in their timing and amplitude of saccadic suppression as well as the latency of the post-saccadic excitatory rebound response after suppression. In particular, we found a higher proportion of neurons in MTC that exhibited an early latency excitatory response that could reflect extra-retinal motor feedback during saccades (Bucklaew et al., VSS, 2023). By recording from neurons in both areas as two monkeys freely viewed either natural images or a blank screen, we observed that area MTC showed a distinct peak in the early response interval (<50ms), which remained significant for the blank screen condition and likely reflects extra-retinal feedback signals. This early peak in response is driven by a subpopulation of cells (18%) that had strong modulations on blank screens, which are seen in both MT and MTC. Here we sought to examine in detail the distribution of these neurons across cortical cell types and layers in MT/MTC. We found no major distinctions in the laminar distribution of early peak neurons as a function of cortical layer. However, units with initial positive modulation were more likely to have narrow shaped spike waveforms while units with initial suppressive modulation were more likely to have broad spike waveforms. These patterns held true across all cortical layers, with broad spiking cells having a lower mean modulation compared to narrow spiking cells, and suggests distinct pathways for motor feedback to influence saccadic suppression.

Funding: AB, SC, and JFM from NIH EY030998, AB from NIH T32EY007125

26.342 FUNCTIONAL ARCHITECTURE OF VISUAL RESPONSES IN DORSAL AND VENTRAL BANKS OF ANTERIOR CINGULATE CORTEX

Pranavan Thirunavukkarasu[1], Steven Errington[2], Amirsaman Sajad[3], Jeffrey D. Schall[1,3]; [1]York University, [2]Washington University in St. Louis, [3]Vanderbilt University

Previous research in humans, macaques, and rodents has demonstrated a role of the medial frontal cortex in detecting errors, registering success, and exerting pro-active control on saccade production. The cortical circuitry accomplishing these computations is unknown. Here, we analyze neural spiking data from two monkeys collected using a linear electrode array to describe the functional properties of neurons across cortical layers in the dorsal and ventral banks of the caudal segments of anterior cingulate cortex (actually midcingulate cortex, MCC) during a visually-guided saccade countermanding task. Monkeys were rewarded for shifting gaze to a visual target unless, in infrequent random trials a stop signal appeared, which instructed the subject to cancel saccade initiation. Despite sampling over 900 neurons in MCC, less than 5% demonstrated significant modulation in response to a visual target. Typically, these responses were sustained, discharging until after saccade production. Around 70% of visually responsive neurons were most sensitive to a visual target appearing in one hemifield. Interestingly, as observed previously in SEF, MCC visual neurons showed an unexpected preference for ipsilateral visual stimuli. MCC visual neurons were modulated significantly later than those in occipital and temporal visual areas, as well as other frontal regions such as frontal and supplementary eye fields. Although we found no difference in visual onset latencies between the dorsal and ventral banks, task-related visual response latency varied across cortical layers. These findings provide the first report of the functional architecture of visual signals in two discrete regions of cingulate cortex and provide important constraints for microcircuit models of these areas.

26.343 STEADY-STATE VISUALLY EVOKED POTENTIALS (SSVEPS) IN THE PRESENCE OF VOLUNTARY EYE AND HEAD MOVEMENTS

Weichen Liu[1] (wel008@ucsd.edu), Chiyuan Chang[2], Russell Cohen Hoffing[3], Steven Thurman[3], Cory Stevenson[4], Tzyy-Ping Jung[4], Ying Choon Wu[4]; [1]UC San Diego, Department of Computer Science and Engineering, [2]Beth Israel Deaconess Medical Center, [3]US DEVCOM Army Research Laboratory, [4]Institute for Neural Computation

This study explored the feasibility of detecting steady state evoked potentials (SSVEPs) and endeavored to uncover the time course of SSVEP signal response under conditions of unrestricted head and eye movement in three dimensional space. Leveraging immersive, head-mounted virtual reality (VR), we recorded continuous eye and head movement data simultaneously with 64-channel electroencephalogram (EEG) during a virtual visual fixation task. Participants were instructed to fixate centrally until prompted to shift their gaze towards a flickering target appearing in their field of view at different distances towards the periphery (near target: 15 degrees, far target: 30 degrees). This gaze shift was accomplished either through a saccade alone or a self-directed head turn paired with saccading. Preprocessed EEG was epoched based upon time-locking to either stimulus, fixation or gaze (based on the gaze intersection point) onset. Canonical Correlation Analysis (CCA) was performed to identify the response frequency of each SSVEP trial using a certain time window. Aligning epochs using the fixation rather than stimulus onset led to higher classification accuracy for both near and far targets (e.g., 500ms - stimulus: 25.9%, fixation: 51.7%; 1500ms - stimulus: 54.9%, fixation: 70.3%). Further, it was found that CCA scores, which reflect the correlation between the single trial EEG and a set of reference frequencies, tended to increase in accuracy even before the onset of the gaze or fixation locking point, while the head and/or the eyes were still moving. In conclusion, we demonstrate that in realistic, unconstrained viewing conditions, SSVEP signal detection can be improved through fixation- and gaze-locking. Additionally, combined head and eye movement did not fully suppress processing of the visual signal, suggesting that the "blanking effect" reported in prior studies may not be accurately characterized as a full suppression of visual processing.

The project is supported by Army Research Laboratory (W911NF2120154)

26.344 PROBING CORRELATES OF SACCADIC SUPPRESSION IN THE PRIMATE SUPERIOR COLLICULUS AND PRIMARY VISUAL CORTEX USING SIMULATED AND REAL SACCADES

Matthias Philipp Baumann[1] (matthias-philipp.baumann@student.uni-tuebingen.de), Ziad M. Hafed[1]; [1]University of Tübingen

Visual sensitivity is strongly impaired around the time of saccades, in a phenomenon known as saccadic suppression of visual sensitivity. Recently, the origins of perceptual saccadic suppression were linked to the rapid visual flow that appears on the retina during saccadic eye movements (Idrees et al. 2020, Idrees et al. 2022). Specifically, these authors found that the visual appearance of the background (e.g. low or high spatial frequency content) that is translated on the retina can affect both the strength and duration of saccadic suppression, and that this effect already starts in the retina. Here, we investigated how visual flows created by rapid image shifts affect the superior colliculus (SC) and primary visual cortex (V1), both downstream of the retina. We recorded SC and V1 neural activity from two monkeys. The monkeys fixated while we presented a rapid image displacement of different textured backgrounds (similar to Idrees et al., 2020). At different times after texture displacement, we presented a brief probe flash (luminance pedestal) within the visual receptive fields (RF's) of recorded neurons. The monkeys also performed the same experiment with real horizontal saccades (SC RF's were away from the saccade endpoint, allowing us to place probe flashes within these RF's and only assess visual responses). Both brain areas responded to texture displacements (whether shifted by saccades or externally) and probe flashes. Moreover, probe flash responses were suppressed in both areas, depending on presentation time relative to image shifts, consistent with (Idrees et al., 2020). However, the two areas exhibited important differences: V1 probe flash responses were similar whether texture displacements were saccade-induced or external; on the other hand, SC responses were much more suppressed by external texture displacements than by saccade-induced image shifts. These results suggest that SC distinguishes between self-induced and external visual stimulation significantly better than V1.

26.345 PREDICTIVE REMAPPING IN NEURAL NETWORKS: A MODEL BASED ON COROLLARY DISCHARGE SIGNALS FOR VISUAL CONTINUITY ACROSS SACCADES

Anirvan Nandy[1,3] (anirvan.nandy@yale.edu), Xize Xu[1,2,3], Monika Jadi[1,2,3]; [1]Department of Neuroscience, Yale University, [2]Department of Psychiatry, Yale University, [3]Kavli Institute for Neuroscience at Yale

When animals are viewing a stable scene, each saccade changes object locations on the retina, yet animals perceive the scene as stable across saccades. Thus, sensory representations must be updated across saccades to align pre-saccadic and post-saccadic visual inputs. One mechanism by which the brain accomplishes this visual continuity is by the predictive remapping of visual receptive fields: in many visual and oculomotor areas, a subset of neurons predictively responds to visual stimuli that will be in their receptive field by an impending saccade even before the saccade is initiated. Corollary discharge (CD), or efference copy has been shown to originate from the intermediate layer of the superior colliculus, where neurons form an orderly map of saccade direction and amplitude, generate a downward motor command to drive saccades and provide a CD of the saccade command to other brain areas. Moreover, the neurons in the visual cortex have been shown to be tuned to saccade direction. Based on these findings, we propose a model that could perform predictive remapping in a neural network. The inputs to the model network convey information about the impending saccade's direction and amplitude, consistent with known corollary discharge signals. The network effectively integrates these inputs to shift the receptive field of each neuron, similar to those observed experimentally. Both forward remapping and convergent remapping can be performed by the network depending on the formulation of the network input. The model suggests that during forward remapping, the pairwise relationship of receptive fields between cells should be preserved.

NIH R01 EY032555, Swartz Foundation, Kavli Institute

Saturday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

Visual Memory: Working memory and neural mechanisms

26.346 REPRESENTATIONS OF IMAGINARY SCENE IN THE ALPHA BAND

Rico Stecher[1], Daniel Kaiser[1,2]; [1]Mathematical Institute, Justus-Liebig-University Gießen, [2]Center for Mind, Brain and Behavior (CMBB), Philipps-University Marburg and Justus-Liebig-University Gießen

Our conscious experience is enriched by our brain's capacity to visualize a myriad of different worlds. How does our brain create such complex mental images? Previous research suggests that our brain generates mental images of individual objects by recalling their visual contents via top-down-related alpha rhythms. Based on this notion, we investigated if more complex visual contents such as natural scenes are also encoded in the alpha band during visual imagery. In our first EEG experiment, participants imagined 16 natural scenes according to detailed three-sentence descriptions and viewed images of them in a separate task. Using multivariate decoding techniques on neural rhythms, we show that imagined scenes and their properties are represented in cortical alpha activity and these representations are partly shared with late scene perception. In a second EEG experiment, we aimed to further characterize these scene representations in the alpha band in individual participants. Here, we tested few participants extensively, having them imagine 16 natural scenes according to short prompts for a total of 10 recording sessions. We then used a latent text-to-image diffusion model to synthesize scene images using the same prompts that were shown to our participants. By comparing layer activations of a scene-trained DNN in response to these AI-generated images to neural scene representations in the alpha band, we show that the contents encoded in imagery-related alpha dynamics of individual participants can be approximated by images „dreamt up" by generative text-to-image models. Overall, our results indicate that our brain creates mental images of complex natural environments by recalling scene-related visual contents via alpha rhythms.

D.K. is supported by the Deutsche Forschungsgemeinschaft (DFG; SFB/TRR 135, project nr. 222641018), an ERC Starting Grant (ERC-2022-STG 101076057) as well as "The Adaptive Mind", funded by the Hessian Ministry of Higher Education, Science, Research and Art.

26.347 SSVEPS REVEAL DYNAMIC (RE-)ALLOCATION OF SPATIAL ATTENTION DURING MAINTENANCE AND UTILIZATION OF VISUAL WORKING MEMORY

Samson Chota[1] (samson.chota@googlemail.com), Arnaud T. Bruat[1], Stefan Van der Stigchel[1], Christoph Strauch[1]; [1]Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, the Netherlands

Visual Working Memory (VWM) allows to store goal-relevant information to guide future behaviour. Prior work suggests that VWM is spatially organized and relies on spatial attention directed towards locations at which memory items were encoded, even if location is task irrelevant. Importantly, attention often needs to be dynamically redistributed between locations, e.g., in preparation for an upcoming probe. Very little is known about how attentional resources are distributed between multiple locations during a VWM task and even less about the dynamic changes governing such attentional shifts over time. This is largely due to the inability to use behavioural outcomes to reveal fast dynamic changes within trials. We here demonstrate EEG Steady-State Visual Evoked Potentials (SSVEPs) to successfully track the dynamic allocation of spatial attention during a VWM task. Participants were presented with to-be-memorized gratings and distractors at two distinct locations, tagged with flickering discs. This allowed us to dynamically track attention allocated to memory and distractor items via their coupling with space by quantifying the amplitude and coherence of SSVEP responses in the EEG signal to flickering stimuli at the former memory and distractor locations. SSVEP responses did not differ between memory and distractor locations during early maintenance. However, shortly before probe comparison, we observed a decrease in SSVEP coherence over distractor locations indicative of a reallocation of spatial attentional resources. Reaction times were shorter when preceded by stronger decreases in SSVEP coherence at distractor locations, likely reflecting attentional shifts from the distractor to the probe or memory location.

26.348 BRAIN NETWORKS INVOLVED IN RECOGNITION MEMORY ARE RECRUITED MORE STRONGLY, AND MORE EXTENSIVELY, BY REAL OBJECTS THAN BY IMAGES OF OBJECTS.

Grant T. Fairchild[1] (grant.t.fairchild@gmail.com), Samantha C. Lee[1], Michael T. Compton[1], Sara Fabbri[1], Christian A.P. Nemeth[1], Lars Strother[1], Jacqueline C. Snow[1]; [1]University of Nevada, Reno

Studies of human visual memory typically rely on artificial stimuli presented in the form of two-dimensional (2-D) images. Yet there is mounting evidence that 2-D images are processed differently in the brain compared to real-world solid objects, and behavioral studies point to a "real object advantage" in memory compared to 2-D images. Here, we used fMRI to compare brain responses to objects presented in real versus 2-D formats. During an initial learning phase in the laboratory, participants were shown a large set of everyday items presented either as real objects or as high-resolution colored 2-D images. During a subsequent recognition phase in the MRI scanner, participants viewed words that corresponded to items that had been presented during the learning phase, or had not been presented at all (i.e., "foils"). The participants' task was to judge whether each item had been presented during the learning phase (or if it was a foil), and if it had, which format it was presented in (real object or image). A univariate analysis found that cortical networks commonly implicated in recognition memory were activated more strongly for real objects than for images; no regions showed the opposite pattern. Next, a multivariate searchlight classifier analysis revealed successful decoding of recognition memory for both real objects and images versus foils, but this decoding was considerably more widespread for real objects. Moreover, an additional multivariate analysis revealed that several of these regions, including the hippocampus and parahippocampal cortex, represented the format in which the stimulus was presented during the study phase. Together, our results show that brain networks implicated in recognition memory are activated more strongly, and more extensively, by real objects than by image displays, and that areas within this network represent the format in which a previously viewed item was seen.

This work was supported by a grant to J.C.S. (NEI-NIH R01EY026701), and by the National Institute of General Medical Sciences of the National Institutes of Health P30 GM145646.

26.349 IS SENSORY VISUAL CORTEX REQUIRED FOR VISUAL WORKING MEMORY: INSIGHTS FROM META-ANALYSIS AND EXPERIMENTAL EVIDENCE

Nikos Konstantinou[1] (nikos.konstantinou@cut.ac.cy), Phivos Phylactou[1]; [1]Cyprus University of Technology

The role of the sensory visual cortex in visual working memory (VWM) remains highly controversial. We report evidence from meta-analyses and experimental work suggesting that the sensory visual cortex has critical involvement in both encoding and maintenance phases of VWM. By systematically reviewing TMS studies, we found strong evidence supporting the sensory visual cortex as a key part of the neural network for encoding and for maintaining visual information. Previous studies may have underestimated its role due to binocular stimulus presentation and complex stimuli use. Addressing these methodological issues, we conducted two TMS experiments using a simple, monocularly presented VWM orientation change-detection task. Results showed that sensory visual cortex stimulation impaired VWM performance during perceptual, early, and late maintenance stages, reinforcing the sensory recruitment hypothesis. Our findings underscore the sensory visual cortex's central role in VWM, emphasizing the shared neural substrates of perception and memory.

26.350 CORTICAL CONTROL OF WORKING MEMORY PRIORITIZATION

Hsin-Hung Li[1,2] (li.14492@osu.edu), Thomas C. Sprague[1,3], Wei Ji Ma[1], Clayton E. Curtis[1]; [1]New York University, [2]The Ohio State University, [3]University of California, Santa Barbara

Humans distribute working memory (WM) resources across items according to their behavioral relevance. Prioritized items are recalled with better precision and less uncertainty. Here, we test the hypothesis that human cortex represents the mnemonic uncertainty of items using a probabilistic neural code whose gain is modulated according to priority. Using fMRI, we scanned participants while they remembered the locations of two targets whose priorities were precued. Priority was operationalized as the probability with which the target would be the goal of a memory-guided saccade generated after a long 12 second delay. Using Bayesian decoding, we estimated the location and uncertainty of each item in WM simultaneously by modifying an existing model of neural uncertainty (van Bergen et al., 2015; Li, Sprague et al., 2021). To do so, we assumed that activity evoked by the two targets was a weighted sum of the activity to each presented alone, and the weights were gain factors based on each target's priority. Supporting our hypothesis, in visual and parietal cortex, we found that low-priority targets were associated with lower gain factors, and the high-priority targets were decoded with smaller errors and lower uncertainty. Moreover, the difference between the decoded uncertainty of the high- and low-priority targets predicted the degree to which participants prioritized the targets behaviorally. To identify the brain regions that control how WM resources are allocated, we conducted a whole-brain GLM with trial-by-trial decoded uncertainty as regressors. Remarkably, we found that neural activity in multiple areas across temporal, parietal and frontal cortex predicted decoded memory uncertainty in higher-level visual cortex. These results support a model in which activity in association cortex is the source of feedback signals that sculpt the gain of WM representations in visual cortex according to behavioral relevance.

R01 EY-027925 to C.E.C. and W.J.M.; R01 EY- 016407 and R01 EY-033925 to C.E.C; Alfred P Sloan Research Fellowship to T.C.S; Swartz Foundation Postdoctoral Fellowship to H.-H. L.

26.351 INVESTIGATING THE EFFECT OF REAL-WORLD BACKGROUND IMAGES ON SPATIAL WORKING MEMORY REPRESENTATIONS

Nicholas Schmitz[1] (nschmitz@vols.utk.edu), Sage Bendickson[1], David Sutterer[1]; [1]University of Tennessee - Knoxville

Recent work finds that the location of remembered stimuli can be reconstructed from patterns of EEG alpha-band activity throughout a 1 to 2 second memory delay (Foster et al., 2016; Sutterer et al., 2021). This foundational work involved objects being presented on a blank background. Thus, an open question is whether this is the way we use spatial working memory in the real world, where many background cues are present in addition to the target object. To answer this question, we used a delayed spatial estimation task in which participants remembered the location of a target dot presented around fixation while the background alternated block-by-block between a conventional blank background and an image of a real-world setting. We hypothesized that background cues could affect spatial memory representations in opposing ways. Background cues could enhance spatial memory representations by providing a physical cue that can be attended during the memory delay. Alternatively, background cues could serve as a placeholder that can be used in lieu of maintaining the target location throughout the delay. To test between these alternatives, we applied an Inverted Encoding Model (IEM) to the topography of EEG alpha-band power to measure the spatial selectivity for the target location for each background condition. We observed higher spatial selectivity throughout the trial for blank background compared to scene background trials. This observation is consistent with the idea that the presence of background information reduces demands on spatial working memory.

26.352 TESTING THE DURATION OF SPONTANEOUS SPATIAL REPRESENTATION IN WORKING MEMORY WHEN ITEMS CAN BE DIFFERENTIATED BY TEMPORAL-ORDER.

Sage Bendickson (sbendick@vols.utk.edu), Nicholas Schmitz, David Sutterer; [1]University of Tennessee Knoxville

Space and time are proposed to play a critical role in working memory (WM) formation and maintenance relative to other object features (Schneegans & Bays, 2017). Alongside perspectives that WM relies on space, recent work has found that stimulus position holds a prioritized status in WM (Rajsic & Wilson, 2014) and that alpha-band EEG activity spontaneously carries information about the location of a stimulus when its location is task- irrelevant (Foster et al., 2017). Recent work also finds that observers make microsaccades towards the encoded hemifield of to-be-remembered items during the maintenance of non-spatial features even when objects are tested based on ordinal position (de Vries et al., 2023), suggesting that space maintains its prioritized role even when temporal information can be used to individuate objects. However, an open question is whether neural representations of object's precise location are maintained throughout a memory delay when items can be individuated by their ordinal position. To answer this question, we used a sequential two-item delayed color estimation task in which each stimulus presentation was followed by a 1- second delay. Observers were cued to report the color of one of the two circles based on its ordinal position (first or second). This ensured that the objects could be individuated without relying on location. We applied an Inverted Encoding Model (IEM) to the topography of EEG alpha-band power to measure the spatial representation of each item in the sequence. Consistent with past work, we found a robust representation of each item's location during the delay immediately following encoding, and the spatial representation of item one did not sustain following the presentation of the second item. As such, our results indicate that spatial locations are spontaneously maintained after initial encoding but can be discarded following the presentation of new information.

26.353 USING RAPID INVISIBLE FREQUENCY TAGGING TO TRACK INTERNAL ATTENTION

Kabir Arora[1] (k.arora@uu.nl), Surya Gayet[1], J. Leon Kenemans[1], Stefan Van der Stigchel[1], Samson Chota[1]; [1]Helmholtz Institute, Utrecht University, 3584 CS, Utrecht, The Netherlands

When multiple items are held in Visual Working Memory (VWM), situational demands may require the prioritization of relevant ones. Previous studies measuring spatial attention using behavior, microsaccades, and electrophysiology suggest that locations where items were encoded play a role during their prioritization, even if their locations need not be explicitly remembered. Whether the engagement of spatial attention during prioritization also enhances sensory processing at encoding locations is currently unclear. Here we implemented a novel EEG paradigm using Rapid Invisible Frequency Tagging (RIFT) to quantify allocation of spatial attention during a VWM task. Participants encoded two items at distinct, frequency-tagged locations, one of which was cued retroactively. We quantified rhythmic brain responses to the individually tagged encoding locations over time as an index of attentional gain. Surprisingly, the RIFT response corresponding to the cued location did not increase following the retro-cue, even though we did observe a distinct cue-locked lateralization of alpha-oscillations - an established neural signature of spatial attention. In a second experiment we successfully verified that our RIFT setup was sensitive enough to measure attention directed to one of two perceived stimuli (as opposed to empty memory locations). Our results suggest that internal prioritization of memory items does not lead to increased visual processing at their encoding locations, despite the presence of other hallmarks of attention. We also show that attentional allocation can reliably be measured with RIFT using EEG (a far more inexpensive option than MEG). While our findings are compatible with previous studies, they reveal striking differences in the way that spatial attention modulates visual processing during the prioritization of memory, as compared to perceptual representations.

26.354 SUPPORT FOR AND APPLICATION OF A MEASURE OF NEURAL EFFICIENCY IN VISUAL PROCESSING

Michael Wenger[1] (michael.j.wenger@ou.edu), James Townsend[2], Sarah Newbolds[1]; [1]The University of Oklahoma, [2]Indiana University

One theme in our work characterizing the real-time characteristics of perceptual processes is the measurement of capacity, or the effects of variations in workload on processing efficiency. A core component is the hazard function of the reaction time (RT) distribution along with its integral. The former can be interpreted as an instantaneous measure of work (known as the intensity function) and the latter can be interpreted as a cumulative measure of work. We have applied this approach to a range of visual tasks, from low-level perceptual learning to high-level face perception. Here we extend that work to combine electroencephalographic (EEG) measures of brain activity and RT to form a neural efficiency score: a ratio of the hazard function of the RT distribution and the instantaneous global field power (GFP) of the EEG. This ratio measures the relationship between work accomplished and brain energy expended to perform that work. We begin by showing that the GFP can be used as a proxy for energy expended by reanalyzing data from a visual Sternberg task performed while simultaneous EEG and metabolic data were recorded.The latter allowed us to quantify metabolic energy expended (in Mj/m), and we compared this against the GFP. We show that there is a strong linear relationship between these two measures, supporting the use of the GFP in quantifying neural efficiency. We then apply this measure to data from a visual contrast detection task performed by women who were either iron deficient and non-anemic or iron sufficien, and show that neural efficiency is related to contrast thresholds and systematically differs as a function of iron status. This suggests the potential for broad use of this measure for both basic and applied questions in vision.

NIH 1 R21 ES027909; Vice President for Research and Partnerships

Saturday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

Undergraduate Just-In-Time 1

26.355 IMPACT OF COLOR PRIMING AND RETRO-CUEING ON VISUOSPATIAL WORKING MEMORY ACCURACY IS AFFECTED BY TARGET CHARACTERISTICS AND RESPONSE TASK

Danica Barron[1], Tanner Lumpkin[1], Courtney Nutt[1], Patsy Folds[1], Ralph Hale[1]; [1]University of North Georgia

Visuospatial working memory (VSWM) has a limited capacity for items that can be accurately encoded and later retrieved. This is impacted by a variety of factors including number of items, item features, item locations, and encoding time. In the present study, we explored the impact of priming on memory retrieval of cued item locations. Participants studied an array of letters arranged in a grid. In half of the trials, a color prime indicated the color of the letter that would later be cued. Half of the color primes came before the study array and half came after (retro-prime). A previous letter location was then cued with an outlined square that remained on screen until a participant response. Prime versus retro-prime trials were completed in blocks and counterbalanced between participants to avoid order effects. Prime and no-prime trials were randomized in both blocks, as were prime colors, letter colors, and order of letters within each array. Lumpkin et al. (2023) found priming to be more effective than retro-cueing using this paradigm. Here, we present findings from five manipulations of this original design: matched perceptual luminance of three hues (red, green, blue); varied physical luminance of a single hue (red); collect color responses instead of letters; reduce congruency of target colors (50% congruent, 50% incongruent); and collect both color and letter responses. As with the Lumpkin et al. (2023) study, results found priming to be effective compared to no prime or retro-cue. However, repeated-measures ANOVAs for each of our five manipulations found accuracy for our priming, retro-cueing, and no-cue conditions were affected differentially based on target characteristics (e.g., matched perceptual luminance, varied physical luminance), congruency of cue (primed and retro-cued), and response task. This study improves our understanding of VSTM and the impact of color priming on this memory system.

26.356 ARTIFICIAL FIXATION POINTS IN READING: DO THEY WORK?

Ali Bhimani[1] (alin.bhimani@gmail.com), A. Caglar Tas[1]; [1]The University of Tennessee, Knoxville

In the past, speed-reading research has found a negative correlation between reading speed and reading comprehension. Methods such as rapid serial visual presentation have attempted to counter this negative correlation to no avail. However, there has been little research testing the effect of artificial fixation points (AFPs) on reading speed and comprehension. AFPs are implemented by altering the characteristics of the letters (e.g., changing the emphasis with bolded letters or changing the color) to cue participants where they should fixate in a given word. In the present study we tested whether AFPs may improve reading speed and text comprehension. We hypothesized that both reading comprehension and reading speed would increase with AFPs compared to no AFPs. To test this, we presented participants with six texts that varied in difficulty (two easy, two moderate, and two difficult). In Experiment 1, participants were asked to read the texts with no emphasis on speed. Text comprehension was measured with five questions at the end of each text. We used eye data collected in Experiment 1 to determine the locations of AFPs in Experiment 2. In Experiment 2, participants were presented with the same six texts, but half of them were altered with AFPs. AFPs were created by changing the font color of the first half of selected words to red. Preliminary results (N=10) showed that text difficulty significantly affected both reading speed and comprehension with difficult texts having slower reading speed and worse comprehension than moderate which was also significantly worse than easy texts (all ps<.014). Further, texts that contained AFPs led to significantly faster reading speed compared to texts with no AFPs (p<.001). However, text altering did not affect comprehension. These results suggest that text altering can benefit reading speed without negatively affecting reading comprehension, contrary to previous findings with speed reading.

26.357 COMMON PERCEPTUAL FEATURES DRIVE BRAILLE LETTER RECOGNITION ACROSS MODALITIES AND LEVELS OF EXPERTISE

Hannah Yun[1] (hannahyun1002@gmail.com), Brenda Rapp[1], Robert Wiley[2], Marina Bedny[1]; [1]Johns Hopkins University, [2]University of North Carolina at Greensboro

Reading is a key example of perceptual and cognitive expertise. Most research on reading focuses on visual print, however reading is also done via touch. Individuals who are blind use braille to read, where each letter or letter combination is represented by a 2x3 matrix of six possible raised dots (e.g., ⠁ for A). We examined what perceptual features are used to recognize braille and their dependence on the level of expertise (novice vs. proficient) and modality (visual vs. tactile). Congenitally blind braille expert readers and sighted novices were presented with braille letter pairs and indicated if they were the same or different by pressing a button. Reaction time and accuracy were measured. Blind experts completed the task by touch (N = 21), while two groups of sighted novices performed it by touch with blindfolds (N = 18) or visually with printed braille letters (N = 20). Blind (acc= 92 ± 5%, d'= 3.28, RT= 1165 ± 279 ms) and sighted participants visually (acc= 98 ± 1%, d'= 4.25, RT= 1443 ± 212 ms) were highly accurate. After 45 minutes of practice with feedback, and 2 hours total on the tactile task, sighted participants achieved an accuracy of 69 ± 10% (d'= 1.02, RT= 2009 ± 251 ms). In mixed-effect regression analysis, performance was influenced by similar perceptual properties across modalities and levels of expertise. Specifically, participants were slower at judging letter pairs with more raised pins, a higher number of possible rotations/reflections (in same trials), and if pair members were rotations/reflections of each other (in different trials). Additionally, participants were slower and less accurate if pair members shared a greater proportion of pins. Perceptual features influencing braille-letter discrimination are invariant across levels of expertise and modality. Visual and tactile systems might encode the representations of braille patterns analogously.

26.358 DRAWINGS REVEAL NO BENEFIT OF SLEEP ON MEMORY

Samuel R. Rosenthal[1] (srosenthal1@uchicago.edu), Emma Megla[1], Wilma A. Bainbridge[1,2]; [1]Department of Psychology, University of Chicago, Chicago, IL USA, [2]Neuroscience Institute, University of Chicago, Chicago, IL USA

It is widely accepted that the activities performed between encoding and recall impact the quality of memory, with sleep better preserving memory (Ellenbogen et al., 2006). However, recent studies have challenged this idea, finding a lack of evidence for a sleep effect (Talamini et al., 2022). Additionally, measuring memory content through drawing has been shown to characterize memories beyond other behavioral methods (Bainbridge et al., 2021). In line with classical work, we hypothesized that sleeping for the duration of a delay period would result in greater recall accuracy in drawings than remaining awake. To test this, we recruited participants (N=190) on Prolific to encode four scene images before drawing them from memory after a 10- hour delay, with participants either awake or asleep over this interval. Additionally, to obtain a baseline of memory without a delay, each participant drew four additional images from memory immediately after seeing them. All drawings were then scored online by separate participants to determine the specific image recalled (N=311) and the number of objects in each drawing (N = 393). We found that participants recalled significantly fewer images after a delay than immediately after encoding, and that recalled images contained fewer objects. However, we found that actions taken between encoding and recall had no significant effect on the quality of memory, with no difference in the number of images remembered or their level of detail between sleep and wake conditions. We found Bayesian evidence for the null hypothesis, suggesting that a lack of difference between sleep and wake was not due to a lack of evidence, but instead a null effect. These results introduce new questions about the commonly accepted belief that sleep benefits memory and suggests that sleep may have minimal impact on the details we recall from a memory.

26.359 EXPLORING THE EFFECTS OF DELAYED VISUAL FEEDBACK ON DYNAMIC POSTURAL CONTROL

Nora Pourhashemi[1], Kayton Jaksic[1,2], Behrang Keshavarz[3,4], Taylor W. Cleworth[1,2]; [1]York University, [2]Centre for Vision Research, [3]Toronto Metropolitan University, [4]KITE-Toronto Rehabilitation Institute

BACKGROUND: Vision provides essential sensory feedback for upright stance. Delayed visual feedback can affect postural control by creating a mismatch between one's motor actions and visual responses. There is limited work examining the relationship between visual delay, postural responses, and visually induced motion sickness (VIMS) during dynamic balance tasks. This study examined the effects of delayed visual feedback on dynamic postural control and VIMS among healthy individuals. METHODS: Twenty young adults stood on a force plate mounted to a motorized platform that continuously translated in the anteroposterior (AP) direction for 60s while wearing a virtual reality head-mounted display (HTC Vive). Full body kinematics using motion capture (Vicon) and center of pressure (COP) displacements were collected. Surface electromyography (EMG) was recorded from the right medial gastrocnemius (MGast), tibialis anterior (TA) and soleus (Sol) muscles. Trials for each delay condition (0ms, 250ms, 500ms) were randomized, and repeated once. Following each trial, questionnaires were completed to assess perceived stability and VIMS. AP and mediolateral (ML) root mean square (RMS), and sample entropy (SE) were used to quantify postural responses. Mean activity and co-contraction between antagonistic muscles were calculated from EMG data. RESULTS: As the amount of delay increased, AP and ML COP RMS, AP COP SE, as well as AP Head and Trunk RMS significantly increased. TA and MGast activity and TA/SOL co-contraction were significantly greater with increased visual delay. Perceived stability significantly decreased, while VIMS remained unaffected with increased delayed visual feedback. During repeated exposure, perceived stability increased, while postural responses decreased. CONCLUSIONS: Overall, increasing the delay between head motion and visual feedback during a dynamic balance task resulted in larger postural responses but did not affect VIMS. However, upon repeated exposure, postural responses decreased. Therefore, participants demonstrated adaptability, integral for developing interventions aimed to reduce fall risk and balance deficits.

Funded by Natural Sciences and Engineering Research Council of Canada (NSERC) and Vision: Science to Applications (VISTA) at York University.

26.360 EXPLORING THE INFLUENCE OF OBJECT AFFORDANCES AND PROXIMITY ON TOP-DOWN VISUAL PROCESSING IN THE DORSAL AND VENTRAL STREAMS

Nisa Hofer[1], Amira Jensen[1], Jenni M. Karl[1]; [1]Thompson Rivers University

Vision research often focuses on bottom-up models, while top-down models receive less attention. We aimed to determine whether high-level knowledge of an object's action possibilities could bias low-level visual processing in favor of the dorsal vs. ventral stream. We hypothesized that objects within reach that afford action would bias processing toward other action-relevant features (e.g., orientation) in the dorsal stream; whereas objects beyond reach that don't afford action would bias processing toward action-irrelevant features (e.g., saturation) in the ventral stream. A 2 x 2 x 2 repeated measures design was used to analyze the effects of object location (within vs. beyond reach), object graspability (graspable vs. non-graspable), and change type (orientation vs. saturation) on participant Accuracy and Reaction Times in a visual change detection task. Participants sat in front of a monitor and viewed a fixation cross (1500ms) followed by 1st object presentation (2500ms), a visual mask (900ms), then 2nd object presentation (up to 3000ms). Between the 1st and 2nd presentation, the object did not change, or increased/decreased in either orientation (14°) or saturation (14%) and the participant pressed a key to indicate the change they observed. The results revealed that participants were faster and more accurate at detecting orientation compared to saturation changes, and at detecting changes in graspable compared to non-graspable objects. They were also more accurate at detecting changes when the object (i.e., monitor) was located beyond reach. The results failed to support our hypothesis, but provide valuable insight for future research. Limitations of the study design such as task difficulty, participant fatigue, and inadequate control of low-level object features are discussed; as is another possibility – that participants may have used bottom-up strategies to complete the task and thus, conscious attention to the object's affordances may be necessary to activate top-down influences on visual processing.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number RGPIN-2017-05995] and Thompson Rivers University.

26.361 IMPACT OF STATIC PHYSICAL AROUSAL ON ATTENTIONAL NETWORKS

Jenna Glotfelty[1], Ming-Ray Liao[1], Brad T. Stilwell[1], Brian A. Anderson[1]; [1]Texas A&M

Humans tend to perform better cognitively when physically aroused. For instance, studies have found that physical arousal can lead to improvements in the executive control of attention. One approach is to use the attentional network task (ANT) after inducing physical arousal. This task is useful because it combines the well-studied spatial cueing and flanker paradigms, which can be used to measure the executive control of visual attention. Dynamic physical arousal can help to alleviate the influence of distracting information in the spatial cueing and flanker paradigms. However, dynamic arousal (e.g., cycling, running, swimming) is difficult to measure in the lab and might not be generalizable (e.g., less conducive to everyday settings like the classroom or office). Whether static arousal improves the control of attention in the same ways as dynamic arousal remains an open question. To test this question, our participants gripped a hand dynamometer for 18 seconds at maximum strength to evoke physical arousal (a technique previously demonstrated to elevate physical arousal). Then, participants completed the ANT. Participants were tasked with reporting the direction of a centrally presented target arrow (left or right) via a key press. The target was flanked by distractors that were either congruent with, incongruent with, or neutral with respect to the target response. Prior to the search array, nonpredictive, salient cues were presented. There were two conditions of physical arousal: grip versus no-grip. Results show that attention was captured by the salient cues (i.e., a cue-validity effect) and the flankers were distracting (i.e., a congruency effect). However, static physical arousal did not modulate either of these effects. These results suggest that, unlike dynamic physical arousal, static physical arousal may not influence the executive control of attention.

26.362 INFORMATION RELIABILITY MODULATES EXPERIENCE-DRIVEN ATTENTION

Alenka Doyle[1], Kamilla Volkova[1], Nick Crotty[1], Nicole Massa[2], Michael Grubb[1]; [1]Trinity College, [2]Mass General Brigham

Attentional control has historically been studied within a dichotomous framework based on top-down goals and bottom-up saliency. However, it has been recently established that one's idiosyncratic history with the visual world can generate attentional biases driven by task-irrelevant and physically non-salient stimuli. In the context of learned reward associations (value-driven attentional capture, VDAC), previous work has shown that the effects of experience-driven attention hinge on reward prediction signals. Recently published work from our lab modified the VDAC training phase by employing pre-cues that provided reliable or unreliable information about the training phase target color; consequently, these pre-cues were either a reliable or an unreliable predictor of the magnitude of the upcoming reward. In the test phase, distractors rendered in colors that had been unreliably pre-cued during training slowed RTs and drew more initial fixations than distractors rendered in reliably pre-cued colors. Surprisingly, this experimental manipulation also eliminated VDAC value-dependency: the magnitude of the associated reward had no impact on RTs or eye movements. That information reliability alone modulated attentional capture in our previous study suggests that the instrumental value of information, as conceptualized in the information-seeking literature, played a critical role. Here, we sought to evaluate the hypothesis that information reliability modulates experience-driven attention more directly. We repeated our modified VDAC protocol but replaced trial-to-trial rewards with accuracy-based feedback in order to isolate the impact of information reliability on experience-driven attention. Consistent with our hypothesis, unreliably pre-cued distractors slowed RTs and drew more initial fixations than reliably pre-cued distractors in the test phase. Given that reliably and unreliably pre-cued distractors had equivalent histories as sought targets, differences between them must be due to our manipulating the reliability of information in the training phase. These results converge with recent work from multiple labs investigating the role of uncertainty in experience-driven attention.

Supported by NSF-2141860 CAREER Award to MAG

26.363 IS THE DEPTH CUE OF FAMILIAR SIZE COMPUTED USING THE BIOLOGICAL EQUIVALENT OF A TRIGONOMETRIC TABLE?

Jhera Darefsky[1] (jdarefsk@asu.edu), Amoolya Chengalasetty[1]; [1]Arizona State University

Familiar size is a pictorial depth cue in which the distance to an object could be computed from an object's known physical size divided by the tangent of its visual angle. If the visual system uses this type of computation, distances, and visual angles that observers have not experienced would be as informative as those experienced. A second approach, not involving trigonometry, is one in which connections are formed between the distance of an object and the visual angle it subtends at a particular distance. Distance is recovered using a look-up table. This study aimed to explore these two explanations by varying the distance and visual angles subtended by an object during the familiarization phase. One hundred thirty-six college students estimated the distance of an identical object that varied in height from 7 to 14 cm. Using a between-subjects design, eliminating relative size, the distance to the display during training was varied. Participants were asked to study an unfamiliar object that was 10 cm tall. The test object was then placed 150 cm from the observer and viewed by one eye. The apparent distance was measured by having observers position a rod, visible to both eyes, to match the distance. Regardless of whether they had previously experienced the test display at its distance and visual angle, a depth illusion was evident. The results suggest that familiar size can be generalized to previously unseen distances, but the process may not necessarily require knowledge of the geometry of triangles. What happens may be a generalization gradient around a stored look-up value. We will have to present differences between experience and testing using more considerable distances and shifts in visual angle to see if this pattern of results occurs in more demanding conditions. Unfortunately, we are testing a null hypothesis.

26.364 MAKING SENSE OF RANDOMNESS: INVESTIGATING PERCEIVED EVENT BOUNDARIES WITHIN SCRAMBLED PICTURE STORIES

Lindsey K. Wilson[1] (lindseykkwilson2001@gmail.com), Prasanth Chandran[1], Maverick E. Smith[2], Karissa B. Payne[1], Lester C. Loschky[1]; [1]Kansas State University, [2]Washington University in St. Louis

How do people perceive discrete events, with boundaries between them, from continuous experience? To measure that, we use picture stories as a highly simplified version of real-world events, and ask participants to segment them into discrete events. To investigate story event segmentation, we need a comparison condition in which segmentation is difficult. Here, we scramble the order of images. But are scrambled-order picture stories completely un-segmentable? Randomization should disrupt viewers' segmentation, but it might be above-chance. Our research question is, what is the nature of such above-chance segmentation of scrambled picture stories? We hypothesized viewers may perceive meaning in scrambled stories based on key images (images preceding or following event boundaries in the unscrambled stories), irrespective of their order. To test that hypothesis, we scrambled the order of six picture stories, had participants view each, then segment the boundaries between events, and then summarize each story. If segmentation agreement is greatest for key images, it would indicate that event segmentation is most influenced by them. Based on a power analysis, we plan to run 20 participants each, for 24 scrambled versions of each story (N = 480 total). Previously, when we presented these picture stories to participants in a coherent order (n=48), there were 38 images across all six stories with >50% boundary agreement. Our incomplete initial results (N = 11) for the randomized condition show a total of nine images with >50% agreement. This reduction from 38 to 9 boundaries (>50% agreement) is consistent with disrupted segmentation in the randomized condition. Three of these nine total event boundaries aligned with the same images from the coherent condition. Those images could be considered key images, irrespective of their order. Nevertheless, we need many more participants to determine the actually agreed-upon boundaries, and whether key images are indeed important for that.

This work was supported by the Kansas State College of Arts and Sciences Undergraduate Research & Discovery Scholarship

26.365 REPRESENTATION-SPECIFIC AND GENERAL COMPONENTS OF THE TASK-EVOKED PUPILLARY RESPONSE IN VISUAL WORKING MEMORY

William Kemball-Cook[1], Sean R. O'Bryan[1], Joo-Hyun Song[1]; [1]Brown University

Pupillometry is a well-established tool for studying arousal and cognitive effort. Research shows that task-evoked pupillary response amplitudes increase with visual working memory (WM) load. However, it is currently unclear whether pupil diameter (PD) is also sensitive to the type of information stored in visual WM (e.g., spatial or feature-based), which may depend on different underlying neural substrates. Here, we directly compared task-evoked PD between spatial and feature-based WM tasks using a within-subjects design to understand if responses differ between the two task types, consistent with a representation-specific view of WM. We analyzed four different PD metrics: baseline, amplitude (relative to baseline), dilation velocity, and time-to-peak. Subjects (N = 74) indicated whether a probe matched either a color or spatial location presented in the encoding display while their pupil diameter (PD) was continuously recorded. WM load was manipulated by varying the number of items in the encoding display from 3-7. Critically, the tasks were visually identical and differed solely in their instructions. Our results revealed that dilation velocity and baseline were sensitive to task type, with consistently larger responses in the feature-based task regardless of WM load. Conversely, amplitude and time-to-peak were sensitive to WM load irrespective of task type, showing larger responses with increasing set sizes. Interestingly, time-to-peak demonstrated an ability to track load demands beyond the typical 4-5 item WM capacity plateau observed with amplitude. Overall, these findings suggest that different components of the task-evoked pupillary response can be used to index representation-specific and representation-general WM processes. Moreover, these results underscore the value of analyzing the pupillary response via separable components rather than a unitary measure (e.g., baseline-corrected amplitude), providing a deeper understanding of cognitive processes involved in working memory tasks and other tasks which require cognitive control.

NSF BCS-1555006

26.366 RHYTHMIC ATTENTIONAL SAMPLING IN VISUAL PERCEPTION AND VISUAL WORKING MEMORY

Khayla Santiago[1] (khaylasan@gmail.com), Chunyue Teng[1]; [1]Lawrence University

Successful goal-directed behavior requires the delicate balance between managing task-relevant information stored in working memory and the continuous processing of incoming sensory input. It remains unclear whether we can concurrently attend to both external sensory inputs and internal representations, or if the focus of attention oscillates between the two sources of information. To test this, we assessed fluctuations in performance during a task that requires simultaneous engagement of working memory and perceptual processing. Participants were instructed to maintain a specific orientation in mind while also observing another orientation patch on the screen. After a variable inter-stimulus-interval (ISI), they were prompted to compare a test probe against either the memorized orientation (memory comparison condition) or the visually monitored orientation (perceptual comparison). Critically, we manipulated the duration of the ISI: 500-1500 ms with a 20 ms step, resulting in a total of 50 ISIs. Response time (RT) and accuracy were analyzed separately for each task. Visual inspection of the time courses revealed notable fluctuations in both RT and accuracy. We performed Fast Fourier transform of the data to extract spectral power and phase angle across different frequencies. Statistical significance was examined through a non-parametric resampling procedure by permutating the RTs and hit/misses across all ISIs. The analysis identified increased power within the theta and low-alpha frequencies (3, 6, and 8 Hz) for both the perceptual and memory tasks. Importantly, the two representations fluctuated at different phase angles at those identified frequencies, indicating a distinct rhythmic alternation in attentional sampling between internal and external visual representations. Together, these results demonstrate the rhythmic nature of attentional shifts between internal and external visual representations, and further highlight the functional relevance of neural oscillations in supporting concurrent visual perception and working memory functions.

26.367 TACTILE STIMULI ARE MIRRORED IN ACCORD WITH EXTERNAL VANTAGE POINTS INDUCED BY VIRTUAL REALITY

Anwesha Das[1] (anwesha.das@caltech.edu), Daw-An Wu[1], Shinsuke Shimojo[1], Matilda Cederblad[1]; [1]California Institute of Technology (Caltech)

Virtual Reality (VR) can give us visual input from a new vantage point. To what extent does one adopt such a vantage point in sensory and motor processing? We examine this with a conditional mirroring effect found in tactile perception, which behaves as if the observer “views” the trace from an egocentric internal visual vantage point. E.g. if I trace the letter ‘d’ on your palm with your hand facing forwards, you perceive ‘b.’ Participants (N=6) wore a VR headset receiving video from a webcam facing the participant. They started the experiment with their eyes closed and reported perceived letters [b,d,p,q] traced on their palm, which could face towards or away from them (tactile baseline). Next, they opened their eyes, and letters were traced over the far side of the hand without touching it (VR-vision-only control). Finally, they kept their eyes open as letters were traced onto the hand (test condition). Following this, there was a short training period where they interacted with real-world objects while viewing the VR perspective. Then all previous tests were repeated in reverse order. In the initial eyes-closed baseline, they felt the tracings from an egocentric vantage point. They tended to mirror letters when the palm was facing away (85±5.7%), but not when it was facing towards them (2.1±2.3%). When they viewed the VR (no tactile input), the external visual perspective dominated. The mirroring effect disappeared on the outward-facing palm (0±0%, $p<<.01$, $t=16$), and started to appear on the inward-facing palm, even though it was not visible to the camera (19±6.1%, $p=.01$, $t=3.7$). Training did not appreciably increase this, thus far (24±5.6%, $p=.51$, $t=.71$). This indicates that external vantage points in virtual reality environments lead to re-mirroring of tactile mappings in accordance with the new visual perspective.

26.368 TESTING WHETHER INDIVIDUAL DIMENSIONS OF SPATIAL LOCATIONS CAN BE PRIORITIZED IN VISUOSPATIAL WORKING MEMORY

Jocelyn Tao[1] (jocelyntao@ucsb.edu), Alison Li[1], Dr. Tommy Sprague; [1]University of California, Santa Barbara

Why can some people easily remember misplaced items, while others forget as soon as they exit the room? This ability to encode past visual information of objects hinges on our visuospatial working memory (VSWM). The VSWM is resource-limited, which necessitates a strategic allocation of neural resources and item feature prioritization for adaptive behaviors. Our study explores the role of precursory cues (pre cues) in enhancing memory. In a 2-D delayed estimation task, participants (n = 20) were tasked to remember a target and report its location on a response cue (a circle for the radial dimension; a line for the tangential dimension). Before target stimulus presentation, participants were sometimes given a pre cue indicating the upcoming response cue. We hypothesized that VSWM would flexibly prioritize one spatial dimension over another, with pre cues guiding prioritization to boost their performance. For example, upon seeing a pre cue for the circle response cue, participants could focus on encoding the tangential dimension over the radial to maximize memory accuracy, whereas without it, they would equally encode both dimensions, spreading resources thin. However, our results found no significant performance difference with or without pre cues, suggesting limitations in dynamically encoding relevant dimensions. This has implications for understanding how we perform tasks that require spatial judgments, like estimating car positions on the highway, a skill which may be impacted in neurodegenerative or psychiatric disorders.

This work was supported by the Alfred P. Sloan Research Fellowship (TCS) and the UCSB Undergraduate Research & Creative Activities Grant.

26.369 THE INFLUENCE OF SACCADE PREDICTABILITY ON FEATURE BINDING AFTER AN EYE MOVEMENT

Isabel S. Jaen[1], Tzu-Yao Chiu[1], Julie D. Golomb[1]; [1]The Ohio State University

People make saccades frequently to sample visual information. Since visual input is naturally coded in retinotopic (eye-centered) coordinates, an updating/remapping process is needed to maintain visual stability across saccades. Previous studies found that spatial attention lingers at the old retinotopic location briefly after making a saccade. This retinotopic attentional trace was shown to disrupt visual feature perception, including one’s ability to bind an object’s location to its features. In this study, we investigated whether the predictability of a saccade influences post-saccadic remapping and feature binding. In an eye-tracking experiment, participants were cued at a target location and asked to report the color of the item subsequently appearing at this location, after making a saccade. At variable post-saccadic timepoints, they were presented with an array of four colored squares, appearing at the cue’s spatiotopic (target) location, its retinotopic (non-target) location, and two other non-target locations. Critically, we manipulated saccade predictability, where saccades in each block were either predictable (consistent saccade target location with fixed saccade cue onset) or unpredictable (variable spatially and temporally). Results showed improved general performance when saccades were predictable. Moreover, in the unpredictable saccade

condition, there were “swap errors” at the early post-saccadic timepoint, a pattern reported previously where participants report the retinotopic non-target color instead of the correct spatiotopic target color. However, the swapping errors disappeared when saccades were made predictable. These results suggest that systematic color misperceptions associated with the retinotopic attentional trace may be malleable to top-down expectations of the upcoming saccade, highlighting the role of predictive coding in maintaining visual stability across saccades.

This study was supported by NIH R01-EY025648 (JG).

SATURDAY AFTERNOON POSTERS IN PAVILION

Saturday, May 18, 2:45 – 6:45 pm, Pavilion

Multisensory Processing: Audiovisual behavior

26.401 ARE EFFECTS OF PERCEPTUAL (DIS)FLUENCY ON SOCIAL JUDGMENTS SPECIFIC TO VISUAL PROCESSING?

Robert Walter-Terrill[1] (robert.walter@yale.edu), Joan Danielle K. Ongchoco[1,2], Brian Scholl[1]; [1]Yale University, [2]Humboldt-Universität zu Berlin

Visual perception is often effortless, but not always: reading, for example, is faster and easier with some fonts compared to others. Such differences in visual *fluency* (the ease, or lack thereof, of information processing) can have profound effects on higher-level cognition and decision-making. When a written passage is held constant, for example, its author may nevertheless be judged as less intelligent (or the message may be judged as less truthful) when the passage is written in a disfluent (but still legible) font. Such effects are often normatively inappropriate, insofar as the ‘vehicle’ of a message (e.g. its font) has no necessary bearing on its actual content (e.g. its truth). Are such effects specific to visual processing, per se, or do they reflect more general perceptual principles? To find out, we explored analogous perceptual (dis)fluency in the auditory domain. In the modern era, the sounds of voices are often determined not only by intrinsic qualities (such as vocal anatomy), but also by extrinsic properties (such as videoconferencing microphone quality). We show that such superficial auditory properties also have surprisingly deep consequences for higher-level social judgments. Listeners heard short narrated passages (e.g. from job application essays), and then made various judgments about the speakers. Critically, the recordings were modified to simulate different microphone qualities, while carefully equating listeners’ comprehension of the words themselves. Common disfluent auditory signals (as in ‘tinny’ speech) led to lower judgments of intelligence, hireability, credibility, and romantic desirability. These effects were robust across speaker gender and accent, and occurred for both human and clearly artificial (computer-synthesized) speech. (And such effects may become more impactful as daily communication via videoconferencing becomes increasingly prevalent.) These results demonstrate that such fluency effects reflect a more abstract form of information processing that transcends visual perception.

RWT was supported by an NSF Graduate Research Fellowship.

26.402 WHAT WE DON’T SEE SHAPES WHAT WE SEE: PERIPHERAL WORD SEMANTICS GATES VISUAL AWARENESS

Shao-Min (Sean) Hung[1,2] (konaes@gmail.com), Sotaro Taniguchi[2], Akira Sarodo[2], Katsumi Watanabe[2]; [1]Waseda Institute for Advanced Study, Waseda University, [2]Faculty of Science and Engineering, Waseda University

Empirical data from vision sciences indicates the linguistic constraint on our perception, particularly showing a categorical benefit from semantically constructing our visual experience. However, in the periphery, our visual acuity decreases dramatically, and extracting semantic information through word recognition becomes inevitably difficult. The current study directly contended with this issue by examining whether peripheral word semantics can influence our vision. We leveraged a peripheral sound-induced flash illusion where the number of visual flashes is often dominated by the auditory beeps delivered. In each trial, two or three Mandarin characters were flashed briefly from left to right in the periphery with number-congruent or number-incongruent beeps. We first successfully replicated the original illusions. That is, incongruent audiovisual presentations led to auditory dominance. For example, when three characters were presented together with two beeps, observers often reported perceiving only two characters. On the other hand, an additional beep induced an illusory visual percept. Crucially, we found that when the three characters formed a word, the lack of a concurrent beep (i.e., 3 characters with 2 beeps) suppressed the awareness of an existing character to a greater extent. Intriguingly, participants’ successful recognition was not crucial. A separate experiment replicated the effect with participants who were unable to recognize the words, corroborating the implicit nature of the effect. When the conventional reading direction was disrupted by reversing the presentation order, the effect disappeared. Furthermore, we adopted Japanese, a language with both logographic (kanji) and phonetic (hiragana and katakana) writing systems, and showed that this effect was specific to the logographic system. These findings demonstrate the capacity of our visual system to extract peripheral semantic information without word recognition, which in turn regulates our visual awareness.

We thank the sub-award under the Aligning Consciousness Research with US Funding Mechanisms by Templeton World Charity Foundation (TWCF: 0495) and Waseda University Grants for Special Research Projects.

26.403 CROSS-MODAL INTERACTIONS DIFFER ACROSS SOUNDS IN THE EXTREME PERIPHERY

Matilda Cederblad[1] (amcederb@caltech.edu), Daw-An Wu[1], Shinsuke Shimojo[1]; [1]California Institute of Technology

Visual detection can be facilitated or inhibited by the simultaneous presentation of non-visual sensory signals. Adding a sound to a difficult visual detection task in the fovea or parafovea typically results in improved task performance. The effect of multisensory stimulation is less straightforward at the edge of the visual field, also known as the extreme periphery. Past research has indicated that not all auditory signals are integrated equally with visual signals in the extreme periphery. Certain sounds may have an inhibiting effect on the detection of visual targets whilst others can facilitate visual detection. Here we systematically investigate the cueing effect of different sounds on detection of extremely peripheral targets. Angular thresholds for visual detection were mapped with separate staircases for left/right peripheral flashes paired with: 1Khz (HIGH) or 300Hz (LOW) beep, White NOISE burst, or NO sound. Observers (N=8) responded "Left", "Right", or "No Target" for the visual target location. Observers were able to detect and correctly localize visual targets at significantly higher eccentricities with LOW sound pairings (M=96.66 degrees) compared to the other conditions (HIGH, M=93.24 degrees; NOISE, M=92.97 degrees; NO, M=92.93 degrees; Repeated Measures ANOVA p<.001, all post-hoc t-test showed LOW pairings p<.05, all others n.s.). The results may relate to differences in localizability for the sounds and their effects on spatial attention. These findings contribute to the understanding of sensory integration at the edge of vision and support previous indications that not all sounds play the same role in visual detection in the extreme periphery of healthy human vision.

26.404 DON'T TALK TO ME! RELEVANT SOUND DISRUPTS VISUAL SEARCH, IRRELEVANT SOUND DOES NOT

Jan Philipp Röer[1] (jan.roeer@uni-wh.de), Ian M. Thornton[2], Ava Mitra[3], Nathan Trinkl[3], Jeremy M. Wolfe[3,4]; [1]Witten/Herdecke University, [2]University of Malta, [3]Brigham and Women's Hospital, [4]Harvard Medical School

Visual search experiments are usually conducted in quiet environments to ensure that participants can fully concentrate on the task. The real world, however, is rarely as quiet as the laboratory. We are more or less constantly exposed to auditory information, some of which we choose to attend to, some of which we try our best to ignore. In three experiments, we examined the effects of background sound on visual search. We used the Multi-Item LOcalization (MILO) task, in which participants clicked through items labeled 1-8 in numerical order as quickly as possible while hearing auditory information through headphones. In Experiment 1, participants needed to engage with the auditory information. In the "listening" condition, participants listened to a news report while performing the MILO task. They were subsequently quizzed about the news. In the "counting" condition, participants counted how many times a specific number was mentioned during a sports commentary. Both conditions significantly disrupted visual search performance compared to a quiet control condition. In Experiment 2, auditory distractors were meaningless sequences of random words that had previously been shown to disrupt visual-verbal working memory. Participants were informed that any background sound was irrelevant and asked to ignore it. It appears that they were able to do so, because there was no effect of auditory distraction on visual search performance. In Experiment 3, we increased the difficulty of the search task by using a "shuffle" manipulation in which the subsequent items in a sequence were randomly repositioned after each localizing response. Even so, search performance again proved to be robust against irrelevant sound. The overall pattern of results suggests that visual search performance can be effectively shielded from auditory distraction, but only if we can choose to ignore the sound and not if we actively listen to it.

NEI EY017001

26.405 DYNAMIC SYNTHETIC FACES IMPROVE THE INTELLIGIBILITY OF NOISY SPEECH, BUT NOT AS MUCH AS REAL FACES

Yingjia Yu[1] (yingjiayu108@gmail.com), Anastasia Lado[1], Yue Zhang[2], John Magnotti[1], Michael S. Beauchamp[1]; [1]University of Pennsylvania, [2]Baylor College of Medicine

Seeing the face of a talker aids speech perception, especially for noisy speech. Advances in computer graphics have made encounters with synthetic faces more frequent, but little is known about their perceptual properties. We examined the benefit for noisy speech perception of two types of synthetic faces, one that used the facial action coding system (FACS) to simulate the musculature underlying jaw and lip movements during speech production, and one generated with a deep neural network (DNN). Audiovisual recordings of 64 single words were combined with pink noise at a signal-to-noise ratio of -12 dB. The words were presented in four formats: noisy auditory-only (An); noisy audiovisual with a real face (AnV:Real) and noisy audiovisual with a synthetic face (AnV:FACS or AnV:DNN). Sixty participants recruited from Amazon Mechanical Turk attempted to identify each word. Within participants, each word was presented in only a single format and counterbalancing across participants ensured that every word was presented in every format. Seeing the real talker's face improved the intelligibility of noisy auditory words (accuracy of 59% for AnV:Real vs. 10% for An). Synthetic faces also improved intelligibility, but by a smaller amount (accuracy of 29% for AnV:FACS and 30% for AnV:DNN vs. 10% for An). A mixed-effects model showed that real faces provided more benefit than synthetic faces (p < 10-16) but there was no difference between synthetic face types (t = 0.2, p = 0.99). The accuracy difference between real and synthetic faces was more pronounced for some speech tokens than others, and was the largest for /th/ and /f/ tokens. These data show that synthetic faces may provide a useful experimental tool for studying audiovisual integration during speech perception and suggest ways to improve the verisimilitude of synthetic faces.

The research was funded by NIH NS065395 and NS113339.

26.406 PERCEPTION OF MATERIALS IN VIRTUAL REALITY BASED ON THEIR AUDIOVISUAL PROPERTIES

Harshitha Koppisetty[1,2], Laurie M. Wilcox[1,2], Robert S. Allison[1,2]; [1]York University, [2]Center for Vision Research

Material perception requires integration of information from multiple senses. In virtual reality (VR) simulations, good agreement between sensory inputs leads to better accuracy, however, the effect of conflicting inputs is less understood. Here, we evaluated the effects of cue conflicts between auditory and visual material information in a virtual environment. To create the auditory stimuli, impact sounds were recorded in a controlled environment with a mechanized rod hitting a panel made of glass, wood, plastic, or metal. We rendered panels made of these same materials mounted on a stand, presented in a virtual room. During testing we combined the visual material textures with each of the recorded impact sounds, to create sixteen different conditions that were interleaved randomly and viewed using a VIVE Pro VR headset. On each trial the target object was presented and struck with a rod to produce an impact sound. The participants then classified the target material, and we recorded their responses and response time. To study the effect of agency, on half the trials, the participant observed an agent striking the target (agent-interaction trials), and in the remaining trials the participant struck the target themselves (self-interaction trials). Our results show that most participants classify materials based on their auditory properties. Further, there was no difference in the classification response between the agent-interaction trials and the self-interaction trials. Interestingly, in one of the sixteen conditions, we observed a potential audiovisual illusion - when observing a wooden target paired with a plastic impact sound, participants predominantly responded ‘metallic’. In sum, attention needs to be paid to incorporating auditory cues in VR, as discordant signals can distort perceived material properties.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) for grant #ALLRP 570802 – 21. Nous remercions le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) de son soutien.

26.408 USING HEARING AND VISION FOR LOCALIZATION, MOTION PERCEPTION, AND MOTION PREDICTION

Yichen Yuan[1] (y.yuan@uu.nl), Nathan Van der Stoep[1], Surya Gayet[1]; [1]Experimental Psychology, Helmholtz Institute, Utrecht University

Predicting motion in noisy environments is essential to everyday behavior, for instance when participating in traffic. Although many objects provide multisensory information, it remains unknown how humans use multisensory information to track moving objects, and how this depends on sensory interruption or interference (e.g., occlusion). In four experiments, we systematically investigated localization performance for auditory, visual, and audiovisual targets in three situations. That is, (1) locating static target objects, (2) locating moving target objects, and (3) predicting the location of target objects moving under occlusion. Performance for audiovisual targets was compared to performance predicted by Maximum Likelihood Estimation (MLE). In Experiment 1, a substantial multisensory benefit was found when participants localized static audiovisual target objects, showing near-optimal audiovisual integration. In Experiment 2, no multisensory precision benefits were found when participants localized moving audiovisual target objects. Yet, localization estimates were in line with MLE predictions. In Experiment 3A, moving targets were occluded by an audiovisual occluder at an unpredictable timepoint, and participants had to infer the final target location from target speed and occlusion duration. In this case, participants relied exclusively on the visual component of the audiovisual target, even though the auditory component demonstrably provided useful location information when presented in isolation. In contrast, when a visual-only occluder was used in Experiment 3B, participants relied primarily on the auditory component of the audiovisual target (which remained available during visual occlusion), even though the visual component demonstrably provided useful location information during occlusion when presented in isolation. In sum, observers use both hearing and vision when tracking moving objects and localizing static objects, but use only unisensory input when predicting motion under occlusion, perhaps to minimize short-term memory load. Moreover, observers can flexibly prioritize one sense over the other, in anticipation of modality-specific interference.

26.409 HOW DOES SUBJECTIVE CONFIDENCE INFLUENCE MULTISENSORY INTEGRATION?

Wei Dou[1] (wdou@ucsc.edu), Sara Soldano[1], Maxwell Volkan[1], Marcella Williams[1], Jason Samaha[1]; [1]Psychology Department, University of California, Santa Cruz

Research on multisensory perception has uncovered the fundamental computations underlying the integration of information from diverse sensory sources. Prevailing theory posits that information from diverse sensory modalities undergo a weighting process, wherein the contribution of each sensory input is determined by the uncertainty associated with its perceptual representation in the brain. Take audio-visual localization, for instance, where visual cues typically yield lower uncertainty estimates than auditory cues, visual cues are assigned greater weight than auditory cues in the final audiovisual percept. However, very little is known about the relationship between metacognition and multisensory integration. In the current study, we are investigating how subjective uncertainty (measured as a confidence judgment) in the visual modality affects how a visual stimulus is weighted in audio-visual integration. In Experiment 1, observers reported the mean position of a cloud of dos on each trial. To isolate subjective confidence, we aimed to manipulate participants’ confidence levels while keeping their performance the same by changing the spatial distributions of the dots (i.e., variability) and the distance between the dot cloud’s epicenter and central fixation (i.e., offset). Participants (N = 18) reported their location judgment (left/right) and their confidence (1-4) in the location decision using a single button press. We observed higher confidence ratings with larger dot variability even though there was no significant change in discrimination accuracy due to compensatory changes in dot offset. Thus, we are able to induce a dissociation between subjective confidence and objective performance in a visual localization task. A secondary experiment is underway to test the influence of this subjective confidence bias on audio-visual weighting in a paradigm using concurrent auditory cues. In sum, we used a novel manipulation to demonstrate a confidence-accuracy dissociation in visual localization, a critical first step to understanding how subjective confidence influences multisensory integration.

26.410 NATURAL HEADING STATISTICS OVER 42 HOURS OF NATURAL ACTIVITY: OBSERVATIONS AND IMPLICATIONS FOR BAYESIAN MODELING

Christian B. Sinnott[1] (christian.sinnott@ski.org), Paul R. MacNeilage[2]; [1]Smith-Kettlewell Eye Research Institute, [2]University of Nevada, Reno

Heading, the direction of linear self-motion in a head-centered reference frame, is estimated by the nervous system using vestibular and visual sensory cues. While accurate heading estimation is important for behaviors like locomotion (Cuturi and MacNeilage, 2013), human heading perception is biased, alternating between underestimation and overestimation of true stimulus values (Cuturi and MacNeilage, 2013; Crane, 2014). Perceptual biases can be modeled using Bayesian approaches that rely upon some representation of an organism’s previous experience, i.e., the prior distribution. The form of the prior can be constrained by measuring natural stimulus statistics, but to date there has been little work that has measured natural statistics of heading in humans. We therefore recorded head movements using a positional tracking camera

(Hausamann et al., 2021) worn by ten participants over 50 hours of unconstrained, natural activity outside the lab. We use a kinematic calibration procedure detailed in previous work assessing human head orientation (Sinnott et al., 2023) to transform data out of sensor coordinates and into a head-centered reference frame. Positional tracking methods also allow for direct estimation of self-motion and heading direction. After pre-processing, we report data from approximately 42 hours of activity. Both heading azimuth (direction in the horizontal plane) and elevation (vertical plane) appear similar across participants, with means close to 0° (straight ahead); both azimuth and elevation exhibit high variability. To investigate this further we partitioned data into low (<0.75 m/s) and high (>0.75 m/s) speed epochs when participants predominantly performed stationary and locomotor tasks, respectively. While low-speed heading retains high variability, high-speed heading has decreased variability and high-speed heading azimuth appears Gaussian. We also present an early implementation of an efficient Bayesian model (Wei and Stocker, 2017) using our empirical measures of natural heading statistics to predict previously observed heading bias.

This research was supported by the NSF under Grant Number OIA-1920896 and NIH under Grant Number P20 GM103650

Saturday, May 18, 2:45 – 6:45 pm, Pavilion

Multisensory Processing: Illusions, recognition

26.411 "MAGNETIC SAND" OR "INTERACTIVITY" ILLUSIONS

Shinsuke Shimojo[1] (sshimojo@caltech.edu), Kensuke Shimojo[2], Eiko Shimojo[1], Daw-An Wu[1]; [1]Caltech, [2]University of California San Diego

We investigate a series of new illusions elicited when moving one's hands closely over a dynamic random-dot (white-noise) display (VSS'23 Demo Night). (1) Draw a letter or a circle or any other patterns, and a trail (typically brighter/whiter, but darker for some) is vividly seen, decaying after 500 ms or so (nicknamed "iconic trace"). (2) Repeatedly open and close your fingers very close to the display as if picking up and releasing the dots (squeeze/move away and open/approach back to the display). Nearby random dots tend to be attracted to and then repulsed by the fingers in concert with your motion ("magnetic sand"). (3) Make small back-and-forth motions with your open palm facing the display. The dots near your palm appear to be "captured" in their movements (4) Make similar motions with your hand behind the display (thus not visible). Again, the corresponding dots appear to be captured. These illusory effects are robust across the range of parameters we tested at least: 10-30 Hz frame rate, from barely-detectable to the highest-possible luminance contrast. Several different mechanisms seem to be involved, such as dynamic occlusion-based contrast adaptation and de-adaptation in (1) and (2), dynamic occlusion-based biases in local motion distribution in (2), and action capture in (3) and (4). Several observations might help further investigation. First, whereas the iconic trace illusion (1) tends to be the strongest in most of the observers, there was no significant rank-order correlation in the strength of the different effects across observers (N=10), indirectly supporting the involvement of multiple mechanisms across illusions. Second, just observing other's actions yields significant effects, though this tends to be weaker than in the own-action condition. Third and most intriguingly, all these effects are perceptually interactive (while not really in the engineering sense).

Kensuke Shimojo is supported by Masason Fundation. The Chen Institute at Caltech

26.412 VISUAL AND AUDITORY STIMULUS FEATURES, AND THEIR CROSSMODAL CORRESPONDENCE, AFFECTS PERCEPTUAL SELECTION IN THE BOUNCE/STREAM ILLUSION

Subin Jeon[1] (subinjeon520@gmail.com), Chai-Youn Kim[1]; [1]Korea University

Various factors influence perceptual selection in bistable perception. For example, two identical disks that approach each other, overlap, then move apart, can be perceived as either streaming past or bouncing off each other (Metzger, 1934). Dubbed the bounce/stream illusion, it can be biased towards bouncing perception by introducing a sound near the point of visual coincidence (Sekuler et al., 1997). There is also evidence that an auditory tone can modulate perception of a static bistable image to favor the visual interpretation crossmodally corresponding to the tone (Zeljko et al., 2021). This study investigates the impact of visual (light/dark) and auditory (high/low) stimulus features, along with their crossmodal correspondence (light-high/dark-low; Melara, 1989), on perception of the bounce/stream illusion. Black or white disks were paired with either a high-pitched (1800Hz) or low-pitched (600Hz) tone. Stimulus onset asynchronies (SOAs) were set between -300ms to +300ms, where negative values indicate that the tone precedes the visual overlap of the disks. No-sound trials were also included. Binary responses of bounce/stream perception were collected from fifteen participants, and percent bounce responses for each lightness/pitch combinations were calculated for analysis. Consistent with prior studies, there was a main effect of SOA irrespective of auditory and visual stimulus features (p<.001). Of more relevance to our purpose, lightness modulated perception; participants exhibited more bounce responses when the disks were black than when they were white (p=0.008). Moreover, perception was further biased when a tone crossmodally corresponding to the disks' lightness was presented (i.e., low–black or high-white) than when an unrelated tone was presented (i.e. high-black or low-white. p=.008 and p=.001, respectively). These findings suggest that beyond the temporal relationship between audiovisual information, the individual features of auditory and visual stimuli, and the crossmodal correspondence between them, shape perceptual selection of an ambiguous motion stimulus.

NRF-2023R1A2C2007289

26.413 TILTING THE BALANCE: DO BALANCE ABILITIES PREDICT THE BODY TILT ILLUSION?

Sophia R. Baia[1] (srbaia@asu.edu), Michael K. McBeath[1]; [1]Arizona State University

Previous research reveals that participants show a consistent bias in overestimating the angle at which their body is tilted. This is called the body tilt illusion. The illusion strengthens with limited sensory information, particularly when visual information is removed, which highlights the potential protective purpose of the illusion. In this sense, a stronger illusion may help one to maintain their balance when valuable sensory information is inaccessible. Although research has examined the contributing factors of audition and vision to the body tilt illusion, the potential role of vestibular information has yet to be studied. Hence, the current study seeks to examine how balance abilities, tilting direction, and visual information may each contribute to the body tilt illusion. Twenty-nine participants completed a series of balancing tasks followed by eight body tilting trials on a human-sized 3D gyroscope and finished with a brief demographics and exercise questionnaire. As expected and consistent with past research, participants experienced a stronger body tilt illusion with their eyes closed compared to eyes open. Additionally, participants exhibited a stronger body tilt illusion when tilted sideways along the coronal plane compared to forward and backward along the sagittal plane. Participants also experienced a stronger body tilt illusion when tilted backward compared to forward. Lastly, stepwise multiple regression analyses suggest that balance is unrelated to the body tilt illusion. Hence, while balance abilities appear to be unrelated to the body tilt illusion, the results support the notion that the body tilt illusion serves a protective purpose when tilting (or falling) in a higher-risk situation. Just as falling with eyes closed is riskier than falling with eyes open, human anatomy suggests that falling sideways along the coronal plane is riskier than falling forward/backward along the sagittal plane and falling backward is riskier than falling forward.

26.414 VISUAL-HAPTIC WEIGHT ILLUSIONS ARE EXPLAINED BY EFFICIENT CODING BASED ON CORRELATED NATURAL STATISTICS

Paul Bays[1]; [1]University of Cambridge

Bayesian accounts of perception hold that observers should take into account prior experience when making sensory judgements. In our everyday experience, the sizes and weights of objects are strongly positively correlated. When objects are lifted, visual information about size can be combined with haptic feedback about weight, and a naive application of Bayes rule predicts that the perceived weight of larger objects should be exaggerated and smaller objects underestimated. Instead it is the smaller of two objects of equal weight that is perceived as heavier, a striking and counterintuitive phenomenon termed the Size-Weight Illusion (SWI). Here we provide a new normative explanation of the SWI based on principles of efficient coding, which dictate that stimulus properties should be encoded with a fidelity that depends on how frequently those properties are encountered in the natural environment. Recent work has shown how efficient coding induces seemingly "anti-Bayesian" biases in individual visual features, including planar orientation, but the theory has not previously been applied to multidimensional or cross-modal stimuli. Here we show that the precision with which human observers estimate object weight varies as a function of both mass and volume in a manner consistent with the joint distribution of those properties observed among everyday objects. We further show that participants' biases in weight perception (the SWI) are quantitatively predicted by Bayesian estimation when taking into account the gradient of discriminability induced by efficient encoding. Results from the related Material-Weight Illusion (MWI), in which objects seemingly made of less dense material are judged as heavier, can also be accounted for on these principles, with surface material mediating the expected relationship between volume and mass.

This research was supported by the Wellcome Trust.

26.415 FACILITATION OF VISUAL AND HAPTIC RECOGNITION AFTER MULTISENSORY ACTIVE CONTROL OF REAL 3D OBJECTS.

Hellen Kyler[1] (hkyler@iu.edu), Karin James[1]; [1]Indiana University

3D objects are often studied in multisensory environments and subsequently recognized through multiple senses. Although several studies have revealed faster recognition of novel objects after active control during exploration relative to passive viewing, the effects that haptic information during active study may have on recognition has not been tested in these paradigms. Here we investigated the effect of active manual control, involving multisensory object information at study, on both unisensory visual and unisensory haptic object recognition. We predicted that regardless of test and exploratory condition, multisensory active exploration of 3D objects would result in more accurate recognition relative to other groups without 3D multisensory information during study. 3D objects were studied through either 3D multisensory active exploration, unisensory passive viewing of a yoked video, multisensory study where the visual information was 2D (seeing one's own actions via a computer screen), multisensory study where the 2D visual information was a video of another participant's exploration or unisensory haptic exploration. Participants subsequently recognized objects either visually on a computer or through haptic exploration of 3D objects (without vision). Haptic test results revealed superior accuracy following multisensory active exploration relative to those who studied objects with unisensory haptics, and who viewed replayed multisensory actions through a computer screen. Multisensory active study facilitated more accurate recognition during the visual test relative to all other groups. Haptic test scores were higher than visual test scores, reflecting the efficacy of haptics during test. Overall, multisensory study of the actual 3D objects resulted in higher accuracy than multisensory study when the visual information was presented through a computer screen (in real time), reflecting the 3D over 2D dominance in the visual modality. We conclude active study of 3D objects with direct vision and haptics facilitates the most efficient recognition, both for subsequent visual and haptic testing.

This investigation was supported by the National Institution of Health, T32 Grant #HD0745.

26.416 THE ILLUSION OF A NEURAL COMMON FACTOR FOR ILLUSIONS

Ayberk Ozkirli[1,2] (ayberk.ozkirli@gmail.com), Maya A. Jastrzębowska[1,2], Aline Cretenoud[1], Bogdan Draganski[2,3], Michael H. Herzog[1]; [1]Laboratory of Psychophysics, Brain Mind Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, [2]Laboratory for Research in Neuroimaging, Department of Clinical Neuroscience, Lausanne University Hospital and University of Lausanne, Switzerland, [3]Neurology Department, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

The source of variability in the susceptibility to visual illusions has been the subject of long-standing debates. Studies have explained variability in susceptibility to size illusions with idiosyncrasies in the surface area of early visual areas or the population receptive field

(pRF) sizes in these areas. However, recent research has revealed weak between-illusion correlations, raising questions about the reproducibility or generalizability of these neural findings. Here, we tested 30 participants on a battery of 13 visual illusions. The illusion battery included three illusions that were previously shown to correlate with either V1 surface area or pRF size (the Ponzo "hallway", Ebbinghaus and Delboeuf illusions), four other size illusions, and six control illusions for which size is irrelevant (contrast, uniform texture, and perceived orientation illusions). The same participants underwent a 3T fMRI experiment that included pRF mapping. For each participant, we estimated the surface areas and pRF sizes in V1 to V4. At the behavioral level, we confirmed our previous results of weak between-illusion correlations and high test-retest reliability. Contrary to previous results, we found no significant correlations between illusion magnitude and visual surface areas or pRF size. Our results fail to support the existence of both a behavioral and a neural common factor underlying visual illusions.

We would like to thank our funding sources: Swiss National Science Foundation (NCCR Synapsy, project grant numbers 32003B_135679, 32003B_159780, 324730_192755, CRSK-3_190185, 176153), the Leenaards Foundation, Fondation ROGER DE SPOELBERCH, and the Partridge Foundation.

Saturday, May 18, 2:45 – 6:45 pm, Pavilion

Object Recognition: High-level features

26.417 TEMPORAL DYNAMICS PROVIDE NEW INSIGHTS INTO THE DIMENSIONS UNDERLYING OBJECT SPACE

Alexis Kidder[1,2] (alexis.kidder.gr@dartmouth.edu), Genevieve Quek[3], Tijl Grootswagers[3]; [1]Dartmouth College, [2]National Institutes of Mental Health, [3]The MARCS Institute for Brain, Behaviour and Development, Western Sydney University

How is object information organized in high-level visual cortex? Recently, a comprehensive model of object space in macaques was proposed, defined by the orthogonal axes of animacy and aspect ratio (Bao et al., 2020). However, when using stimuli that dissociated category, animacy, and aspect ratio in humans, no tuning of aspect ratio was observed in fMRI data (Yargholi & Op de Beeck, 2023). This difference could be a result of different stimuli, or the limited temporal resolution of fMRI. Here, we asked if and when information about aspect ratio and/or animacy is available over time. We collected whole-brain electroencephalography (EEG) data while participants (N = 20) viewed the stimulus set used by Yargholi & Op de Beeck (2023). To mask object details and increase reliance on shape information, we also presented binarized versions of the stimuli. Stimuli were presented in 5Hz streams using rapid serial visual presentation. Intact and binarized stimuli sets were shown in separate streams. Using standard multivariate decoding pipelines and representational similarity analysis, we found that both aspect ratio and animacy are represented during visual object processing. The dominant dimension was modulated by stimulus type, demonstrating that the observable dimensions of object space depend on the nature of the stimuli presented. Taken together, these findings demonstrate that the dimension of aspect ratio is represented during object processing, however earlier and more transiently than categorical dimensions, such as animacy. By focusing on underlying temporal dynamics, our results provide clear new insights into the contradicting findings reported in previous work.

26.418 THE REPRESENTATIONAL DYNAMICS OF VISUAL EXPECTATIONS IN THE BRAIN

Laurent Caplette[1] (laurent.caplette@yale.edu), Tetsu Kurumisawa[1], Helen Borges[1], Jose Cortes-Briones[1,2,3], Nicholas B. Turk-Browne[1]; [1]Yale University, [2]Veterans Affairs Connecticut Healthcare System, [3]Connecticut Mental Health Center

Visual perception is modulated by expectations resulting from prior knowledge. Despite significant progress in recent decades, the neural mechanisms underlying this phenomenon remain unclear. Notably, the features in which expectations of real-world objects are represented in the brain are largely unknown: Are expected objects represented as detailed images with both low- and high-level features or are they represented only in terms of some features? Which features play a part in the modulation of sensory processing once an object is seen? In this study, participants were shown cues followed by object images. There were 8 cues associated with 8 object images, with a 58% validity; these associations were not explicitly learned. Participants had to categorize objects as animate or inanimate while their brain activity was recorded using magnetoencephalography (MEG). We used representational similarity analysis and a convolutional neural network to assess the features in which expected and perceived objects were represented during the task. Perceived objects were first represented in low-level features on posterior sensors, and then in high-level features on anterior sensors. During that time, expected objects were represented in high-level features, on anterior sensors. Interestingly, a low-level representation of expected objects was observed during cue presentation prior to object onset (starting around 300 ms after cue onset). These results suggest that expected objects are represented both with low- and high-level features but that only high-level features play a role in the integration of expectations with sensory information. The fact that this high-level representation was only visible on anterior sensors, throughout all object processing, suggests that this integration happens in high-level brain areas. The precise loci of these phenomena will be further investigated using source-level analyses.

Funding: NSF CCF 1839308 and NSERC Postdoctoral Fellowship

26.419 CATEGORICAL OBJECT PROPERTIES OUTWEIGH LOCAL VISUAL INFORMATION IN OBJECT RECOGNITION

Elsa Scialom[1] (elsa.scialom@epfl.ch), Zehra Merchant[1], Udo A. Ernst[2], David Rotermund[2], Ben Lonnqvist[1], Michael H. Herzog[1]; [1]EPFL, Lausanne, Switzerland, [2]University of Bremen, Bremen, Germany

Certain parts of an object are more critical for recognition than others. Despite decades of research, it is not well understood what these parts are and how they interact with each other. This study investigates if and to which extent categorical object properties interact with local visual information in object recognition. Fifty participants classified fragmented objects from 12 categories, each consisting of 4 unique exemplars (48 objects in total). Exemplars differed in orientation and shape. Their outlines were presented as either curved segments or dots, providing high and low local visual information respectively.

Spatial frequencies and fragment's size were comparable. We gradually increased the number of fragments to quantify the minimum number of fragments necessary to recognize objects. A linear mixed model indicates that participants required significantly fewer curved segments than dots to recognize objects ($F(1,48) = 248.05$, $p<0.001$, partial $\eta2 = 0.12$). Additionally, the number of fragments necessary for recognition varied across different categories ($F(11,38) = 5.93$, $p<0.001$, partial $\eta2 = 0.65$). There was only a weak interaction between fragment type and object category indicating independence between these factors ($F(11,38) = 3.66$, $p<0.001$, partial $\eta2 = 0.02$). As a consequence, when examining object categories individually, differences in exemplars' recognition were remarkably stable across the two types of fragmentation ($r = 0.75$, $p < 0.01$). The effect sizes observed indicate that the minimum visual information needed to recognize object primarily depends on object category, while local visual information plays a secondary role. Thus, we argue that studying the complexity of object categories is key to determine the minimum visual information needed for object recognition. Our results will help to identify the objects' parts that are crucial for object recognition in visual prostheses where the number of electrodes is restricted.

Authors are funded by ERA-NET NEURON under grant NEURON-051. UAE and DR are also funded by the "Iris und Hartmut Jürgens Stiftung: Chance auf ein neues Leben" and the "Stiftung Bremer Wertpapierbörse (SBWB)".

26.420 QUANTIFYING THE ROLE OF PERCEIVED CURVATURE IN THE PROCESSING OF NATURAL OBJECT IMAGES

Laura M. Stoinski[1,2,3] (stoinski@cbs.mpg.de), Katya Müller[2], Martin N. Hebart[2,4]; [1]University of Leipzig, [2]Max Planck Institute for Human Cognitive & Brain Sciences, [3]International Max Planck Research School, [4]Justus-Liebig-Universität Gießen

Curvature has been suggested to play a crucial role in supporting visual object processing and functional selectivity in high-level visual regions. While for artificial stimuli there may be a clear definition of curvature, this is more challenging for natural images, and the definition of curvature can vary from the curvature of the global shape of objects cropped from background to local elements of textures, both of which may deviate from our subjective percept of the curviness of individual natural images. How can we quantify perceived curvature of natural images, and how does this perceived curvature relate to patterns of brain activity? To improve our understanding of perceived curvature, we gathered extensive curvature ratings for 1,854 objects across 27,961 natural images of the THINGS database, compared their alignment with fMRI responses to computed curvature measures (Li & Bonner, 2020; Walther & Shen, 2014), and developed a neural network model that predicted perceived curvature for new images. Perceived curvature exhibited high reliability ($r = 0.93$). Computed curvature only weakly correlated with perceived curvature ($r = 0.27$ and $r = 0.30$) but also weakly correlated with each other ($r = 0.22$). In the human visual system, perceived curvature generally accounted for more variance across higher-level visual cortex than other measures and corresponded best to known category selectivities (e.g., Li & Bonner, 2020; Long et al., 2018). Given the validity of this curvature measure, we aimed at providing an automated quantification of perceived curvature for novel images. To this end, we finetuned a convolutional neural network to predict the perceived curvature of images, achieving notable performance (cross-validated $R2= 64\%$). Together, our results highlight the importance of perceived curvature as a mid-level summary statistic and provide an approach for the automated quantification of perceived curvature in natural object images.

This work was supported by a Max Planck Research Group Grant (M.TN.A.NEPF0009) awarded to MNH and the ERC Starting Grant COREDIM (ERC-StG-2021-101039712) awarded to MNH.

26.421 THE CONTRIBUTION OF FEATURES, SHAPE, AND SEMANTICS TO OBJECT SIMILARITY

Brent Pitchford[1,2] (bpitchford@hi.is), Inga María Ólafsdóttir[3,2], Marelle Maeekalle[1,2], Heida Maria Sigurdardottir[1,2]; [1]University of Iceland, [2]Icelandic Vision Lab, [3]Reykjavik University

Object similarity may not be an abstract construct that can be defined outside of the operational definition of task context. We asked people to assess the similarity of objects by rating their semantic relatedness, overall shape, and internal features. Shape similarity was assessed by rating object silhouettes with no internal features. Featural similarity was assessed by rating grayscale objects where global shape was distorted. Object pairs were either different at the basic level (e.g., hairbrush, pipe) or at the subordinate level (e.g., two different bowties). Semantic similarity of objects differing at the basic level was measured by rating similarity in meaning of word pairs. We then assessed to which degree semantics, shape, and features predicted a) explicit judgments of visual similarity of objects, b) implicit measures of object similarity as assessed by object foraging, and c) similarity in an object space derived from activations of a deep layer of a convolutional neural network trained on object classification. Explicit judgments of visual similarity were predicted both by features and shapes, but not semantics. Unlike explicit judgments, implicit object similarity depended on whether people searched for target objects among distractors of the same or different category. If targets and distractors differed at the basic level, both shape and semantic similarity predicted unique variability in foraging not accounted for by features. If objects belonged to the same category, featural similarity predicted unique variability not accounted for by shape. Contrary to previous suggestions that neural networks are primarily feature-based, shape uniquely explained variability in object space distance not accounted for by features in cases where objects differed at the basic level. Different information therefore contributes to people's explicit vs. implicit judgments of object qualities – and can also be distinguished from measures of similarity extracted from artificial neural networks trained on object classification.

This work was supported by The Icelandic Research Fund (Grants No. 228916 and 218092) and the University of Iceland Research Fund

26.422 A RADIAL TEMPLATE SPACE REVEALS ORGANIZATION OF FEATURE AND TASK-SELECTIVE REGIONS IN LATERAL OCCIPITOTEMPORAL CORTEX

Mark D. Lescroart[1], Matthew W. Shinkle[1], Arnab Biswas[1], Yu Zhao[1]; [1]University of Nevada, Reno

Human lateral occipitotemporal cortex (LOTC) represents information about objects and object properties, motion, and bodies, and is also modulated by tasks, including mental imagery, working memory, and attention to different features. However, the organization of this region is still debated, partly because variability in anatomy blurs cross-subject functional maps. We address this by developing a flexible template space for LOTC to average responses across participants. Following work that has suggested organizing principles centered on motion-selective region hMT+, we use hMT+ to define the center of a radial grid. We fix the polar angles of the grid to evenly sample space between anatomically and functionally defined regions including the superior temporal sulcus, the Occipital Place Area, and the Fusiform Face Area. We define the radial dimension by distance along the cortical surface from the hMT+ centroid. We then use data from multiple experiments to investigate the organization of feature and task representations in LOTC. In previous experiments, we developed fMRI encoding models based on motion energy, object boundary contours, and body parts and their locations. We project the prediction accuracy of these models into the radial grid space and average across subjects. The models each predict unique variance (relative to each other) in fMRI responses to rendered and naturalistic stimuli in different locations around hMT+: the boundary contour model predicts responses posterior to hMT+, the body part model superior and anterior, and the motion energy model in the center of hMT+. Separate analyses of experiments in which participants performed tasks while watching stimuli reveal that object-focused tasks (searching for objects and a one-back task comparing object parts) activate regions ventral and anterior to hMT+, while a self-motion focused task activates a region anterior and ventral to hMT+. Together, these results describe a discrete organization of feature and task selectivity surrounding hMT+.

Supported by National Institute of General Medical Sciences of the National Institutes of Health under grant numbers P30 GM145646 and P20 GM103650

26.423 NATURALISTIC DATASET AUGMENTATION AND SELF-SUPERVISED LEARNING LEAD TO MORE HUMAN-LIKE RECOGNITION OF OCCLUDED OBJECTS IN CONVOLUTIONAL NEURAL NETWORKS

David Coggan[1] (beaked.rugs-0r@icloud.com), Frank Tong[1]; [1]Vanderbilt University

Human observers can readily perceive and recognize visual objects, even when occluding stimuli obscure much of the object from view. By contrast, state-of-the-art convolutional neural networks (CNNs) perform poorly at classifying occluded objects. In previous work, we evaluated 30 humans and various CNNs on an occluded object benchmark containing nine occluder types (e.g., mud-splashes, bars, polkadots) and six visibility levels. We showed that augmenting CNN training datasets with artificial occluders led to higher classification accuracy, but a less human-like pattern of accuracy across the different occluder types. In the present study, we explored whether a human-like form of occlusion robustness could be acquired through more naturalistic modifications to the learning environment that better reflect human visual experience. First, we trained three CNNs with identical architectures to classify differently augmented ImageNet databases: unaltered (baseline), occlusion by uniformly coloured, computer-generated shapes (artificial), and occlusion by other objects extracted from photographs (natural). After training each model, we measured classification accuracy and human-likeness using the occluded object benchmark. Human-likeness was measured through both image-wise error-consistency and by correlating the profile of accuracies across conditions. Both types of occlusion training increased accuracy compared to baseline. However, natural occlusion training increased human-likeness (both measures), while artificial occlusion training showed mixed results (higher image-wise, lower condition-wise). Improvements in both accuracy and human-likeness from natural occlusion training were partially reduced when the natural occluder texture was replaced by uniform colour during training, suggesting that both the shape and texture of natural occluders play a role in human occlusion robustness. Finally, for each dataset augmentation, substituting supervised classification training with a more naturalistic, self-supervised task (contrastive learning) led to equal-or-better human-likeness. Taken together, these results indicate that occlusion-robust object recognition in humans emerges in part from unsupervised engagement with the specific forms of occlusion that occur in nature.

Supported by NIH grants R01EY029278 and R01EY035157 to FT

26.424 PROBING UNEXPLORED AREAS IN HIGH-DIMENSIONAL FMRI VOXEL SPACE USING AN ENCODING MODEL AND IMAGE SYNTHESIS

Ryuto Yashiro[1,2] (ryuto-yashiro@g.ecc.u-tokyo.ac.jp), Masataka Sawayama[1], Ayumu Yamashita[1], Kaoru Amano[1]; [1]Graduate School of Information Science and Technology, The University of Tokyo, [2]Japan Society for the Promotion of Science

Recent studies have adopted multivariate data-driven approaches like representational similarity analysis to examine the neural representation that underpins object recognition. This approach has been greatly facilitated by the recent development of large-scale datasets and neural encoding models precisely predicting neural responses to visual stimuli. The present study builds upon these advancements to elucidate neural representations underlying object recognition that have remained elusive with existing methods and datasets. We firstly applied principal component analysis on voxel patterns of publicly available fMRI data derived from thousands of visual images in the Natural Scenes Dataset (NSD; Allen et al, 2022). This analysis allowed us to capture the image distributions along each principal component and, more importantly, identify areas with sparse or no corresponding images within high-dimensional voxel response space. To understand the visual information represented in these blank areas, the latent vector of a generative model (autoencoder or BigGAN-deep) was optimized to generate an image eliciting a voxel pattern corresponding to the blank area. This was achieved by integrating the generative model with an encoding model (the feature-weighted receptive field model; St-Yves & Naselaris, 2018), which was trained to predict a voxel pattern to a visual image. We found that the autoencoder successfully synthesized visual images anticipated to elicit desired voxel patterns for each brain region. In contrast, BigGAN-deep failed to synthesize such images, likely due to the strong constraint imposed by its class embedding. Our approach enables the exploration of “unexplored” areas in the high-dimensional voxel response space, potentially leading to the discovery of novel neural representations. Additionally, image synthesis with the encoding model may offer a more feasible means of inducing a specific voxel pattern to enhance brain function or behavior, providing an advantage over conventional methods like decoded neurofeedback, where subjects voluntarily control their voxel patterns based on real-time feedback.

26.425 REVISITING SHAPE VERSUS TEXTURE BIAS IN PRIMATE VISION: CONTRASTING HUMAN VS. MONKEY PERCEPTUAL STRATEGIES

Syed Suleman Abbas Zaidi[1] (ssazaidi@mit.edu), Kohitij Kar[2]; [1]Massachusetts Institute of Technology, [2]York University

Visual perception in artificial and biological systems often diverges in the processing of complex stimuli. Geirhos et al. (2019) underscored this distinction, demonstrating that unlike convolutional neural networks (CNNs) trained on ImageNet, human vision is biased by shape over texture in object recognition. To delineate the mechanistic differences in visual perception between CNNs and biological systems, we probed the object shape versus texture biases of rhesus macaques—an animal model where finer-grained neural measurements are feasible. We trained two macaques on binary object discrimination tasks using the Microsoft COCO dataset across ten object categories. They were subsequently tested on cue-conflict images (from Geirhos et al. 2019), wherein images featured either texture-shape congruence or conflict – designed to assess whether macaques exhibit shape-bias like humans. Our results revealed a nuanced perceptual strategy in macaques. Consistent with previous studies, we observed high accuracy when the images contained no shape-texture conflicts –indicating that macaques are adept at shape-based recognition, with performances ranging from 0.80 to 0.89 across shapes. However, the introduction of conflicting textures led to variable outcomes. In particular, the accuracy for recognizing 'Elephant' shapes with 'Chair' textures dropped sharply to 0.14, highlighting a substantial influence of texture on the recognition process. The performance gradient across various shape-texture pairings suggests a complex interplay in the macaques' visual processing, differing significantly from the consistent human shape bias reported earlier. Next, we asked how these behavioral biases were driven by activity in the macaque IT cortex (critical for object recognition). We observed a significant alignment (consistency of 0.36) between neural activity and cue-conflict confusion pattern. In conclusion, our results reveal that macaques' reliance on shape versus texture is context-dependent and less robust than in humans. These insights motivate further exploration of the factors influencing distinct perceptual biases and the evolution of visual processing across species.

Google Research, CFREF, Brain Canada, SFARI

26.426 SPURIOUS RECONSTRUCTION FROM BRAIN ACTIVITY: THE THIN LINE BETWEEN RECONSTRUCTION, CLASSIFICATION, AND HALLUCINATION

Ken Shirakawa[1,2] (shirakawa.ken.38w@st.kyoto-u.ac.jp), Yoshihiro Nagano[1,2], Misato Tanaka[1,2], Shuntaro C. Aoki[1,2], Kei Majima[3], Yusuke Muraki[1], Yukiyasu Kamitani[1,2]; [1]Graduate School of Informatics, Kyoto University, [2]ATR Computational Neuroscience Laboratories, [3]National Institutes for Quantum Science and Technology

Visual image reconstruction aims to recover arbitrary stimulus/perceived images from brain activity. To achieve reconstruction over diverse images, especially with limited training data, it is crucial that the model leverages a compositional representation that spans the image space, with each feature effectively mapped to brain activity. In light of these considerations, we critically assessed recently reported photorealistic reconstructions based on text-to-image diffusion models applied to a large-scale fMRI/stimulus dataset (Natural Scene Dataset, NSD). We found a notable decrease in the reconstruction performance of these models with a different dataset (Deeprecon) specifically designed to prevent category overlaps between the training and test sets. UMAP visualization of the target features (CLIP text/semantic features) with NSD images revealed a strikingly limited diversity with only ~40 distinct semantic clusters overlapping between the training and test sets. Further, CLIP feature decoders trained on NSD highlighted significant challenges in predicting novel semantic clusters not present in the training set. Simulations also revealed the inability to predict new clusters when the training set was restricted to a small number of clusters. Clustered training samples appear to restrict the feature dimensions that could be predicted from brain activity. Conversely, by diversifying the training set to ensure a broader distribution in the feature dimensions, the decoders exhibited improved generalizability beyond the trained clusters. Nonetheless, it is important to note that text/semantic features alone are insufficient for a complete mapping to the visual space, even if they are perfectly predicted from brain activity. Building on these observations, we argue that the recent photorealistic reconstructions may predominantly be a blend of classification into trained semantic categories and the generation of convincing yet inauthentic images (hallucinations) through text-to-image diffusion. To avoid such spurious reconstructions, we offer guidelines for developing generalizable methods and conducting reliable evaluations.

This work was supported by the JSPS (KAKENHI grants JP20H05954, JP20H05705, JP21K17821 and 22KJ1801), JST (CREST grants JPMJCR18A5, and JPMJCR22P3), and NEDO (commissioned project, JPNP20006)

SATURDAY, MAY 18, 2:45 – 6:45 PM, PAVILION

Object Recognition: Visual preference

26.427 TEMPORAL DYNAMICS OF FOVEAL AND PERIPHERAL VISUAL DISCRIMINATION DURING FIXATION

Sanjana Kapisthalam[1] (skapisth@ur.rochester.edu), Martina Poletti[1]; [1]University of Rochester

The ability to simultaneously monitor stimuli at the center of gaze and across the rest of the visual field during fixation is vital for survival. Here we examine how visual discrimination for concurrently monitored foveal and peripheral stimuli changes during the course of fixation. Two Gabors (8cpd, 2 degrees diam), were presented, one at the center of the display and another 8 degrees away. Participants (N=6) were instructed to saccade to the central stimulus, and after a variable delay (0-450ms), one Gabor briefly (50ms) changed orientation. Subjects reported the direction of the orientation change and its location. To prevent visual fading the peripheral Gabor was jittered throughout the presentation time. Stimulus contrast was adjusted separately for foveal and peripheral stimuli to yield ~65% of correct responses when the orientation change occurred between 150-300 ms after saccade landing. Trials with microsaccades were discarded. Our results show that in the initial 150ms of fixation, performance was higher for foveal vs. peripheral stimuli (0.67±0.07 vs 0.50±0.02, p = 0.002). However, this foveal advantage diminished over time, and eventually performance for foveal stimuli reached chance level after 300 ms. Conversely, performance steadily improved during fixation for peripheral stimuli. 300 ms after saccade landing performance was better for peripheral stimuli (0.50±0.09 vs 0.67±0.10, p = 0.003). Here we show that visual discrimination varies greatly even during short fixation periods. Early on during fixation visual discrimination is better for foveal stimuli whereas later on this pattern flips. As in most circumstances, fixation periods are shorter than 500ms, these results may reflect the natural tendency of the visuomotor system to start preparing for the next saccade after 300 ms of fixation. This study provides critical insights into the interplay between foveal and peripheral processing during fixation and advances our understanding of the temporal dynamics of visual perception.

Meta Platforms inc.

26.428 CHARACTERIZING FREQUENCY RESPONSE FUNCTIONS OF LOW-LEVEL AND HIGH-LEVEL STIMULI IN THE HUMAN BRAIN

Na Zhao[1,2] (zhaona19@mails.ucas.ac.cn), LiNan Shi[1], Sheng He[1,2]; [1]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, [2]University of Chinese Academy of Sciences, Beijing

Brain oscillatory activities are strongly linked with cognitive functions. For example, gamma band oscillations are associated with higher-order cognitive processes. Recent reports suggest that the 40 Hz gamma band rhythm can stimulate microglial activation and facilitate the clearance of amyloid-β, thereby potentially improving cognitive function. However, such a suggestion was questioned by other studies in mice showing that sensory stimulation at this frequency does not consistently induce oscillations in brain regions such as the hippocampus. This study explores the efficacy of flicker stimulation in entraining deeper brain structures, including the hippocampus. We employed three types of stimuli (checkerboard patterns, famous faces, and landscapes) flickering at varying frequencies (20, 30, 40, and 48 Hz). Functional Magnetic Resonance Imaging (fMRI) revealed that high-level stimuli (faces and landscapes) triggered broader activation in the temporal lobe compared to low-level stimuli (checkerboard patterns). Compared with high-level stimuli, the activation of lower-level stimuli diminished more rapidly in the temporal cortex as the flicker frequency increased. Additionally, using magnetoencephalography (MEG), we examined if the frequency-specific activation signals in the temporal cortex. The power spectral analysis revealed a decrease in signal power with increasing flicker frequency across all conditions, with high-level stimuli showing a slightly more gradual decline in signal power. However, there was no significant advantage of high-level over low-level stimuli at a 40Hz frequency. Collectively, our findings suggest that high-level stimuli may be more effective in inducing gamma band entrainment in deep brain structures compared to low-level stimuli.

26.429 THE CATEGORIZATION DIFFICULTY CONTRIBUTES TO THE UNCANNY VALLEY WITHOUT ANIMACY

Kota Sasaki[1] (k.s20000424@gmail.com), Fumiya Yonemitsu[1], Atsunori Ariga[1]; [1]Chuo University

We feel strong eeriness for non-human objects that are highly similar to humans (a phenomenon known as uncanny valley). Previous research has accounted for this phenomenon based on animacy perception (e.g., avoidance to threat or anxiety for mortality) or cognitive processes irrespective of animacy (categorization difficulty hypothesis). However, whether animacy perception is responsible for uncanny valley to be elicited is unclear at present. This is because all the previous studies used objects related to animacy as stimuli (e.g., faces). Therefore, the present study examined whether uncanny valley occurs for non-animacy objects, using geometric figures (square, circle, and triangle). In Experiment 1, we made stimulus figures by systematically morphing an original geometric figure (e.g., circle) into another figure (e.g., triangle). Participants first categorized which original figure (e.g., circle or triangle) the morphed figure was perceived as and then evaluated its likability with a 7-point Likert scale. As the result, we found the uncanny-valley-like phenomenon with non-animacy objects; that is, the figures that took longer times to be categorized (difficult-to-categorize objects) were rated as being less likeable as compared to the figures that were quickly categorized (easy-to-categorize objects). Our results were in favor with the categorization difficulty hypothesis, suggesting that animacy perception is not necessary for the occurrence of uncanny valley. In the following experiments, we hypothesized that the categorization difficulty deteriorates perceptual and cognitive fluency of the morphed figures, resulting in the lower evaluation for them. To assess this hypothesis, we independently manipulated the perceptual (Experiment 2) and cognitive (Experiment 3) fluency of to-be-evaluated figures and investigated how the uncanny-valley-like phenomenon is modulated in response to the processing fluency. Based on these results, we discuss not only the underlying mechanism(s) for the uncanny valley phenomenon but also how to overcome the phenomenon in industry.

This research was financially supported by JSPS KAKENHI Grant Numbers JP21J01431 to FY and JP 22K03210 to AA.

26.430 THE PSYCHOPHYSICS OF STYLE

Tal Boger[1] (tboger1@jhu.edu), Chaz Firestone[1]; [1]Johns Hopkins University

Images vary not only in content, but also in style. When viewing a Monet painting, for example, we see both the scenery it depicts (lilies dotting a pond) and the manner in which it does so (broken brushstrokes, blended colors, etc.). Parsing images in this way is a remarkable perceptual achievement, akin to separating illumination and reflectance to achieve color constancy, or disentangling letter-identities from typefaces when reading. What is the nature of this process, and what are its psychophysical signatures? Here, 9 experiments reveal 3 new phenomena of style perception. (1) Style tuning. Using neural style-transfer models, we rendered natural scenes in the styles of famous artists. Then, inspired by 'font tuning' (wherein text is easier to read in a single typeface than multiple typefaces), we asked observers to scan arrays of images and enumerate all scenes of one type (e.g., mountains). Observers were faster and more accurate in same-style arrays than mixed-style arrays [E1–E2]. Such tuning accumulated over time [E3] and survived controls for color and luminance [E4]. (2) Style discounting. Analogous to 'discounting the illuminant' in color constancy, we find that vision 'discounts' style. Changes to a scene's content (e.g., Monet-pond → Monet-building) were more easily detected than changes to its style (Monet-pond → Klimt-pond; E5), even when low-level image statistics predicted the opposite [E6]. (3) Style extrapolation. After viewing items in a given style (e.g., a fork and knife from one cutlery set), observers misremembered seeing additional items from that style (the spoon from that set; E7), even with low-level similarity equated across lures [E8–E9]. Such errors suggest spontaneous representation of the unseen items — as if mentally 'rendering' objects in newly learned styles. While we typically associate style with more qualitative approaches, our work explores how tools from vision research can illuminate its psychological basis.

NSF BCS 2021053, NSF GRFP 2023351964

26.431 SEEING BEAUTY EVEN WHEN NONE MAY EXIST

Chenxiao Guan[1] (chenxiaoguan@zju.edu.cn), Yongqi Li[1], Mowei Shen[1], Hui Chen[1]; [1]Zhejiang University

Belief, serving as a foundational source of insights for navigating the world, plays a crucial role in shaping cognitive processes such as perception, attention and memory. Considering aesthetics, typically framed as subjective experience, we explore the question: can belief influence or even reshape this experience? In a series of experiments, we designed a specific task demonstrating the role belief plays in aesthetics. In a competition task, two subjects respond as fast as possible to artworks after presented for 3s, followed by feedback on winning or losing. Subsequently, a recognition task and a rating task were carried out. During the recognition task, participants were presented with artworks from the competition and novel pieces. They were prompted to indicate whether they had seen this artwork in the competition and specify whether they won or lost. In the rating task, they provided aesthetic ratings for both the competition and novel artworks. All the artworks were selected from The Vienna Art Picture System and the initial aesthetic ratings between conditions were counterbalanced. The results unveil intriguing patterns. In Exp.1, participants completed the recognition task followed by the rating task, revealing higher aesthetic ratings for novel artworks compared to those actually shown. Surprisingly, artworks recognized (or reported) as previously displayed received significantly higher aesthetic ratings than those recognized as not shown (even when they might not displayed), indicating that the belief of having seen the artwork impacts aesthetic evaluations. Exp.2 altered the task order to avoid the influence of the recognition task on ratings, yet the same effect persisted. In Exp.3, we introduced a more realistic competitive environment and found consistent results. In Exp.4, participants from Exp.3 underwent the same recognition and rating tasks a month later, and the effect endured. Collectively, these findings suggest that the belief possesses the capacity to alter aesthetic ratings.

This work was supported by grants from National Natural Science Foundation of China (No.32171046), Science and Technology Innovation 2030- "Brain Science and Brain-like Research" Major Project (No.2022ZD0210800) awarded to author H.C.

Saturday, May 18, 2:45 – 6:45 pm, Pavilion

Color, Light and Materials: Surfaces, materials

26.432 SELF-SUPERVISED MODELS OF HUMAN TEXTURE-DISCRIMINATION

Abhranil Das[1] (abhranil.das@utexas.edu), Wilson Geisler[1]; [1]University of Texas at Austin

Our ability to interpret scenes builds upon segmenting images into same-texture regions, which are usually the same physical 'stuff'. How do we do this so immediately on arbitrary new images with novel textures, under varying lighting and geometry, with a visual system whose resolution degrades rapidly away from fixation? We conduct the first direct measure of human ability to identify if two 1°×1° grayscale texture patches (widely sampled to prevent learning specific textures) are the same 'stuff', when presented at the same or two different locations, which include the fovea and three eccentricities. We also develop the first models for such general discrimination, using two image-computable approaches that can incorporate biological properties like eye optics and the ganglion-cell sampling resolution at the stimulus location. In the first model, we define important texture features like the luminance histogram, power spectrum, and edge properties. Then we measure and model the statistical distribution of these features across textures, using which we build a Bayesian ideal observer for same-different discrimination. With a single fixed decision boundary on only two features, this discriminates arbitrary texture patches with 93% accuracy, and aligns qualitatively with human performance. The second class of models are convolutional neural networks that can discover new features, using which we also achieve good discrimination. Biological vision does not need numerous labelled samples to learn texture discrimination. Instead, it is likely self-supervised: primitive vision plausibly used coarse features to segment images, then evolution boot-strapped these labels to learn texture discrimination. We implemented this by training the Bayesian texture-discriminator on natural image patches segmented by proximity and color, and the same decision boundary emerged as with explicit texture labels.

26.433 THE EFFECT OF ATTENTIONAL LOAD ON MODAL AND AMODAL COMPLETION

Zeyu Wang[1] (wangzeyu221@mails.ucas.edu.cn), Jinyou Zou[2], Peng Zhang[1]; [1]Institute of Biophysics, Chinese Academy of Sciences, [2]Aier Institute of Optometry and Vision Science, Changsha, China

Perceptual completion or filling-in is a remarkable ability of visual system to interpolate missing information from retinal input. It remains unclear whether modal and amodal completion or filling-in involve different or common neural mechanisms. In this study, we measured tilt after-effect (TAE) after prolonged adaptation to modal and amodal illusory gratings and no-filling-in control stimuli under different attentional load. Four moving gratings were presented in the apertures of a dark occluder. Modal or amodal completion was achieved by manipulating the relative depth between the inducers and the occluder. The phase and speed of the moving gratings were scrambled to generate the no-filling-in control stimuli. In the passive (or no-load) condition, subjects maintained fixation and judged the orientation of probes. In the low and high attentional load conditions, subjects performed either single- or conjunction-feature letter detection tasks in a rapid sequence visual presentation (RSVP) of colored letters. In the passive condition, both modal and amodal filling-in produced stronger TAE than their control counterparts, at similar magnitude. Diverting attention away from the stimuli almost eliminated TAE in both conditions, except for a marginal effect in modal filling-in under low attentional load. Our findings suggest that modal and amodal completion generate similar orientation representations at the early stage of visual processing, and both require attention. Whether lateral or feedback mechanisms are differentially involved in these two filling-in phenomena requires further investigation with neuroimaging techniques.

26.434 CAN MATERIAL-ROBUST DETECTION OF 3D NON-RIGID DEFORMATION BE EXPLAINED BY PREDICTIVE PROCESSING THROUGH GENERATIVE MODELS?

Shin'ya Nishida[1,2] (shinyanishida@mac.com), Mitchell van Zuijlen[1], Yung-Hao Yang[1], Jan Jaap van Assen[3]; [1]Cognitive Informatics Lab, Graduate School of Informatics, Kyoto University, [2]NTT Communication Science Labs, Nippon Telegraph and Telephone Corporation, [3]Perceptual Intelligence Lab, Industrial Design Engineering, Delft University of Technology

Depending on the optical material property of the object (e.g., matte, glossy, transparent), the optical flow generated by a non-rigid deformation of a 3D object dramatically changes. Nevertheless, a recent study (van Zuijlen et al., VSS2022) shows that the sensitivity to detect deformation of a rotating object is similar for matte and glossy objects, and only slightly worse for transparent objects. What makes deformation perception robust to material changes? One possibility is that the visual system constructs a generative model for each object that can correctly predict how the image should change if the object rigidly moves, being able to detect deformation when there is a significant deviation from the prediction. According to this hypothesis, the deformation detection sensitivity would be impaired when extra image deviations from the model predictions are additionally produced by unusual global movements in the surrounding lightfield. In the experiment, the target object was an infinite knot stimulus rotating around a vertical axis, rendered with one of four optical properties (dot-textured matte, glossy, mirror-like, and transparent). The object was deformed by an inward pulling force in seven levels of intensity (including a rigid condition). Using Maxwell Renderer, the movie of each object was rendered under one of three lightfield conditions: static, imploding, or rotating. The object's background was black-masked to make the lightfield change directly invisible to observers. Observers performed a 2-IFC task to choose which of the two stimuli (one being always rigid) deformed more. The results do not support the prediction made by the generative model: light-field manipulation had no significant influence on the deformation detection threshold, nor on the effect of material on the threshold. The results rather support the idea that the visual system effectively ignores the complex flow produced by material-dependent features (e.g., highlights, refractions) in deformation detection.

Supported by JSPS Kakenhi JP20H05957 and a Marie-Skłodowska-Curie Actions Individual Fellowship (H2020-MSCA-IF-2019-FLOW).

26.435 NEURAL REPRESENTATION OF TRANSLUCENT AND OPAQUE OBJECTS IMAGES IN MACAQUE INFERIOR TEMPORAL CORTEX

Hoko Nakada[1] (h.nakada@aist.go.jp), Daiki Nakamura[1], Ryusuke Hayashi[1]; [1]National Institute of Advanced Industrial Science and Technology (AIST) , Human Informatics and Interaction Research Institute

Translucency and opacity are optical properties of materials characterized by the extent of light passing through them. Although recent psychophysical studies have revealed the visual properties associated with the perception of translucency, the neural substrates involved in its perception remain unknown. In this study, we conducted electrophysiological experiments using visual stimuli of objects with various shapes and varying degrees of translucency. To regulate the visual attributes systematically, we used the Translucent Appearance Generation (TAG) model, an unsupervised artificial neural network designed to synthesize images that represent the material appearance. The latent variables of this model encode human-interpretable visual attributes, such as object color in the fine-scale layer and shape in the coarse-scale layers. Notably, variables in the middle layers of the model encode feature information related to whether they appear translucent or opaque. We created visual stimuli by generating object images from sampled latent variables (27 points from coarse-scale layers and 7 points from middle layers) using the TAG model and converting them to grayscale. Mean luminance, Michelson contrast, and the area of the object region were equalized to minimize the effect of these factors on the experimental results. We recorded neural responses using four multi-electrode arrays (each consisting of 128 channels) from the inferior temporal cortex of one macaque monkey. The arrays were placed in three sites in the TE area, and one was in the TEO area. The results of representational similarity analysis showed higher correlations between semi-translucent images than between translucent and opaque images after averaging responses across different shapes. The results of unit-level response analysis indicated that some neurons exhibited selectivity to translucent or opaque objects. These findings suggest that neurons in the inferior temporal cortex represent differences in the object's appearance related to translucency or opacity at both population and unit levels.

The study was supported by the Japan Science and Technology Agency, Moonshot Research & Development Program grant JPMJMS2012, and the National Institute of Information and Communications Technology (NICT) grant NICT 22301.

26.436 SKEWNESS ADAPTATION INDUCED AN ASYMMETRIC EFFECT IN GLOSSINESS PERCEPTION BUT NOT IN TRANSLUCENCY

Hiroaki Kiyokawa[1], Keigo Yoshida[1], Ichiro Kuriki[1]; [1]Saitama University

Human vision can judge a lot of materials and surface qualities, such as glossiness, translucency, etc., just at a glance. However, it is still unclear whether human visual system use different mechanisms for evaluating different material classes. Several previous studies reported that perceived translucency is enhanced by adding clear specular highlights on the object surface. Also, a recent study reported that rating scores between perceived glossiness and translucency are significantly correlated. These findings suggest that some classes of material properties share a common underlying mechanism. We tested this hypothesis using the adaptation effect to the skewness of luminance histogram, which was previously used in a study on glossiness-perception. In our experiment, adaptation effects were tested for whether glossiness and translucency judgments would be affected in the same way or not using adapting stimuli with different polarities of skewness. They were texture images with the same mean luminance and luminance contrast but had skewed luminance histograms, either positively or negatively. The test images were of glossy or translucent objects generated on a computer by Blender software. Participants were instructed to rate perceived glossiness or translucency in a 9-point scale after adapting to images with positive/negative skewness or without adaptation. The apparent glossiness or translucency of the test images were perceptually equated in a preliminary session. The results showed a significant asymmetry in glossiness judgments between the two skewness conditions; glossiness scores were significantly reduced after adaptation to positive skewness than without adaptation, while the scores after adaptation to negative skewness showed no reduction. In contrast, translucency judgments showed almost no effects under the two skewness conditions. These results suggest that the mechanisms for the glossiness and translucency perception could be distinct. Specifically, the mechanisms associated with the luminance histogram skewness appear to differ, at least to some extent.

26.437 PROBING THE RELATIONSHIP BETWEEN MATERIAL CATEGORIZATION AND MATERIAL PROPERTY ESTIMATION USING AMBIGUOUS VISUAL STIMULI

Chenxi Liao[1] (cl6070a@american.edu), Masataka Sawayama[2], Jacob Cheeseman[3], Filipp Schmidt[3], Roland W. Fleming[3], Bei Xiao[1]; [1]American University, [2]The University of Tokyo, [3]Justus Liebig University Giessen

We routinely interact with a wide range of materials. Through visual inspection, we can classify them into categories (e.g., rock), as well as infer their diverse optical (e.g., translucency) and mechanical properties (e.g., rigidity). To investigate if and how human material categorization affects the estimation of material properties, we developed a framework to systematically create images of ambiguous materials. Specifically, by training an unsupervised image generation model (StyleGAN) with transfer learning, we obtained models that synthesize images from three material classes: soaps, rocks, and squishy toys. Via linear interpolation of models' latent spaces and weights, we can smoothly morph one material to another. We sampled ten morphing sequences in which a soap is gradually transformed into a rock and then into a squishy toy in 13 steps. In Experiment 1, ten participants rated each image on five attributes: translucency, glossiness, surface smoothness, rigidity, and brittleness. In Experiment 2, the same participants performed a 10-AFC task of material identification on the same set of images. We found that estimations of mechanical and tactile properties (e.g., rigidity, brittleness, and smoothness) were modulated by morphing. Notably, the rigidity ratings gradually increased along the morphing from soap to rock, followed by a decrease from rock to squishy toy. In contrast, optical properties (e.g., translucency, glossiness) were not correlated with morphing. Finally, participants were uncertain about the material identity of images close to the midpoint of cross-material morphing, sometimes perceiving them as entirely different materials like jelly, candy, wax, and glass. Such morphed materials with high category ambiguity also show high variance in estimations of mechanical properties. Together, our results suggest that material categorization significantly impacts the inference of mechanical properties, especially when material identity is ambiguous.

NIH-1R15EY033512-01A1

Saturday, May 18, 2:45 – 6:45 pm, Pavilion

Color, Light and Materials: Art, cognition

26.438 CHANGING THE STRUCTURE OF COLOR CATEGORIES CAUSALLY INFLUENCES COLOR-CONCEPT ASSOCIATION GENERALIZATION

Melissa A. Schoenlein[1] (schoenlein@wisc.edu), Melina Mueller[1], Karen B. Schloss[1]; [1]University of Wisconsin-Madison

People form associations between colors and concepts from experience (Schoenlein & Schloss, 2022), but questions remain concerning how continuous color-concept association distributions are populated from sparse input. The category extrapolation hypothesis proposes that associations for an input color (e.g., a blue) extrapolate to all other colors with the same color category (i.e., all blues) (Rathore et al., 2020). Schoenlein and Schloss (VSS-2022) found that color category boundaries predicted patterns of color-concept association generalization, but this result was correlational. Here, we tested whether category boundaries causally influence association generalization when participants form novel associations between colors and the concepts of two alien species, Slubs and Filks. The category extrapolation hypothesis implies that if participants learn to merge hues from different categories (e.g., blue and purple) into a single overarching category (e.g., blue+purple), they will show greater generalization across category boundaries compared to participants who keep those categories separate. We tested this hypothesis in an experiment with two groups (blue+purple mergers and green+yellow mergers), who completed three tasks. In the first, color naming task, blue+purple mergers learned new color names which merged blues/purples and separated greens/yellows (green+yellow mergers learned the opposite). In the second, color-concept association formation task, participants learned to associate Slubs with a particular blue and Filks a particular green. In the last, color-concept association task, participants indicated (yes/no) whether they associated each species with colors in sequences spanning blue-to-purple and green-to-yellow. These judgments produced psychometric functions over color sequences and the bias term indicated degree of generalization. The results showed that merging hues from different categories resulted in greater generalization (interaction between merging group and color sequence, $p=.008$). This study provides the first causal evidence that color category boundaries influence color-concept associations, indicating that category extrapolation influences how continuous color-concept association distributions are populated from sparse input.

NSF award BCS-1945303 to KBS

26.439 ESTIMATING HUMAN COLOR-CONCEPT ASSOCIATIONS FROM MULTIMODAL LANGUAGE MODELS

Kushin Mukherjee[1] (kmukherjee2@wisc.edu), Timothy T. Rogers, Karen B. Schloss; [1]University of Wisconsin Madison, [2]University of Wisconsin Madison, [3]University of Wisconsin Madison

Color-concept associations are important for many facets of visual cognition from object recognition to interpretation of information visualizations. Thus, a key goal in vision science is developing efficient methods for estimating color-concept association distributions over color space. Such methods may also inform how people form associations between colors and abstract concepts despite these color-concept pairs never co-occurring in the natural world. To this end, we investigated the extent to which GPT-4, a multimodal large language model (LLM), could estimate human-like color-concept associations without any additional training. We first collected human association ratings between 70 concepts and a set of 71 colors spanning perceptual color space (UW-71 colors). We then queried GPT-4 to generate analogous ratings when given concepts as words and colors as hexadecimal codes, and compared these association ratings to the human data. Color-concept association ratings generated by GPT-4 were correlated with human ratings (mean r across concepts = .67) at a level comparable to state-of-the-art methods for automatically estimating such associations from images. The correlations between GPT-4 and human ratings varied across concepts (range: r = .08 – .93), with the correlation strength itself predicted by the specificity (inverse entropy) of the color-concept concept association distributions (r = .57, p < .001). Although GPT-4's performance was also predicted by concept abstractness (r = -.42, p < .001), this effect was dominated by specificity when both factors were entered into a model together (specificity: p < .001, abstractness: p = .25). These results suggest that GPT-4 can be used as a tool for estimating associations between concepts and perceptual properties, like color, with better accuracy for high-specificity concepts. They also suggest that learning both word-to-percept structure and word-to-word structure, as multimodal LLMs do, might be one way to acquire associations between colors and abstract concept words.

NSF award BCS-1945303 to K.B.S.

26.440 A BAYESIAN ANALYSIS OF THE PHYLOGENETIC DEVELOPMENT OF MONOLEXEMIC COLOR TERMS IN DRAVIDIAN LANGUAGES.

Male Shiva Ram[1] (shivaram.ram180@gmail.com), Phani Krishna P[2]; [1]School of Medical Sciences, University of Hyderabad, [2]Department of Linguistics, University of Delhi

Color naming and use is a diverse concept, however English and other languages use eleven primary-color prototypes (Berlin and Kay, 1969), pioneering study on color names neglected Dravidian languages that are old. Color vocabulary exists in South Dravidian (Kapp, 2004). We investigated the monolexemic theory of color and proposed employing corpus and Bayesian interference models to build color word evolution and existence. (Male, 2023) on color word categorization may provide insight into the migration of color terminology in Dravidian and other languages. Tamil is a Dravidian language (Kolipakam, 2018) that goes back 4500 years. These theories may aid in our understanding of how cognitive and cultural links generate language. The phylogenetic models' primary results highlighted the history and development of monolexmic color words in Dravidian languages. These evolutionary characteristics of color words demonstrate that color term universality is not advocated in Dravidian languages. Interestingly, monolexmic color terms exist exclusively in Tamil, which is the principal reference that has spread to other languages during the history of development, civilization, and linguistic migrations among the Dravidian culture and geographical territories in India. Implementing phylogenetic models in these monolexmic color terms may disclose new color terms' existence or discovery in Dravidian ancient languages.

Prof. Arulmozi, S, CALTS, UoH for Corpus database assistance and The Institute of Eminence-University of Hyderabad

26.441 ENSEMBLE CODING OF COLOR IN A PILE-SORT TASK

Delwin Lindsey[1] (lindsey.43@osu.edu), Angela Brown; [1]The Ohio State University

In sequential color sorts, subjects partition a palette of diverse color samples into n-piles (n=2…N). The composition of the piles often approximates the basic color categories associated with Berlin & Kay's theory of color term evolution. Here, we show that this result depends on the to-be-sorted colors. Subjects saw a randomly-ordered array of iPad color samples and "slid" them into virtual "bins". We obtained the standard result when 44 subjects sorted a 25-color palette ("P1") that varied smoothly in hue at approximately constant Munsell value (lightness), and contained good examples of the basic color categories. However, we obtained something quite different when 73 subjects sorted a similarly-sized palette ("P2") that sparsely sampled Munsell color space, with multiple values for many hues, while still containing good examples of the basic categories. P2 subjects most frequently sorted by value or by complex combinations of hue and value. We then obtained sequential pile-sorts (n=2…6) of palette "P3", which varied smoothly in hue, with one random Munsell value for each hue. For P3, the iPad tracked the order in which subjects populated the bins. In aggregate, 49 subjects' P3 results were indistinguishable from the P1 results. We propose that pile-sorts are mediated by processes analogous to Ensemble Coding, whereby subjects plan pile-sort strategies based on pre-determined color categories, prior to executing an n-sort. Consistent with this hypothesis, P3 subjects often populated the bins sequentially (43% of piles), placing most samples belonging to one category in one bin, before moving on to place the next category of samples into another bin. This result is inconsistent with the proposal that subjects generally adopt an ad-hoc strategy based on the pairwise similarities among color samples. Unlike traditional categorical parsing of colors by hue, our results also highlight the importance of the covariation of both hue and lightness.

BCS-1152841 to DTL

26.442 TOP-DOWN KNOWLEDGE CAN AFFECT PERCEPTION WHEN THE INPUT IS AMBIGUOUS

Michael Cohen[1,2] (michaelthecohen@gmail.com), Andrew Rios[1], Esther Min[1], Rosa Lafer-Sousa[3]; [1]Amherst College, [2]MIT, [3]National Institute of Mental Health, NIH

Does what we see depend on what we know? Many findings suggest that top-down factors such as emotions, desires, and categorical knowledge affect perception and can change an item's appearance. However, a variety of methodological flaws, replication failures, and extremely small effect sizes render these findings inconclusive, leading to tremendous debate in the literature. Here, we searched for an uncontroversial example of top-down factors altering the appearance of items using a novel variant of the memory color effect. The memory color effect is a phenomenon in which familiar grey items look like their canonical color (e.g., a grey banana looks faintly yellow, Hansen et al., 2006). Although it has been thoroughly studied, the memory color effect remains extremely controversial, with many researchers doubting its very existence (Zeimbekis, 2013; Gross et al., 2014; Valenti & Firestone, 2019). Here, we asked, can the memory color effect provide us with clear evidence of top-down factors affecting perception? We theorized that under globally ambiguous viewing conditions, categorical knowledge may impact perception since that knowledge will be used to disambiguate underdetermined sensory input. Thus, we focused on one specific globally ambiguous viewing condition: very dim light. Specifically, we showed participants images of highly familiar grey items printed on poster paper under very dim lighting (i.e., mesopic), which created a lighting condition in which discerning all the attributes of the item is difficult. As predicted, we found an extremely strong memory color effect for familiar items under dim light, but not under bright, unambiguous light. No such effect was found for unfamiliar items. Moreover, this effect could be subjectively appreciated by observers who reported that the item just "looks" colorful, even though it is not. These results demonstrate that in certain situations, top-down factors can directly affect perceptual experience and appreciably alter how items appear.

26.443 LONG-TERM SEMANTIC KNOWLEDGE PREDICTS CHANGES IN COLOR PERCEPTION

Alexis Green[1], Bevil Conway[1], Rosa Lafer-Sousa[2]; [1]Laboratory of Sensorimotor Research, National Eye Institute, NIH, [2]Laboratory of Neuropsychology, National Institute of Mental Health, NIH

Do top-down factors like knowledge affect perception (Fodor, 1983; Pylyshyn, 1999)? We address the question by assessing the impact of long-term semantic knowledge on perception of an ambiguous color image, #theDress. The image became viral around the globe in 2015; it is of a closely cropped photograph of a striped dress, which is perceived by some as blue and black (BK) and by others as white and gold (WG). In reports that followed shortly after, the population was divided roughly equally into these two groups of perceivers. In the ensuing years, the "true" color of the dress became widely known: in real life, the dress is unambiguously perceived as blue and black. We wondered whether this knowledge has changed how people perceive the image. If long-term semantic knowledge impacts color perception, we predict that the proportion of BK observers will have increased. On the other hand, if perception of the image is solely determined by retinal signals, we predict the proportion of BK and WG observers will be unchanged. We collected a new data set (n=688), replicating the online experiment we performed eight years earlier (n=264) in the days after the image emerged. Subjects had to view the image and provide verbal color reports, perform color matching, and answer questions about their experience with the image. In both populations, we found that having experience with the image significantly increased the frequency of BK reporting (permutation test, $p < 0.0001$, $p < 0.004$, 2015 and 2023 datasets respectively) and decreased WG reporting ($p = 0.02$, $p = 0.03$). Further, between the experienced observers from both datasets, BK reporting increased substantially over time (permutation test, $p = 0.009$) while WG reporting decreased ($p = 0.004$), consistent with the hypothesis that long-term semantic knowledge can change color perception.

26.444 EYES, STILL LIFES & EIDOLONS – THE ROLE OF COLORS AND CONTOURS WHEN VIEWING STILL-LIFE PAINTINGS?

Doris I. Braun[1] (doris.braun@psychol.uni-giessen.de), Karl R. Gegenfurtner[1]; [1]Giessen University Dept. of Psychology

How important are object colors and contours for the guidance of eye movements? We selected high-quality digital reproductions of 20 still-life paintings of fruit and flower arrangements from the 17th-18th century from a recent database (Van Zuijlen et al., 2021). The "eidolon factory" (Koenderink et al., 2017) was used to disrupt object recognition while maintaining local image structure. To investigate the influence of color and contour information on visual exploration we presented each still life image on a computer monitor in four versions: as original in color and grayscale and as Eidolon in color and grayscale. 26 young participants first had to indicate the most interesting part of each image after it was briefly flashed for 120 ms. In a second experiment, they were asked to view each image for 5 s and then to rate how much they liked it on a 7-point scale. During the image exploration eye movements were recorded with an EyeLink 1000. For each participant and image we analyzed the number and amplitude of saccades, the position and duration of fixations and the extent of image exploration (spread). Eye movements reflected content and structure of each image. When objects were defined by contours as in both versions of the originals, significantly more saccades were made, saccade amplitudes and fixation durations were significantly shorter, and spreads were larger. Also inter-subject agreement was significantly higher when contour information was present. This was even further improved by color. For both Eidolon versions, impaired object recognition prolonged fixation durations and increased saccade amplitudes. For eidolon images, color information had a major effect on which image regions were preferentially fixated. We conclude that contour and color information results in a more thorough exploration of the paintings with shorter fixations, smaller saccades, a larger spread and higher intersubject agreement.

Supported by ERC Advanced Grant Color 3.0 (Grant no. 884116) and Deutsche Forschungsgemeinschaft SFB/TRR 135 A1,C2 (project number 222641018)

26.445 ADAPTING TO ART: ADAPTATION ALTERS IMPRESSIONS OF IMPRESSIONIST PAINTING STYLES

Fatemeh Charkhtab Basim[1] (f.ch.basim@gmail.com), Michael A. Webster; [1]University of Nevada, Reno

Painting styles such as impressionism involve coarse application of color, resulting in a blotchy perception of discrete strokes or daubs when viewed from a close distance while appearing more fused and natural from a far distance. But how does it look to the painter? We asked whether artists might adapt to the visual textural qualities of their paintings. Stimuli were images of 1/f color noise formed by combining separate noise images for the red, green, and blue components, and subtended 7.5 by 10 deg on a larger gray monitor background. The noise images were altered by averaging the colors into 0.25 deg blocks or filtering with the 'David Mills' impressionism MATLAB function. For each of three test images we then generated a set of 100 blends that varied from the original noise to fully pixelated. Six observers adjusted the blend level using a staircase procedure until the pixilation was noticeable, either before or after adapting to a rapid succession of either the blocked or impressionist images that were randomly resampled every 0.2 sec for 180 sec. Thresholds for detecting the stylization increased by roughly 35% after adaptation – a difference that was significant for all images from both texture types - showing that adaptation reduced sensitivity to the manipulated textures. In a second experiment, the blocked and impressionist images were blended in 100 steps, and a staircase was used to estimate the boundary for classifying the style. Adaptation shifted the boundaries toward the adapting style consistent with a selective sensitivity loss for the style. These results are consistent with other studies demonstrating adaptation to visual texture or to image noise, and suggest that artists might perceive their works in unique ways because they may be uniquely adapted to them. Supported by EY-010834

26.446 UGLY COLOURS: CHROMATIC DETERMINANTS OF IMAGE AESTHETIC VALENCE

Jan Kučera[1] (j.kucera1@ncl.ac.uk), Dalya Al-Shahrabi[2], Fatima Felisberti[3], Alejandro Parraga[4], Anya Hurlbert[1]; [1]Newcastle University, [2]Northumbria University, [3]Kingston University London, [4]Universitat Autonoma de Barcelona

Affective and aesthetic responses to images may be partly determined by spatiochromatic statistics, independently of semantic meaning or content (e.g. McAdams et al. 2023, J. Vis.). Such relationships are generally explored through positive rather than negative responses – e.g. liking or beauty ratings, not disliking or ugliness. Here we examine the extent to which chromatic properties alone may predict both perceived beauty and ugliness, in photographs. We further examine whether differences in colour distributions between beautiful and ugly photographs are predictable from abstract colour preferences; in Western culture, people of both sexes and different ages tend to prefer colours with cooler hues and higher saturation and lightness and to dislike dark yellow hues (browns). Methods: In an online experiment, 94 participants (94% of which were aged 16-25 years) each submitted four recently taken photographs, two that they considered ugly and two beautiful, with an indoor and outdoor scene for each aesthetic category. Photographs were restricted to scenes without people or other animate objects. In a subsidiary online experiment, 6 new participants independently ranked hue salience in each submitted image. Results: Analysis of the image statistics showed that across both scene types, colourfulness was significantly higher for beautiful vs. ugly pictures, and for outdoor images only, mean image chroma was significantly higher for beautiful pictures. Hue distributions - whether obtained by an automated colour category classification algorithm or from the independent human ratings - also differed between beautiful and ugly pictures. In outdoor pictures, the proportion of "blue" was significantly lower and "brown" and "grey" significantly higher in ugly vs. beautiful categories. The results suggest that aesthetic responses to pictures are partially predicted by affective responses to colour, both positive and negative. The results further suggest that perception of ugliness is not merely the opposite of beauty, but a distinct process.

26.447 ART HAS NO GENDER, ONLY GENDER BIAS

Claudia Damiano[1] (claudia.damiano@utoronto.ca), Stefanie De Winter[2], Nicole Ruta[2], Johan Wagemans[2]; [1]University of Toronto, [2]KU Leuven

The Abstract Expressionism art movement was predominantly showcased by men, leading to the exclusion of women artists from the art canon. Art critics of the time also deemed women's art as "decoration" whereas men's art was considered "real art", further marginalizing women artists. Based on this historical context, our study examined whether and how the appreciation of art is influenced by gender biases. Across three experiments (N=800) using 160 Abstract Expressionist paintings by an equal number of men (Pollock, Louis, Twombly, Kline) and women (Krasner, Frankenthaler, Mitchell, Hartigan) artists, we tested whether people show a bias for attributing authorship of artworks to men, and investigated whether artworks created by men and women are evaluated differently. Results confirmed that participants were significantly more likely to judge that the artworks were painted by men (57.07%, $p < 0.001$), regardless of the actual gender of the artist, specifically if the artworks contained higher levels of black paint (beta = 0.36, $t(466) = 6.28$, $p < 0.001$) and angular lines (beta = 0.15, $t(466) = 2.37$, $p = 0.018$). Experiments 2 and 3 revealed that artworks thought to be painted by men received significantly lower scores for attributes describing so-called "woman art" (e.g., "decorative", "childish"), but also for attributes describing "good art" (e.g., "vibrant", "memorable"), as well as for liking, pleasure, interest, and order (all $p < 0.005$). These findings suggest that art criticism of the 60s is no longer reflected in art perception by a general audience nowadays. In summary, while there appears to be a general bias in assuming the paintings were created by men, the actual gender of the artists did not significantly impact overall appreciation of the artworks. Instead, the authorship bias, influenced by stylistic elements within the paintings, plays a more significant role in aesthetic evaluations of the paintings.

This work was funded by an FWO Junior Postdoctoral Fellowship to S.D.W. (12E0723N), as well as Methusalem Funding (METH/21/02) and an ERC Advanced Grant to J.W. (GRAPPA, 101053925).

Saturday, May 18, 2:45 – 6:45 pm, Pavilion

3D Perception: Depth cue integration, neural mechanisms

26.448 DISCRIMINATION THRESHOLDS REFLECT TASK-RELATED, COGNITIVE PROCESSES RATHER THAN CUE UNCERTAINTY IN DEPTH PERCEPTION

Jovan Kemp[1] (jovan_kemp@brown.edu), Fulvio Domini[1]; [1]Brown University

Sensory uncertainty typically refers to the combined effect of the ambiguity in world-to-percept mappings and the inherent stochasticity of the neural processes underlying sensory processing. In normative probabilistic models of sensory cue integration, uncertainty is assumed to be the primary cause of variance in perceptual judgments, while the effects of later-stage, task-related cognitive processes (such as working memory (WM) maintenance) are often considered

negligible. This assumption is critical in testing normative models for optimally combining ambiguous cues to reduce judgment variance. However, the presupposition that variance in perceptual judgments is solely related to sensory uncertainty is rarely assessed. Here, we examine the role of task-related cognitive processes in visual depth judgments. Specifically, we investigated the relationship between WM and depth discrimination thresholds, which describe the depth at which one stimulus can reliably be discriminated from another, for virtual 3D paraboloids specified by binocular disparity and texture cues. We conducted a 2-interval forced choice (2-IFC) task whereby participants must retain the depth of two stimuli presented at different time intervals in WM. We then manipulated the difficulty of WM encoding and maintenance. In one experiment, we employed a backward mask to selectively disrupt the WM fidelity of the stimuli presented during the 2-IFC task. In another experiment, we manipulated the interval length between the presentation of the two stimuli, allowing the WM encoding to degrade before a judgment was made. Overall, we show that these manipulations reliably increase the measured thresholds and therefore demonstrate an integral role of cognitive processes on sensory judgment variance. Critically, we show that predictions for variance reduction still hold after WM manipulation. In agreement with previous findings predicted by a deterministic theory of 3D cue integration termed Intrinsic Constraint, these results provide converging evidence that judgment variance solely reflects cognitive processes and not cue uncertainty.

This material is based upon work supported by the National Science Foundation under Grant No. 2120610

26.449 CONTRIBUTIONS OF ABSOLUTE BINOCULAR DISPARITY, MOTION PARALLAX AND ANGULAR DECLINATION TO ABSOLUTE TARGET LOCALIZATION

Lingling Bai[1], Lizhu Yan[1], Teng Leng Ooi[2], Zijiang He[1]; [1]University of Louisville, [2]The Ohio State University

One accurately localizes a target in the full-cue environment using external depth cues and the ground surface. However, in the dark, a target beyond 3-4 m is often perceived at the intersection between its projection line from the eyes and an implicit curved surface (intrinsic bias). Thus, perceived target distance increases as the target's angular declination decreases. Yet, it is unknown how angular declination information interacts with absolute binocular disparity and absolute motion parallax to determine perceived target location in the dark. We employed the blind walking-gesturing paradigm to measure judged location of a 0.2 degree target in four conditions (monocular-static, binocular-static, monocular-motion-parallax & binocular-motion-parallax). Motion-parallax was initiated by the observer laterally displacing their head and body by 0.4 m (2 cycles) while judging the target. Twelve target locations [4 distances (1.5, 3.25, 5.75, and 7.0 m) x 3 heights (0.14, 0.74 m, and eye level)] were tested. The average results (n=8) reveal judged target locations in the monocular-static condition transcribed a curvilinear profile, reflecting the intrinsic bias. In the other three testing conditions that carried binocular disparity and/or motion parallax cues, the targets at or nearer than 3.25 m were judged significantly nearer to their physical distances (more accurately) than at the intrinsic bias, whereas targets at or farther than 5.75 m were not judged significantly farther beyond the intrinsic bias. For example, for the targets at eye level, the two nearer ones (1.5 & 3.25 m) were perceived significantly nearer (more accurately) in the binocular-motion parallax than in the monocular-static condition ($p<0.001$ & $p<0.05$); whereas perceived distances of the two farther targets (5.75 & 7.0 m) were similar. This indicates absolute binocular disparity and motion parallax are effective depth cues within a 3.25 m distance range in the dark, and beyond that, angular declination influences target localization.

NIH R01EY033190

26.450 NATURAL SIZE-DISTANCE SCALING REDUCES, BUT DOES NOT ELIMINATE, DEPTH MATCHING ERRORS FROM CONFLICTING OCCLUSION AND STEREOPSIS

Domenic Au[1] (domau@my.yorku.ca), Robert Allison[1], Laurie Wilcox[1]; [1]York University

Stereoscopic depth matching is significantly degraded when occlusion information conflicts with depth from binocular disparity. However, in these studies the visual angle of the target was held constant while in natural viewing image size changes linearly with object distance. Thus, it is not clear whether the disruption in performance can be solely attributed to the discrepant occlusion and disparity signals. Here we evaluated the combined effects of size, occlusion and binocular disparity using a depth matching paradigm in mixed-reality. The virtual stimulus was a green letter 'A' presented using an augmented reality display. It was superimposed on a physical surface with variable transparency fixed at 1.2 m. The target letter was placed at one of eight distances (0.9 - 1.6 m), including the surface location. The letter was rendered either with a fixed size (variable retinal angle) or with size scaled to maintain a constant visual angle. Observers matched the distance of a virtual probe to the perceived distance of the letter in three conditions where the surface was: opaque, transparent or absent. We found that depth matches were accurate and there was no effect of size scaling in both the 'no surface' and 'transparent surface' conditions. This was also the case when the letter appeared in front of the surface. However, when the letter was positioned beyond the opaque surface (maximum cue conflict) its position was systematically underestimated. Introducing correct size scaling reduced this error but did not eliminate it. In sum, when occlusion and disparity are in conflict our results show that using a fixed retinal size exacerbates depth matching errors. When retinal size varies (as in the natural world) depth matching is more accurate but significant underestimates remain. When occlusion and disparity signals are in agreement, size has little impact on performance.

Natural Sciences and Engineering Research Council

26.451 THE EFFECT OF REFLECTANCE, DEPTH GAIN, AND SCENE COMPLEXITY ON PERCEIVED DEPTH

Rebecca Hornsey[1,2] (rlhornsey@outlook.com), Loes van Dam[2,3], Rebecca Ranson[2], Paul Hibbard[1,2]; [1]University of Stirling, [2]University of Essex, [3]Technical University of Darmstadt

3D shape can be inferred from a variety of different visual cues. Here, the way surface reflectance, depth gain, and scene complexity may influence our perception of depth on 2D images was tested. The perceived 3D shape from images of household objects was measured using a gauge figure: the slant and tilt of which was adjusted by the participant, at multiple sample points on the surface of the target object. These settings were then used to create the pictorial relief for the objects. First, the effect of three surface conditions on perceived shape was assessed: 1) full colour and texture; 2) a grey-scale representation, which preserved texture and reflectance cues; 3) a uniform grey reflectance, which provided only shape from shading cues to global shape and surface texture. Following this, the effect of scene complexity was explored, to assess whether perceived depth of individual objects can be affected by the number of objects observed. Finally, the gain was varied by squashing and stretching the depth of the target object, to assess the extent to which observers could perceive metric depth from pictorial cues. The results show that perceived depth was closely related to the ground truth shape, but was not affected by the surface manipulations, or scene complexity. These results show that observers were able to make use of shape from shading and surface texture cues to perceive 3D shape. However, pictorial relief did not vary with the depth gain applied to the objects. These results show that, for these viewing conditions, pictorial relief was limited to a bas relief ambiguity, and observers were not able to recover scaled metric depth.

The Leverhulme Trust, Economic and Social Research Council

26.452 QUANTIFYING THE MECHANISMS FOR THE ROLE OF VISUAL CONTEXT ON ORIENTATION JUDGMENTS

Ernestine Brannon[1], Paul Dassonville[1]; [1]University of Oregon

Our ability to judge the orientation of an object is dependent on the effects of context at multiple levels of processing. In early visual processing, orientation contrast effects can enhance the perceived disparities between edges in the object and surround. In later stages of processing, visuovestibular cues in the environment (edges of window frames, doors, desktops, etc.) provide cues to gravitational vertical that inform the observer's egocentric reference frame. A combination of these mechanisms is thought to underlie the rod-and-frame illusion (RFI), where the perceived orientation of a line is distorted in the presence of a surrounding, tilted frame. To test this, we compared the perceptual effect of the RFI to separate measures of the visuovestibular distortion and orientation contrast caused by tilted frames of different sizes. Visuovestibular distortions were measured in a task that had participants saccade from a fixation point to the "topmost point" on a surrounding response circle, which was itself surrounded by a tilted frame. Orientation contrast effects were measured by having participants saccade from a fixation point to the location where the rod of the RFI would intersect the surrounding response circle if it were extended upward. While the magnitude of the visuovestibular effect remained constant across frame sizes within the range tested here, the orientation contrast effect was greatest for the smallest frame. As predicted, the sum of the separately-measured visuovestibular and orientation contrast effects were found to approximate the overall perceptual effect of the illusion regardless of frame size, and this was true on average as well as when assessing individual differences. These findings support the hypothesis that perceived orientation is affected by both orientation contrast effects and visuovestibular distortions, and that the magnitudes of these effects can be separately quantified with these two saccade tasks.

26.453 SLANT DISCRIMINATION PERFORMANCE FOLLOWS PATTERNS PREDICTED BY BINOCULAR VIEWING GEOMETRY

Stephanie M Shields[1] (smshields@utexas.edu); [1]The University of Texas at Austin

The visual system uses information in the two 2D retinal images to estimate the 3D structure of the environment, including the 3D orientation of surfaces. The information available depends on projective geometry, the mapping of the 3D environment onto the two retinae. One could expect that, as a result, projective geometry has shaped how the system processes 3D scene structure and can be used to at least partially explain patterns in perceptual performance. To test whether that prediction holds true for the perception of 3D orientation, the present experiments measured human performance on a two-interval forced choice slant discrimination task. Within an experiment, stimulus slant and viewing distance varied, and across experiments, the slant cues in stimuli varied. Observers likely combine information from multiple cues when estimating slant in everyday scenarios, so the first experiment used a cue-rich stimulus, specifically a square plane with texture that approximated pink noise. To assess the impact of various sources of information on performance, the cues available to observers were reduced in the three subsequent experiments, either by removing binocular disparity, texture, or global shape from the cue-rich stimulus. Results confirmed that observers weighted cues differently. Removing texture perturbed performance the least, suggesting that monocular and binocular shape cues most strongly influenced performance. A notable source of shape cues was the projection of the stimulus's vertical borders, so the geometry of that projection was used to predict performance patterns. In line with those predictions, observers generally performed best at high magnitude slants (beyond approximately +/- 70 degrees), and their performance typically worsened as slant approached frontoparallel (0 degrees) and as viewing distance increased. Thus, the results of these experiments both further understanding of slant perception and support the theory that projective geometry can explain patterns in perceptual performance.

National Science Foundation Graduate Research Fellowship Program

26.455 BINOCULAR DEPTH INFORMATION MODULATES OBJECT-SELECTIVE ACTIVATION IN HIGH-LEVEL VISUAL CORTEX

Tasfia Ahsan[1] (ahsant@my.yorku.ca), Eva Deligiannis[2], Karsten P. Babin[2], Rebecca L. Hornsey[1], Laurie M. Wilcox[1], Jody C. Culham[2], Erez Freud[1]; [1]Department of Psychology, York University, [2]Department of Psychology, Western University

Despite the fact that we know a great deal about the processing of binocular disparity for synthetic stimuli (such as random dot stereograms) in early stages of visual processing (e.g., V1, V2), little is understood about the contribution of stereopsis to the processing of more naturalistic stimuli in later stages of the visual hierarchy. Here we used functional magnetic resonance imaging (fMRI) to investigate how binocular vision contributes to high-level neural processing of visual objects. Participants viewed objects and scrambled objects (along with faces, scenes and bodies, not discussed here) as 2D or stereoscopic 3D images with a high-quality 3D MRI projector and polarized eyewear. High-resolution 3D objects models were used and displayed using the average human interpupillary distance. The objects were re-scaled and rotated to attain a similar disparity range across the set. Blender software was used to segment the objects into similar parts; importantly each fragment preserved its original disparity profile. The parts were then redistributed to create a cloud of fragments. In the 2D condition one of the images of the stereopair was presented to both eyes. Additionally, we ensured that the viewing geometry was consistent with natural viewing of real objects. Our results show stronger activation for 3D stimuli versus 2D stimuli in both the dorsal and ventral visual streams. More surprisingly, 3D viewing decreased object-selectivity (objects – scrambled objects) in shape-selective regions such as LOC and V3A. This modulation in selectivity was due to a greater increase in activation with the addition of stereopsis (3D – 2D) for scrambled than intact objects. Our results suggest the high-level regions that process different visual categories might be differentially sensitive to availability of binocular disparity.

Natural Sciences Research Council of Canada (NSERC); CF-REF program Vision Sciences to Applications (VISTA)

26.456 INVOLVEMENT OF CEREBELLAR VERMIS IN THE PERCEPTION OF DEPTH FROM MOTION EXPLORED WITH TRANSCRANIAL MAGNETIC STIMULATION

Emily Johnson[1], Shane Corbett[1], Abuk Akech[1], Haylee Hardin[1], Samuel A. Birkholz[1], Mark Delisi[1], Jeffrey S. Johnson[1], Mark Nawrot[1]; [1]North Dakota State University

Current models suggest that the unambiguous perception of depth from motion parallax (MP) relies on the integration of retinal image motion with a pursuit eye movement signal, which previous work has suggested may be generated by the frontal eye fields (FEF). In the present study, we used Transcranial Magnetic Stimulation (TMS) to explore a possible role of the Cerebellar vermis (CV) in these computations. Both physical lesions to and TMS of the CV, but not the cerebellar hemispheres, have been found to disrupt motion perception and produce smooth pursuit deficits. However, a possible role of the CV in the computation of depth from MP has not been explored. In the present study, we used TMS to investigate this possibility. Triple-pulse (33 hz) TMS was applied to mid-line CV (1 cm below the inion) at the time of visual stimulus onset (0 msec ISI) during three tasks: i) pursuit, ii) motion perception, and iii) MP depth perception. The pursuit task was a step-ramp. Psychophysical tasks required the observer to report perceived depth or perceived motion direction (2AFC) of a computer-generated random-dot stimulus making a single translation, leftward or rightward, of duration (t). Between trials, t varied in two interleaved staircases, one for each direction of stimulus translation. Average performance from 35 participants indicates that TMS of CV produced an overall decrease in pursuit latency (13 msec for leftward and 11 msec for rightward translation). Similarly, for depth perception, TMS produced a decrease in pursuit latency in both directions (25 msec for leftward and 18 msec for rightward translation). TMS produced no change in motion perception latency for either direction of stimulus translation. These results suggest that cerebellar vermis may play a role in the integration of the pursuit signal needed for the unambiguous perception of depth from MP.

Supported by NIH NEI R15EY031129

26.457 INVOLVEMENT OF CORTICAL AREA MT IN THE PERCEPTION OF DEPTH FROM MOTION EXPLORED WITH TRANSCRANIAL MAGNETIC STIMULATION

Shane Corbett[1], Emily Johnson[1], Haylee Hardin[1], Abuk Akech[1], Samuel A. Birkholz[1], Mark Delisi[1], Jeffrey S. Johnson[1], Mark Nawrot[1]; [1]North Dakota State University

The unambiguous perception of depth from motion requires the integration of visual retinal motion with a signal indicating pursuit eye movement direction. In previous work (Nawrot & Johnson, 2023, ECVP), we used Transcranial Magnetic Stimulation (TMS) to show that the Frontal Eye Field (FEF) may be the source of the necessary pursuit signal. In the current study, we used TMS to investigate the role of cortical area MT in these computations. Right hemisphere MT was localized as the region approximately 5 cm lateral and 3 cm superior to the inion that, when stimulated, produced visible phosphenes in the contralateral visual field. Following localization, triple-pulse (33 Hz) TMS was applied to the right MT of 35 participants either 30 msec before (early stimulation) or 90 msec after visual stimulus onset (late stimulation) during the performance of three different psychophysical tasks: i) pursuit, ii) translational motion perception, and iii) MP depth perception. The pursuit task was a step-ramp. Psychophysical tasks required the observer to report perceived depth or perceived motion direction (2AFC) of a computer-generated random-dot stimulus making a single translation (leftward or rightward) of duration (t). Between trials, t varied in two interleaved staircases, one for each direction of translation. TMS of right hemisphere MT produced a small decrease in pursuit latency in the contraversive direction (7 msec early and 5 msec late), but no change in ipsiversive pursuit. For depth perception, TMS produced 14 msec (early) and 20 msec (late) of speeding for leftward stimulus translation. For rightward stimulus translation, TMS produce 13 msec of speeding for late, but no change for early stimulation. There was no change in motion perception latency for either direction of stimulus translation, for either early or late TMS. The concordance of pursuit and depth effects suggests MT may have a role in their integration.

Supported by NIH NEI R15EY031129

26.458 STEREOMOTION SCOTOMAS: AN IMPAIRMENT OF VELOCITY-BASED MECHANISMS REVEALED BY VARIATION OF STIMULUS SPEED

Amna Dogar[1] (aad9096@nyu.edu), Caterina Pedersini[1], Bas Rokers[1,2]; [1]Psychology, New York University Abu Dhabi, [2]Center for Brain and Health, New York University Abu Dhabi

In otherwise typical observers stereomotion scotomas can occur, an intriguing impairment restricted to specific locations in the visual field in the perception of motion-in-depth based on binocular cues. Potential causes for the impairment are failures to process changes in binocular disparity over time (CD) and interocular velocity differences (IOVD). Previous work using cue-isolating stimuli has suggested that IOVD failures are the primary contributor (Barendregt, Dumoulin, Rokers; 2016). Here we explored a method to identify the source of failure using stimuli that always contained both cues. Participants discriminated the direction (toward/away) of a field stereomotion of dots moving for 1 second within circular apertures (1.5 deg radius) distributed across 32 locations in the visual field. Stimuli contained both cues, but moved at either slow (0.25 deg/sec) or fast (4.0

deg/sec) speeds. These speeds primarily targeted CD and IOVD mechanisms respectively. We conducted multiple behavioral sessions with a group of typical participants. Our results confirmed that around 50% of the participants experienced impairment in stereomotion processing. For fast motion, the location of impairment varied between participants but remained stable across sessions. For slow motion, while impairments were also observed, their location was not stable over time. Furthermore, we did not find a correlation in the distribution of perceptual impairment between the two speeds. Varying the speed of stimuli containing both CD and IOVD cues supports prior work suggesting that stereomotion scotomas are based on an impairment in velocity-based mechanisms underlying motion-in-depth perception.

NYUAD Center for Brain and Health, funded by Tamkeen under grant CG012, ASPIRE Precision Medicine Research Institute Abu Dhabi (ASPIREPMRIAD) award grant number VRI-20-10.

26.459 REVERSED DEPTH ILLUSION IN RANDOM-DOT STEREOGRAMS BECOMES MORE VISIBLE WHEN THE STEREOGRAMS ARE MORE DYNAMIC IN BOTH CENTRAL AND PERIPHERAL VISION

Li Zhaoping[1] (li.zhaoping@tuebingen.mpg.de); [1]University of Tuebingen, Max Planck Institute for Biological Cybernetics

In anticorrelated random-dot stereograms (RDSs), a black dot in one eye corresponds to a white dot in the other eye to depict visual depth of object surfaces. Neurons in the primary visual cortex (V1) respond as if the depth order between surfaces is reversed from that defined by binocular disparities (Cummings & Parker 1997). The resulting reversed depth illusion is visible in peripheral but not in central vision (Zhaoping & Ackermann 2018). A Central-peripheral Dichotomy (CPD) theory explains this lack of central illusion as follows. Feedback from higher to lower visual areas to aid recognition mainly targets the central, rather than peripheral, visual fields. Such feedback verifies feedforward sensory signals, and thereby vetoes misleading V1 signals that cause the illusion. However, this illusion becomes visible in central vision using dynamic RDSs in which the random set of dots was replaced every 10 milliseconds by another randomly generated set while keeping the scene and stimulus design unchanged (Zhaoping ECVP 2021). The subsequent dots backward mask the previous dots, thereby compromising the feedback process that normally vetoes the illusion. In this study, we test whether this illusion also becomes more likely perceived in peripheral vision (at about 10 degrees eccentricity in the lower visual field) when the RDSs are made dynamic, and examine whether this illusion is more visible when the RDS becomes more dynamic, i.e., when the duration (e.g., 10 or 20 milliseconds) of each random set of dots becomes shorter. We report that, while this illusion remains more visible in peripheral vision, this visibility is enhanced in both central and peripheral vision when the RDSs are more dynamic. These findings suggest that the feedback verification is also present, albeit weaker, in peripheral vision. Whether backward masking enhances the reversed depth illusion more in central than peripheral vision will be reported.

26.460 OCULOMOTOR CONTROL IN FINE SHAPE AND STEREO JUDGEMENTS DURING NATURAL HEAD MOVEMENTS

Michele A. Cox[1] (michele.a.cox@gmail.com), Ashley M. Clark[1], Paul Jolly[1], Sanjana Kapisthalam[1], Yuanhao H. Li[1], Ruitao Lin[1], Soma Mizobuchi[1], T. Scott Murdison[2], Alina Neverodska[1], Jie Z. Wang[1], Zhetuo Zhao[1], Michele Rucci[1]; [1]University of Rochester, [2]Reality Labs

Recent studies have shown that humans exhibit remarkable precision in controlling eye movements during fixation. However, because of technical challenges in measuring eye movements, most of these studies were conducted as observers examined stimuli rendered on fixed-distance displays with their heads immobilized. Thus, little is known about the precision of natural, head-free fixation on real, three-dimensional objects. To overcome these limitations, we developed a system capable of measuring eye movements at high resolution during normal head movements. This device consists of a specifically-designed magnetic induction eye-tracker integrated with a motion capture system, together enabling simultaneous measurements of head and eye movements with arcminute (1/60th degree) resolution. Using this apparatus, we examined oculomotor control during fine discrimination of real objects at nearby working distances (30-60 cm). Targets were designed for either shape or a depth judgement and 3D printed from parametric models to scale angularly (1 degree) with viewing distance. The shape judgement consisted of reporting the position (left or right) of a 0.5 arcminute gap in a Landolt C. The depth judgement consisted of reporting whether the central prong of a 3-pronged fork was closer or farther than the two flankers (0.5 arcminute disparity; horizontal prong spacing and diameter 0.2 degrees). Normal-sighted observers (N=10) performed these tasks while seated and were allowed to freely move their heads. Trials were self-paced and ended with verbal reports. Our data show that head-eye control differs in shape and depth judgements. Head rotations and translations were attenuated and fixations considerably longer in the stereo judgements. The eye-in-head speed of intersaccadic fixation was also reduced during the stereo judgement, an effect that persisted even after controlling for head motion differences across tasks. Computational modeling is used to explore the benefits of longer, slower fixations for fine depth discrimination.

Supported by Reality Labs and NIH grants R01EY018363 and P30EY001319.

SATURDAY, MAY 18, 2:45 – 6:45 PM, PAVILION

Scene Perception: Neural mechanisms

26.461 AN INTRACRANIAL EEG NATURAL SCENES DATASET TO INTEGRATE ELECTROPHYSIOLOGY WITH FMRI

Harvey Huang[1] (huang.harvey@mayo.edu), Zeeshan Qadir[1], Morgan Montoya[1], Michael Jensen[1], Gabriela Ojeda Valencia[1], Kai Miller[1], Gregory Worrell[1], Kendrick Kay[2], Dora Hermes[1]; [1]Mayo Clinic, [2]University of Minnesota

Human brain networks involved in visual perception have been extensively investigated with fMRI studies such as the Natural Scenes Dataset (NSD), in which 8 healthy participants viewed thousands of naturalistic images. However, spectro-temporal characteristics are limited in fMRI. In contrast, intracranial EEG (iEEG) boasts sub-millisecond resolution and meaningful spectral power changes, but is limited by sparse spatial sampling. By collecting a large, high-quality iEEG dataset with matching stimuli to an fMRI dataset, we enable integration of the two modalities to leverage benefits of both. We recorded iEEG data in 12 human participants implanted for epilepsy monitoring, while they viewed the same 1000 stimuli presented to the NSD-fMRI participants. At each electrode, broadband power was calculated as a measure of local neuronal activity. Electrodes with significant signal-to-noise ratio across 6 repetitions of a 100-stimuli subset were considered visually responsive. Of 1650 total electrodes, 92 were visually responsive: 16 early visual, 10 lateral occipital, 5 dorsal occipital, 31 ventral temporal, 6 lateral temporal, 2 temporal pole, 18 frontal, 2 cingulate, and 2 amygdala. To integrate the iEEG and fMRI datasets, broadband power at each electrode was correlated with fMRI beta weights at all vertices across stimuli, yielding brain-wide correlation maps. 73 electrodes exhibited significantly positive correlations with fMRI vertices within a 3 mm radius, indicating robust local correspondence. Globally, most electrodes showed one of five spatial correlation patterns, which each consisted of positive and negative correlations across early visual areas and ventral, lateral, and dorsal streams; and which showed preference for contrast, places, faces, food, and bodies. This novel dataset integrates iEEG and fMRI on the same set of naturalistic stimuli. The recorded iEEG activity correlates well with fMRI and reveals global patterns of stimulus preference. This dataset opens the door to time-frequency analyses that can elucidate high-resolution dynamics within these networks.

We thank the patients in this study for their participation, Cindy Nelson and Karla Crockett for their assistance, and Peter Brunner for support with BCI2000. Research reported in this publication was supported by the NEI (R01EY035533, R01EY023384) and the NIGMS (T32GM145408).

26.462 HOW DO VISUAL TASKS ALTER THE REPRESENTATIONAL SPACE OF IDENTICAL SCENES? INSIGHTS FROM A BRAIN-SUPERVISED CONVOLUTIONAL NEURAL NETWORK

Bruce C. Hansen[1] (bchansen@colgate.edu), Henry A.S. Lewinsohn[1], Michelle R. Greene[2]; [1]Colgate University, [2]Barnard College

The neural representation of visual information is not a static pattern, but instead undergoes multiple transformations over time (Hansen et al., 2021), and supports different feature use with differing task demands (Greene & Hansen, 2020). However, exactly how task-relevant information is built up and subsequently used by the observer is only vaguely understood. To model that process, we constructed a novel convolutional neural network (CNN) where the convolutional layers were independently supervised by EEG responses at different time points. The CNN's goal was to use image information evaluated against neural responses to differentiate between two different tasks performed on identical real-world scenes. Participants (n = 24) viewed repeated presentations of 80 scenes while making cued assessments about either the presence of an object in the scene, or whether the scene afforded the ability to perform a function. Neural data were gathered via 128-channel EEG in a standard visual evoked potential (VEP) paradigm. Deconvolution was used to back-project onto image space activations across the layers of our brain-supervised CNN to reveal how the neural responses guided the differentiation of identical scenes at different image locations. The distribution of local activations was then compared to behavioral assessments of task-relevant information at each image location obtained through a crowd-sourced experiment. The behavioral data showed that the central region of image space was frequently informative for the object task, with the ground plane being most often informative for the function task. Crucially, our brain-supervised CNN used those task-relevant regions more to differentiate between identical sets of stimuli at ~70ms and ~250ms. Interestingly, the brain-supervised CNN made differential use of the task-relevant information within the early and late time points, suggesting a two-stage analysis of behaviorally-relevant scene locations. Our findings suggest that the observer's task-specific engagement with visual information substantially alters early neural representations.

James S. McDonnell Foundation grant (220020430) to BCH; National Science Foundation grant (1736394) to BCH and MRG.

26.463 LINGUISTIC AND VISUAL SIMILARITY JUDGEMENTS PREDICT EEG REPRESENTATIONAL DYNAMICS IN VISUAL PERCEPTION AND SENTENCE READING

Katerina Marie Simkova[1] (kms863@student.bham.ac.uk), Jasper JF van den Bosch[2], Clayton Hickey[1], Ian Charest[3]; [1]CHBH, School of Psychology, University of Birmingham, [2]School of Psychology, University of Leeds, [3]cerebrUM, Département de Psychologie, Université de Montréal

Emerging evidence in cognitive and computational neuroscience suggests that multi-modal computational models converge on representations that improve the performance in each of the modalities used. This latent representational space also enables the prediction of brain response profiles across modalities but it remains unclear how vision and linguistics share meaningful representations in the human brain. Here, we collected ~7 hours of electroencephalography (EEG) data from each of six participants passively viewing 100 natural scene images or actively reading 100 sentence captions describing the images. The activity pattern similarity was estimated using a cross-validated Mahalanobis distance computed on a spatiotemporal transformation of the modality-specific EEG data across all pairs of conditions. To establish the presence of shared representations in both modalities and to assess their behavioural relevance, we collected behavioural similarity judgements through multiple arrangement (MA) tasks on the set of images and sentences from two independent groups of participants (n = [24, 22]). This was used to construct the visual and linguistic fixed model RDMs each characterising the unique similarity structure of the two modalities. We then quantified the extent to which the behavioural model RDMs generalise to the visual and linguistic EEG RDMs using cosine similarity. We observed a significant relationship between the visually evoked EEG RDMs and both MA models (visual MA: 0.138 ± 0.024; linguistic MA: 0.136 ± 0.024). Interestingly, both MA models also revealed significant overlap with the linguistic EEG RDMs (visual MA: 0.042 ± 0.007; linguistic MA: 0.043 ± 0.007, all $p < 0.001$). These results remain when controlling for the potential influence of prior exposure to the cross-modal stimuli. We demonstrate that a similar representation emerges regardless of whether participants viewed an image or read its sentence caption and provide further evidence for behaviourally relevant shared representations in vision and language.

26.464 MAPPING CONTOUR PROPERTIES ACROSS VISUAL CORTEX

Seohee Han[1] (seohee.han@mail.utoronto.ca), Dirk B. Walther[2]; [1]University of Toronto

Detecting and integrating contours that delineate the boundaries of objects, surfaces, and other scene elements is a crucial function of what is loosely called "mid-level vision." How are such contours and their properties processed by the brain? We here explore this question using the high-resolution Natural Scenes Dataset (Allen et al, 2021). We analyzed the BOLD activity related to eight participants viewing subsets of 73,000 images of objects and scenes within V1, V2, V3, and hV4. Using the population receptive fields of individual voxels, we sample contour properties in a spatially specific manner to construct individual regressors for each voxel. This technique, first described by Roth et al. (2022), allows us to determine to what extent voxels within the visual cortex contribute to the representation of contour properties within their receptive field across thousands of images. When analyzing the salient contours in the images, we find a strong preference for horizontal orientations, consistent with the importance for scene layout, such as the horizon line. Interestingly, this finding contrasts with a similar analysis that relies on analyzing the orientation-specific Fourier energy in the photographs, which showed a primarily radial organization of orientation preference across the visual field. We present direct comparisons of the two methods. The technique of sampling contour properties with spatial specificity opens the door to exploring a range of other contour properties, such as contour curvature, contour junctions, as well as relationships between contours, such as parallelism and symmetry. Observing the neural representations of these properties and relationships across visual regions will bring us closer to a mechanistic understanding of how our perceptual information is organized in mid-level vision.

26.465 MAPPING NEURAL SIMILARITY SPACES FOR SCENES WITH GENERATIVE ADVERSARIAL NETWORKS

Gaeun Son[1] (gaeun.son@mail.utoronto.ca), Dirk B. Walther[1], Michael L. Mack[1]; [1]University of Toronto

Recent progress in vision science has focused on characterizing how the perceptual similarity of visual stimuli is reflected in the similarity of neural representations. While such neural similarity spaces are well-established in simple feature domains (e.g., orientation columns in V1), a correspondent finding with complex real-world stimuli has yet to be demonstrated. We explored this topic using scene wheels (Son et al. 2021), a continuous scene stimulus space created with scene-trained generative adversarial networks. We generated four different indoor scene wheels in which various global scene properties – spatial layouts, surface texture, component objects, colour schemes, etc. – changed gradually along a circular continuum. Participants were shown scene wheel images during fMRI scanning with a continuous carry-over design to provide stable estimates of scene-specific neural patterns. After scanning, participants rated pairwise perceptual similarity for the same scene wheel images. We also computed two types of physical similarity measures, the angular distances of the images on the scene wheels and their pixel-wise correlation, as well as a semantic similarity measure which indicated scene category as determined by a classifier network. We performed representational similarity analysis by comparing the similarity of scene-specific voxel patterns across multiple high-level visual regions as measures of physical, perceptual, and semantic similarity. We found that for scene wheels constrained to a single scene category (e.g., dining room), the neural patterns in visual cortex mainly represented the physical similarity of the scenes. However, when the scene wheels contain notable category boundaries (e.g., dining rooms and living rooms), both perceptual and category similarity structures were present in neural pattern similarity. These results provide important evidence that similarity structures defined by the complex feature spaces of real-world scenes are coded in neural representations and that such neural representations flexibly code for physical, perceptual, and categorical information.

This work was supported by an NSERC Discovery Grant (RGPIN-2020-04097 to DBW and RGPIN-2017-06753 to MLM) and Brain Canada (Future Leaders in Canadian Brain Research Grant to MLM).

26.466 NEURAL INTERPOLATION OF DYNAMIC VISUAL INFORMATION IN NATURAL SCENES

Lu-Chun Yeh[1] (luchun.yeh@gmail.com), Max Bardelang[1], Daniel Kaiser[1,2]; [1]Mathematical Institute, Department of Mathematics and Computer Science, Physics, Geography, Justus Liebig University Gießen, [2]Center for Mind, Brain and Behavior (CMBB), Philipps University Marburg and Justus Liebig University Gießen

Adaptive natural vision requires our brain to interpolate missing information about occluded objects in the environment. Previous studies suggest that this process is supported by visual cortex „filling in" occluded parts of scene images. However, we live in a dynamic world, and objects keep moving in and out of occlusion (e.g., when trains move through tunnels). Here, we used multivariate pattern analysis on time-frequency-resolved EEG data to track neural representations during dynamic occlusion. Participants watched 4-second videos of a person walking across a scene (either left-to-right or right-to-left) while performing an unrelated fixation task. The videos featured three conditions: The person walking across a blank background (isolated condition), across the scene without occlusion (visible condition), or across the scene while being dynamically occluded between 1.5 and 3 seconds (occluded condition). We trained linear classifiers on EEG response patterns to discriminate rightward- and leftward-walking in the isolated condition and tested them on the visible and occluded conditions. Classifiers trained on time-locked broadband responses, as well as on alpha (8-12Hz) and beta (13-30Hz) rhythms, successfully discriminated walking direction in the visible condition. However, only classifiers trained on alpha rhythms could discriminate walking direction in the occluded condition. Critically, we introduced an additional condition during which the person stopped in front of a natural obstacle (e.g., a river). We found that alpha dynamics tracked the termination of motion in this condition, even when it was hidden by the occluder. Together, our results provide evidence for an automatic interpolation of information during dynamic occlusion. The alpha dynamics that mediate this interpolation may constitute a neural correlate of top-down processes that „fill in" missing information based on context.

26.467 OBJECT-SELECTIVE CORTEX INCORPORATES PREDICTIONS FROM SCENE CONTEXT TO AID OBJECT RECOGNITION

Surya Gayet[1,2] (surya.gayet@gmail.com), Marco Gandolfo[2], Marius Peelen[2]; [1]Utrecht University, Helmholtz Institute, [2]Radboud University, Donders Institute, [3]Radboud University, Donders Institute

Our ability to recognize objects is strongly facilitated by scene context. Scene context does not only predict the identity of an object (a road is likely to contain a car), but also its appearance (e.g., a car is about as wide as the lane on which it's driving). Using a multi-method approach, we investigated how object processing is influenced by the predicted (retinal) size of an object, as inferred from the current viewing distance. To this end, we created a large set of outdoor scenes, each containing one object, positioned relatively nearby in one condition (thus producing a relatively large retinal image) and relatively far away in the other (producing a relatively small retinal image). For each scene, we additionally created two conditions with coherently sized objects, by swapping the positions of the 'large' and 'small' objects in the scene. A series of behavioral experiments showed that coherently sized objects are more easily recognized than incoherently sized objects. This recognition advantage was severely reduced when removing global scene information. Using fMRI searchlight analyses, we found that activity patterns evoked by different object categories were more dissociable (by a classifier trained on isolated objects) when objects were coherently compared to incoherently sized. This sharpened representation of size-coherent objects was specific to object-selective cortices, lateral occipital complex (LOC) and posterior fusiform sulcus (pFs), and predicted the magnitude of the behavioral recognition advantage (on an image-by-image basis) in the behavioral experiment. Using online chronometric TMS, we found that stimulating LOC from 160ms after scene onset reduced this behavioral effect to baseline. This establishes a causal role of object-selective cortex in instantiating this viewpoint-dependent behavioral recognition advantage. We conclude that the predicted real-world size of an object, as inferred from scene context, contributes to object recognition and sharpens object representations in object-selective visual cortex.

This project received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement no. 725970, granted to M.V.P.) and from the Netherlands Organisation for Scientific Research (VI.Veni.191G.085, granted to S.G.).

26.468 OCCIPITO-VENTRAL PATHWAY DYNAMICALLY TRANSFORMS IMAGES INTO LOW-DIMENSIONAL FEATURE MANIFOLDS ALIGNED ON THOSE SUPPORTING BEHAVIOR

Yaocong Duan[1] (y.duan.1@research.gla.ac.uk), Robin Ince[1], Joachim Gross[2], Philippe Schyns[1]; [1]University of Glasgow, [2]University of Muenster

Theories and models of visual categorization suggest that the brain must actively transform its representations of complex input scenes into the low-dimensional manifolds that represent the task. With its involvements in multiple face, object and scene categorizations, the occipito-ventral pathway should be pivotal to such task-dependent representational transformations. However, previous studies using full images didn't allow us to observe these transformations. Here, we tracked them directly in the brain. We used an experiment where participants (N=10) performed four different 2-Alternative-Forced-Choice categorizations of the same 64 base images of a realistic city scene in different blocks of 1,536 trials. These images comprised varying embedded targets (8 face identities representing 2 genders x 2 expressions x 2 vehicles x 2 pedestrian). Each trial started with a fixation cross, followed by one base image for 150ms, whose pixels were randomly sampled with the Bubbles procedure. We concurrently recorded participants' categorization responses and source-localized MEG activity. First, we determined the features each participant used for behavior in each task–computing Mutual Information(Pixel visibility; Correct vs. Incorrect). We also reconstructed their dynamic representations of each image pixel on each MEG source–computing MI(Pixel visibility; MEGt source amplitude). We then tracked the representational transformations across the occipito-ventral pathway layers. In each participant and task, we discovered that 80 and 150ms post-stimulus, the broad initial representation of image pixels in occipital cortex progressively transforms across ventral pathway layers into the low-dimensional task-specific feature manifolds (whose contents align with those supporting task-behavior). Occipital cortex also reduces task-irrelevant pixels (until 150ms), when the ventral pathway has identified the task-relevant feature manifolds. Our findings offer new insights into how the occipito-ventral pathways dynamically aligns its features to those of behavior in multiple face, object and scene categorization tasks.

P.G.S. received support from the Wellcome Trust (Senior Investigator Award, UK; 107802) and the MURI/Engineering and Physical Sciences Research Council (USA, UK; 172046-01). RI was supported by the Wellcome Trust [214120/Z/18/Z].

26.469 RECOGNIZING PLACES VERSUS NAVIGATING THROUGH THEM ARE DIFFERENTLY SENSITIVE TO INCREASINGLY PERIPHERAL VISUAL INFORMATION

Yaelan Jung[1] (jung.yaelan@gmail.com), Andrew Persichetti[2], Daniel Dilks[1]; [1]Emory University, [2]NIH

It is well-documented that the two visual streams for object processing (i.e., vision-for-perception and vision-for-action) are differently sensitive to increasingly peripheral stimuli, with peripheral visual information especially important for the vision-for-action system. Here we hypothesize the same is true for scene processing, with peripheral visual information especially important for "visually-guided navigation" (like vision-for-action) compared to "scene categorization" (like vision-for-perception). We directly tested this hypothesis in two ways. First, using eye tracking, we monitored the eye movements of adult participants while they performed a visually-guided navigation task and a scene categorization task. Consistent with our hypothesis, we found that participants primarily used peripheral vision during the visually-guided navigation task and central vision during the scene categorization task. Second, using resting-state fMRI, we investigated the functional connectivity between the foveal to increasingly more peripheral regions of the primary visual cortex (V1) and two scene-selective regions – the occipital place area (OPA), which is implicated in visually-guided navigation, and the parahippocampal place area (PPA), which is implicated in scene categorization. Again, consistent with our hypothesis, we found that the OPA shows an exclusive connection to peripheral V1, while the PPA exhibits a gradient of connectivity from foveal to peripheral V1. Taken together, these findings provide converging behavioral and neural evidence that recognizing places versus navigating through them are differently sensitive to increasingly peripheral stimuli, with peripheral visual information especially important for navigating through a place.

26.470 REPRESENTATION OF NAVIGATIONAL AFFORDANCES AND EGO-MOTION IN THE OCCIPITAL PLACE AREA

Frederik S. Kamps[1] (fkamps@mit.edu), Emily M. Chen[1], Nancy Kanwisher[1], Rebecca Saxe[1]; [1]Massachusetts Institute of Technology

Humans effortlessly use vision to plan and guide navigation through the local environment, or "scene". A network of three cortical regions respond selectively to visual scene information, including the occipital place area (OPA), parahippocampal place area (PPA), and retrosplenial complex (RSC) – but how this network supports visually-guided navigation is unclear. Recent evidence suggests that one region in particular, the OPA, supports visual representations for navigation, while PPA and RSC support other aspects of scene processing. However, most previous studies tested only static scene images which lack the dynamic experience of navigating through scenes. Accordingly, here we used dynamic movie stimuli to test whether OPA, PPA, and RSC represent two critical kinds of navigationally-relevant information: navigational affordances (e.g., can I walk to the left or right?) and ego-motion (e.g., am I turning left or right?). We found that OPA is sensitive to both affordances and ego-motion, as well as the conflict between these cues – e.g., turning toward versus away from an open doorway. These effects were significantly weaker or absent in PPA and RSC. Responses in OPA were also dissociable from those in early visual cortex, consistent with the idea that OPA responses are not merely explained by lower-level visual features. OPA responses to affordances and ego-motion were stronger in the contralateral than ipsilateral visual field, suggesting that OPA encodes navigationally relevant information within an egocentric reference frame. Taken together, these results support the hypothesis that OPA contains visual representations that are useful for planning and guiding navigation through scenes.

This work was supported by NIH grant 1R01HD103847-01A1 (awarded to RS)

26.471 UNVEILING TASK-DEPENDENT ACTION AFFORDANCE REPRESENTATIONS: INSIGHTS FROM SCENE-SELECTIVE CORTEX AND DEEP NEURAL NETWORKS

Clemens G. Bartnik[1] (c.g.bartnik@uva.nl), Nikolina Vukšić[2], Steven Bommer[2], Iris I.A. Groen[1,2]; [1]Video & Image Sense Lab, Informatics Institute, University of Amsterdam, The Netherlands, [2]Psychology Research Institute, University of Amsterdam Amsterdam, The Netherlands

Humans effortlessly know how and where to move in the immediate environment using a wide range of navigational actions, from walking and driving to climbing. Yet little is known about where and how action affordances are computed in the brain. Some work implicates scene-selective cortex in navigational affordance representation, reflecting visual features computed in mid-level DNN layers (Bonner et al., 2017, 2018), while others report a lack of affordance representation therein (Groen et al., 2018). Here, we curated a novel set of real-world scenes that afford distinct navigational actions in both indoor and outdoor environments, for which we collected rich behavioral annotations (N=152) for seven commonly used visual properties. The behavioral annotations indicate that navigational actions form a distinct space separate from representations of objects or materials; even in combination, visual properties explain only around 20% of the variance in navigational action annotations. We collected human fMRI measurements (N=20) to a subset of 90 images while subjects performed three distinct tasks (action affordance recognition, object recognition, and fixation). Using representational similarity analysis, we confirm that scene-selective brain regions, especially the Parahippocampal Place Area and Occipital Place Area, represent navigational action affordances. Furthermore, elevated behavioral correlations in scene-selective regions during action affordance and object recognition tasks relative to fixation suggests these representations are task-dependent. Unlike prior findings, however, we find that DNNs trained for scene and object classification poorly represent these action affordances. Interestingly, language-supervised models like Contrastive Language-Image Pre-training (CLIP) show enhanced predictions for behavior and brain activity, suggesting they capture affordance representation better. These findings strengthen evidence for action affordances in the scene-selective cortex and reveal their task dependency. However, the underlying computations remain elusive, but our work suggests that integrating semantic information in computational models of affordance perception is a promising direction.

SUNDAY MORNING POSTERS IN BANYAN BREEZEWAY

Sunday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

Plasticity and Learning: Electrophysiology, brain stimulation

33.301 VARIATION IN CORTICAL RESPONSES TO NEUROMODULATION: MOTOR THRESHOLDS VS. VISUAL PHOSPHENES

Remy Cohan[1,2,3,4] (rcohan@yorku.ca), Jennifer K.E. Steeves[1,2,3,4]; [1]Centre for Integrative and Applied Neuroscience, [2]Centre for Vision Research, [3]Department of Psychology, [4]York University, Toronto, Canada.

Theta burst stimulation (TBS) is a repetitive transcranial magnetic stimulation (rTMS) protocol with the added benefit of shorter stimulation time (compared to traditional rTMS protocols). This increases efficiency and compliance in both research and clinical settings. Even though TBS applications in the primary motor cortex (M1) have been well-explored, scant attention has been paid to its impact on the primary visual cortex (V1). TBS includes two variants: Intermittent (iTBS; excitatory) and continuous (cTBS; inhibitory). In our lab, previous neuroimaging studies including magnetic resonance spectroscopy and resting state functional magnetic resonance imaging did not lead to notable alterations in GABA concentration in V1 or functional connectivity when V1 was stimulated with either cTBS or iTBS. However, we previously found significant effects when applying 1Hz rTMS to V1. This emphasises the need to understand TBS effects on visual brain areas and optimise TBS protocols accordingly. In this study, we explored cTBS' influence on M1 and V1 utilising stereotactic neuronavigation. We investigated the disparities in biophysical parameters such as scalp-to-cortex distance (SCD), electric fields at hotspots and stimulation intensities between M1 and V1. We used motor thresholds (MTs) for M1 stimulation and phosphene thresholds (PTs) for V1 stimulation as markers. Our preliminary results suggest that PTs are significantly higher than MTs despite longer SCD at V1, and shorter SCD at M1. In addition, we found that post-cTBS PTs increased in comparison to sham stimulation, hinting at an inhibitory aftereffect. However, both active and sham cTBS in M1 yielded to increased MTs. A profound understanding of the varied stimulation parameter effects and their location-specific variability is crucial for effective and efficient TMS protocol application in research and therapeutic contexts. These insights will facilitate tailoring TMS approaches to each target region's unique requirements, enhancing the overall effectiveness of TBS in vision research and clinical settings.

This research was funded by the Natural Science And Engineering Research Council of Canada, Canada First Research Excellence Fund (CFREF), Vision Science to Applications (VISTA) .

33.302 NAP AFTER ANODAL TRANSCRANIAL DIRECT CURRENT STIMULATION (TDCS) DISRUPTS CONSOLIDATION OF VISUAL PERCEPTUAL LEARNING- AN EEG STUDY

Xinyi Zhu[1,2,3,4] (anguszhu324@stu.pku.edu.cn), Qing He[5], Fang Fang[1,2,3,4]; [1]School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, 100871, Beijing, China, [2]Key Laboratory of Machine Perception, Ministry of Education, Peking University, 100871, Beijing, China, [3]IDG/McGovern Institute for Brain Research, Peking University, 100871, Beijing, China, [4]Peking-Tsinghua Center for Life Sciences, Peking University, 100871, Beijing, China, [5]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

In visual perceptual learning (VPL), consolidation is a crucial process that stabilizes and even further improves performance on trained tasks without additional practice. Our previous study found that sleep after transcranial direct current stimulation (tDCS) disrupted VPL consolidation, yet the underlying neural mechanisms of such phenomenon are still unclear. The current study incorporated an adaptation sleep and two intervention days containing formal sleep. On each intervention day, participants were trained on an orientation discrimination task (ODT) in the morning following resting-state EEG recording. In each ODT trial, two Gaussian-enveloped sinusoidal gratings (Gabor patch) with slightly different orientations were presented in random order in the lower left or right visual field (VF). The order of VFs was counterbalanced across subjects, and the orientations of Gabors in the two VFs were orthogonal. After training, 2.0 mA anodal tDCS was delivered over the occipital regions contralateral to the VF in training for 25 minutes for active and 30 seconds for sham conditions, respectively. Immediately after brain stimulation, resting-state EEG was recorded for a second time. In the afternoon, participants napped in the sleep room with polysomnography (PSG) recording for 60-90 minutes. A post-test was completed after the nap. Results showed that, following a nap, participants' orientation discrimination thresholds declined in the sham condition but not in the active tDCS condition. Moreover, in N2 sleep, delta power in the tDCS-stimulated occipital area was lower in the active condition compared to that in the sham condition, and the sleep spindle power in the frontal area contralateral to the stimulated hemisphere was also lower in the active condition. Our findings indicate that the consolidation of VPL can be disrupted by the interplay of post-training anodal tDCS and subsequent nap, suggesting that anodal tDCS-induced hyper-excited brain activity might disturb the system consolidation process during sleep.

33.303 STRONGER ADAPTATION OF MIDDLE-TO-LATE ERP COMPONENTS TO OBJECT SILHOUETTE IMAGES BEFORE VERSUS AFTER OBJECT PRIMING IN APHANTASIA

NITHIT SINGTOKUM[1] (nithitsingtokum136@gmail.com), Kanyarat Benjasupawan[1], Chattarin Poungtubtim[1], Sedthapong Chunamchai[2], Chaipat Chunharas[2]; [1]Faculty of medicine chulalongkorn university, Bangkok, Thailand, [2]Cognitive Clinical and Computational Neuroscience Lab, KCMH Chula Neuroscience Center, Thai Red Cross society, Chulalongkorn University, Bangkok, Thailand

Visual imagery is the ability to reactivate and manipulate visual representation in the absence of visual stimuli. This ability, however, cannot be accomplished by people with aphantasia. Previous research has indicated that the intensity of visual imagery aligns with the activation of distinct brain regions, such as the fusiform area. However, it remains uncertain whether this process is influenced solely by voluntary attempts at imagination by aphantasics or if it can also occur involuntarily. The patient NP abruptly lost her ability to "picture things in her mind". The results of detailed clinical neuropsychological tests were normal except that she could not identify objects from their silhouette images especially when they were presented from unusual viewpoints. To explore the involuntary nature of this process, we conducted an EEG experiment where Participants had to monitor a sequence of images on a screen and click a button when an image tilted clockwise or counterclockwise (occurring only 10% of the time). Unrelated to the task, stimuli were intentionally presented as triplets: a silhouette image (pre-silhouette), followed by the corresponding object image, and then the same silhouette image as the pre-silhouette (post-silhouette). Greater ERP differences between post-silhouette and pre-silhouette images would suggest the advantage of image preview or the absence of object presentation in the pre-silhouette condition. Our findings revealed that patient NP exhibited significantly higher average ERP changes between 300 and 800 ms after stimulus onset compared to control groups in occipital ($p<0.001$, t-value = 5.25), frontal ($p<0.001$, t-value = 5.87), and parietal ($p<0.01$, t-value = 3.68) areas. The increased adaptation after an image preview, even when object recognition was unrelated to the task, implies a potential impairment in the top-down signal that facilitates visual imagery in individuals with aphantasia, and notably, this impairment may occur involuntarily.

33.304 EEG MEASURES OF CONSOLIDATION OF VISUAL PERCEPTUAL LEARNING DURING WAKEFULNESS ARE DISTINCT FROM CONSOLIDATION DURING SLEEP

Aaron Cochrane[1] (aaron_cochrane@brown.edu), Theodore LaBonte-Clark[1], Kiley Haberkorn[1], Takeo Watanabe[1], Yuka Sasaki[1]; [1]Brown University

Visual perception is malleable in response to training, called visual perceptual learning (VPL). Thus, improvements on one VPL task must be consolidated to resist retrograde interference from subsequent training. Previous findings have shown that, during post-training REM sleep, the power of EEG theta-band oscillations, as well as a lower excitation-dominance between excitatory and inhibitory neurotransmitters (E/I), were each reliably related to learners' abilities to resist such retrograde interference. Notably, while consolidation of VPL also occurs during wakefulness, the underlying neural mechanisms of the consolidation of VPL during wakefulness have yet to be fully characterized. Here, we tested whether theta power and EEG-based E/I balance similarly indexed increased resistance to interference in awake humans. Participants were trained with one texture discrimination task (TDT) set and rested for one hour (post-training rest). During the rest period, posterior EEG was recorded. Training on a second set of TDT was then completed, which has the potential to interfere with the learning from the first TDT. While previously reported results linked higher interference to lower theta power during post-training REM sleep, we did not observe evidence for such a relationship in theta power. Instead, we found a similar link between higher interference and lower alpha power during post-training rest (contralateral to trained hemifield; ρ = .45). While the lower E/I balance measured by neurotransmitters was linked to increased resistance to interference during post-training REM sleep previously, the lower E/I balance measured by EEG was linked to increased susceptibility to interference during an awake post-training rest period (theta-band ρ = .47; alpha-band ρ = .36). These suggest that while consolidation during REM sleep is associated with inhibition, consolidation during wakefulness is associated with excitation. Neural mechanisms for the consolidation of VPL are thus likely to differ between wakefulness and sleep.

Acknowledgements: NIH (R01EY031705, R01EY019466, R01EY027841)

33.305 LTP-LIKE ACTIVITY INDUCED BY POST-TRAINING RHYTHMIC FLICKER CONSOLIDATES VISUAL PERCEPTUAL LEARNING

Xin-Yue Yang[1-4], Qing He[5], Fang Fang[1-4,]; [1]School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China, [2]Key Laboratory of Machine Perception, Ministry of Education, Peking University, Beijing, China, [3]IDG/McGovern Institute for Brain Research, Peking University, Beijing, China, [4]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China, [5]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China*

Understanding how newly encoded information is consolidated in visual perceptual learning (VPL) is significant in contemporary vision science. Consolidation can be manifested in different forms at different stages of VPL, among which the early consolidation post-initial training remains largely uncharted, with numerous aspects still unknown regarding neural mechanisms and potential interventions. In this study, we investigated whether repetitive visual stimulation (RVS), known for inducing LTP- or LTD-like effects on neural activities and measures of behavioral performance in a frequency-dependent manner, could be used to promote early consolidation of VPL and investigated the neuroelectrical activities supporting such effects. Forty-five participants were equally assigned to three groups with 10-Hz, 1-Hz, and 0-Hz (static) flickers. Participants underwent training on an orientation discrimination task (ODT), an RVS session, a four-hour break, and a post-test on ODT. In an ODT trial, participants fixated on a central fixation point, two embedded-in-noise Gabors with slightly different orientations successively appeared in the periphery, and the participants reported the rotation (clockwise or counter-clockwise) with a keypress. In RVS, a sinusoidal grating flicker was flashed at the corresponding frequency in the same location as the Gabor in ODT. During all tasks, we continuously monitored eye movement using a video-based eye tracker and recorded electroencephalogram (EEG) using 64-channel electrode caps. Results showed a decline in discrimination threshold with training in all three groups, indicating that learning had occurred. Ten-Hz flicker resulted in decreased threshold in the post-test, while 1-Hz flicker increased the threshold, and static flicker yielded no effect. The 10-Hz flicker-induced improvement was accompanied by a stronger N145 component of visually evoked potential (VEP) recorded in the contralateral occipito-parietal cortex in the post-test. Our findings suggest that VPL-involved neural traces exhibit plasticity in early consolidation, and 10-Hz RVS induces LTP-like activity in the visual areas, facilitating early consolidation of VPL.

SUNDAY, MAY 19, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Decision Making: Perceptual decision making 2

33.306 DERIVING THE FUNCTIONAL FORM TO FIT CONFIDENCE RATINGS IN PSYCHOPHYSICAL EXPERIMENTS

Keith A. Schneider[1] (keithas@udel.edu); [1]University of Delaware

Introduction: Measuring subjects' confidence, or metacognition, permits insight into subjects' decision making processes. Subjects can report their confidence in each decision they make during a psychophysical experiment. Typically, researchers fit an arbitrary function, if anything, to describe these data. Here I describe a simple procedure to derive a functional form for confidence ratings, based on the psychometric curve. Methods: A subject's confidence during a psychophysical task should be proportional to the total amount of information available for each decision. If the psychometric function is known, fitting their decision data, then a subject's uncertainty is proportional to the slope of that function, plus any intrinsic information in the stimuli. For example, if a subject must discriminate between two stimuli of different contrasts, then the total amount of information would be proportional to the derivative of the psychometric function plus the sum of the log contrast of each stimulus; i.e., subjects are more confident for suprathreshold stimuli. Results: I demonstrate this using two experiments in which subjects discriminated between Gabors. In one experiment, subjects judged which stimulus had the higher contrast (comparative judgement), and in the second, subjects judged whether the two stimuli were equal (equality judgment). These decisions have different functional forms. The comparative judgement is monotonic, with a maximum slope at the point of subjective equality. The equality judgment is a negative convex function with an absolute maximum—therefore its derivative has two humps. In both experiments, a linear term fits the sum of the stimuli log contrasts. Conclusion: As an application, I show that, even when the psychometric functions are identical, subjects' uncertainty can be used to discriminate intrinsically biased decisions, of which subjects are unaware, from explicit bias, e.g. instructions to make a given choice unless subjects are certain that another choice is correct.

33.307 DISSECTING THE REACTION TIMES OF GLOBAL AND LOCAL PROCESSING

daniel lougen[1] (dlougen2@uoregon.edu), Steven Shofner[2], Paul Dassonvile[3]; [1]University of Oregon, [2]University of Oregon, [3]University of Oregon

Reaction time studies using hierarchical visual stimuli typically demonstrate that global processing is faster than local processing (global precedence effect), and that global interference during a local task has a greater impact than the opposite (global interference effect). Previous examinations of individual differences in these reaction times have sometimes provided evidence that these effects are modulated by various clinical conditions (e.g., autism), but these reports are inconsistent in the literature. This may be because reaction times are a blunt measure of processing speed, since they provide only a sum of the durations of the many different processes that occur between stimulus and response (including afferent transmission, decision, and efferent transmission), and they are prone to distortion due to speed-accuracy tradeoffs. Here, we used a compelled response task (Stanford et al., 2010) to tease apart the duration of decision processes from that of non-decision processes (the sum of afferent and efferent delays). Participants viewed hierarchical C's (large C's made of small c's, both of which could be open to the right or left in a mirror-reversed fashion) and reported their orientations at either the global or local level (in a blocked design) with right or left button presses. We found that although the global interference effect was reflected in almost equal measures in non-decision and decision times, the global precedence effect was mostly comprised of an increased duration for local decisions compared to global decisions with very little difference in the durations of non-decision processes. The ability of the compelled response task to provide more precise measures of the durations of the underlying components of a response indicates benefits of its use over typical reaction time measures for examining individual differences in sensorimotor processing.

33.308 INHOMOGENEITIES IN HUMAN RESPONSES TO ZERO-COHERENCE DOT MOTION PERSIST AS INCREASINGLY MORE SENSORY EVIDENCE IS ADDED

Kanathip Jongmekwamsuk[1,2] (kanathipjong@gmail.com), Chaipat Chunharas[2], Rosanne L. Rademaker[1]; [1]Rademaker Lab, Ernst Strüngmann Institute for Neuroscience in cooperation with Max Planck Society, Frankfurt, Germany, [2]Cognitive Clinical & Computational Neuroscience, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

The human brain combines prior knowledge with sensory information to make inferences about the environment and construct our perception of the world. To find out if people prefer certain responses over others when recalling a direction of motion, and to see if such inhomogeneities in response preference persist as increasingly more sensory evidence is provided, we had human participants report the direction of motion of moving dot stimuli shown for 1s at four possible levels of coherence (0%, 25%, 50%, and 75%). Across trials with coherent motion, we ensured uniform sampling of all possible directions of motion, which means that in the case of accurate recall (plus noise), we expect a uniform distribution over all possible response options. Similarly, on trials without coherent motion, random responses should also lead to a uniform response distribution. Our data revealed striking inhomogeneities in reported direction in the absence of coherent motion, with many more reports at cardinal and oblique directions compared to other directions. And while performance increased with higher coherence, with reports becoming more uniform across motion directions, clear non-uniformities persisted even on high coherence trials. Specifically, people continued to over-report oblique (but not cardinal) directions even for 75% coherent dot motion. This suggests an integration of existing inhomogeneities in response preference (revealed in the 0% coherence condition) and sensory information, where direction estimation is increasingly less affected by a "response prior" as sensory certainty goes up, but the weight of the response prior appears to differ between cardinals and obliques.

33.309 METACOGNITION IN PUTATIVE MAGNO- AND PARVOCELLULAR VISION

Vrishab Nukala[1] (vnukala@ucsc.edu), April Pilipenko[2], Jessica De La Torre[1], jason Samaha[1]; [1]University of California, Santa Cruz, [2]University of California, San Diego

A major distinction in early visual processing is the magnocellular (MC) and parvocellular (PC) pathways. The MC pathway tends to process motion, transient events, and low spatial frequencies, while the PC pathway mostly processes color, sustained events, and high spatial frequencies. Prior work suggests that the PC pathway more strongly contributes to conscious object recognition via projections to the ventral "what" visual pathway, whereas the MC pathway contributes to non-conscious motion and localization processing via the dorsal stream "where/how" pathway. This invites the question: Are we equally aware of activity in both pathways? And if not, do task demands interact with which pathway is more accessible to awareness? We investigated this question in a set of two studies measuring participant's metacognition (their ability to judge the accuracy of their own perceptual decisions) for stimuli biased towards MC or PC processing. The "Steady/Pulsed Paradigm" presents brief, near-threshold stimuli under two conditions thought to favor either pathway. In the "pulsed" condition, the target appears atop a brief luminance pedestal which theoretically saturates the transient MC response and leaves the PC pathway to process the stimulus. In the "steady" condition, the stimulus is identical except the luminance pedestal is constant throughout the trial, rather than flashed along with the target. This theoretically adapts the PC neurons and leaves MC for processing. Experiment 1 was a spatial localization task thought to rely on information relayed from the MC pathway. Using a model-based approach to quantify participants' metacognition about their task performance, we found higher metacognition in the steady (MC-biased) condition. Experiment 2 was an orientation-discrimination task more reliant on PC pathway information. Preliminary results show an abolishment of the MC pathway advantage seen in experiment 1 and suggests that the metacognitive advantage for MC processing may hold for stimulus localization tasks only. (edited)

33.310 METACOGNITION IS MENTALLY DEMANDING: REVEALING THE COSTS AND CONSEQUENCES OF METACOGNITIVE EFFORT

Julian Matthews[1] (julian.r.matthews@gmail.com), Atsushi Kikumoto[1,2], Kentaro Miyamoto[1], Kazuhisa Shibata[1]; [1]RIKEN Center for Brain Science, [2]Brown University

Metacognition, the capacity to monitor and control our own mental processes, is thought to automatically and effortlessly accompany decision-making. This assumption underlies all models of metacognition, which were developed using experiments involving hundreds of trials with no incentives to motivate metacognitive performance. However, evidence shows that many other cognitive functions are mentally demanding, and the costs of cognitive effort lead people to disengage in effortful tasks unless suitably motivated. Here, we propose that metacognition is mentally demanding, leading people to employ heuristics to avoid metacognitive effort. We develop a flexible, effort-based decision-making paradigm that allows participants to trade-off rewards for reduced metacognitive effort. We operationalise metacognitive effort as the precision of confidence judgments, participants can maintain more confidence criteria (higher precision) at the expense of mental exertion. Critically, our design incentivises metacognition while controlling for fatigue and the effortfulness of the primary task. Across three experiments, we demonstrate that individuals sacrifice rewards to avoid metacognitive effort. Then, an important question arises: is effort a practical concern for understanding metacognition? To illustrate the consequences of metacognitive effort we take a famous observation in metacognition research—confidence leaks (correlations in confidence ratings for independent tasks performed in close temporal proximity). Introducing incentives into a confidence leak paradigm, we show that confidence leaks are significantly reduced when metacognition is rewarded. This indicates that confidence leaks, previously considered intrinsic to metacognition, can emerge due to a lack of incentives for investing metacognitive effort. Our research shows that metacognition is sensitive to effort and value, challenging the assumption that it automatically accompanies decision-making. We demonstrate that metacognition is effortful and when effort is counteracted by reward, metacognitive inefficiencies are reduced. Our model proposes that these inefficiencies may be rational heuristics by a system that faces substantial computational demands with few incentives to exert metacognitive effort.

Japan Society for the Promotion of Science Postdoctoral Fellowship P23010

33.311 NEAR-OPTIMAL METACOGNITION ACROSS THE VISUAL PERIPHERY

Brian Odegaard[1] (bodegaard@ufl.edu), Joseph Pruitt[1]; [1]University of Florida

How effectively does perceptual confidence track task accuracy across the visual field? Previous research has provided an array of findings, including underconfidence in the visual periphery (Toscani, Mamassian, and Valsecchi, 2021), overconfidence in incorrect trials (Odegaard, Chang, Lau, and Cheung, 2018), and anecdotal accounts of being surprised at how poorly we perform visual tasks at eccentric locations (Cohen, Dennett, and Kanwisher, 2016). In this investigation, we aimed to systematically characterize metacognitive efficiency for visual detection judgments up to forty degrees away from fixation. On each trial, twenty-four participants performed a 2AFC task to report signal (a Gabor) or noise at a specific location along a horizontal axis, and rate confidence in their judgment. On most trials, participants were cued to the location of the upcoming stimulus (valid trials), but on a small proportion of trials, participants were cued to an incorrect location (invalid trials). Results showed that, as expected, perceptual sensitivity declined with eccentricity, and valid trials had higher perceptual sensitivity than invalid trials. Importantly, metacognitive efficiency, as measured by the metric meta-d'/d', was nearly optimal at each eccentric location, showing that confidence judgments effectively tracked task accuracy. Additionally, we revealed a unique profile of decision criteria across eccentric locations as a “check-mark” pattern emerged, with very conservative criteria at the 40-degree location, and unbiased criteria closer to fixation. Decision criteria were minimally impacted by attentional cueing, and a second experiment without cues replicated the “check-mark” pattern across eccentric locations. Finally, results showed that declining metacognitive sensitivity in eccentric locations was driven primarily by declining confidence in correct trials, and surprisingly consistent confidence on incorrect trials across eccentricities. Together, these results challenge accounts of liberal criteria and suboptimal confidence in the periphery, and reveal that our sense of confidence may track performance decrements more effectively than previously supposed.

This work was supported by an Office of Naval Research Young Investigator Award (#N00014-22-1-2534)

33.312 THE EFFECTS OF SPATIOTEMPORAL UNCERTAINTY ON METACOGNITION IN ORIENTATION ENSEMBLE PERCEPTION

Alan L. F. Lee[1] (alanlee@ln.edu.hk), Gabriel Tam[1], Derek H. Arnold[2]; [1]Lingnan University, Hong Kong, [2]University of Queensland

Humans can extract statistical representations quickly over a set of elements that vary over space and time. Previous studies have investigated the computational processes underlying ensemble perception. Here, we investigated how observers make judgments about their own performance amid spatiotemporal uncertainty in an ensemble-perception task. We adapted the stimulus from Yashiro et al. (2020) and compared two types of ensembles with different spatiotemporal uncertainty: Temporally-smooth ensembles had orientations varied across locations, but with the same average orientation across frames (i.e., temporal SD = 0), while spatially-smooth ensembles had orientations varied across frames, but with the same orientation across locations in every frame (i.e., spatial SD = 0). To obtain a performance-matched confidence comparison between the two ensemble types, we adopted the confidence forced-choice paradigm (Barthelmé & Mamassian, 2010; Knotts et al., 2018). On each trial, participants performed a clockwise-vs-counterclockwise orientation-discrimination task on the average orientation of the two ensemble types, one after the other in randomized order, and then chose the response in which they were more confident in being correct. We manipulated task difficulty by varying the ensemble's spatiotemporal average orientation relative to the vertical reference. In Experiment 1, when locations of elements varied from frame to frame for both ensemble types, participants chose spatially-smooth ensembles more frequently than temporally-smooth ensembles even when performance was matched. This systematic metacognitive bias was eliminated in Experiment 2 when element locations were kept constant both across frames and between the two ensemble types. Furthermore, we replicated the "recency effect" (Yashiro et al., 2020) on confidence judgments in that later frames in an ensemble had greater influence on confidence choices than earlier ones. Our findings suggest that the structure of spatiotemporal uncertainty of a stimulus ensemble could influence not only first-order performance, but also second-order metacognitive judgments.

This work was supported by the Faculty Research Grant from Lingnan University (LU Fund Code: 103433) and the Research Matching Grant Scheme (LU Fund Code: 185218) from the University Grants Committee of the Hong Kong SAR

33.313 TREATING LOGICAL FALLACIES IN A NORMATIVE COMPUTATIONAL FRAMEWORK OF PERCEPTUAL DECISION MAKING

Barnabás Molnár[1] (molnar_barnabas@phd.ceu.edu), Ádám Koblinger[1], József Fiser[1]; [1]Central European University

Cognitive fallacies are examples of breakdowns in human reasoning in which observers make irrational decisions as evaluated by Bayesian probability calculus. These examples were used for arguing that human reasoning does not operate by the rules of probabilistic computation in contrast with the surging trend of studies demonstrating that at the level of perceptual decisions, human behavior can be described well by probabilistic models. While multiple studies pointed out flaws in the investigations of cognitive fallacies, a comprehensive and quantitative treatment of the topic is missing. We provide such a treatment by placing perceptual decision making into a new framework and linking it to the problem of the "Base-rate fallacy" (BRF), one of the most celebrated cognitive fallacies. In BRF, individuals participating in vignette studies apparently do not consider the base-rate probabilities of events (priors) when making judgmental decisions. We created a standard 2-AFC perceptual decision making paradigm (N=23) where observers decided which of two shapes embedded in noise was presented in the trial, added one moment in the trial sequence (change point, CP) where significant change occurred to the conditions of the trials and measured behavior in trials well after the CP. We uncovered that humans' decision making under such conditions shows a far more complex but still probabilistic behavior than reported before. Generalizing this process, we found that keeping the process identical except for changing higher-level noise characteristics of the setup at the CP humans flip between interpretations of the input relying vs. not relying on assumed differences in the base rates perfectly mimicking the BRF. In conclusion, instead of being evidence for the lack of probabilistic treatment of the input, cognitive fallacies might be indicators of the same internal model based on probabilistic computations seamlessly transitioning into a particular unconscious interpretation of the current situation.

33.314 WHEN DOES RESPONSE DURATION TRACK PERFORMANCE?

Hanbei Zhou[1] (hzhou43@jhu.edu), Rui Zhe Goh[1], Ian Phillips[1], Chaz Firestone[1]; [1]Johns Hopkins University

A founding insight of psychophysics was to link internal mental processes to the timing of the behaviors they produce. Perhaps the most obvious and well-characterized example is the relationship between performance and response time, as when salient targets are found faster in visual search or when more confident perceptual decisions are made more quickly. But what is "response time"? Whereas nearly all psychophysical studies that measure the timing of behavior focus on the time taken to initiate a response, another potentially relevant magnitude is the duration of the response itself — e.g., not just how long it takes between the appearance of a stimulus and the onset of a keypress, but also how long one holds down the key before letting it go. Recent work makes a theoretical case that response duration may be a neglected source of data about visual processing (Pfister et al., 2023); here, 4 experiments provide empirical support for this proposal. Subjects completed a detection task in which a field of white noise either contained or didn't contain a face, with difficulty manipulated by varying the face's opacity. Subjects responded with a keypress (with both keyUp and keyDown events recorded separately). Remarkably, on more difficult trials, subjects not only took longer to initiate a response but also held down the response key for longer, as if answering in a tentative fashion. Response duration also tracked accuracy, with subjects holding down the response key for longer on incorrect as opposed to correct trials. These effects emerged again in a direct replication, but not in follow-up experiments using easier tasks. Overall, our results suggest that response duration may be an untapped source of information about performance — especially in tasks with high uncertainty — raising a wealth of avenues for future investigation.

This work is supported by NSF BCS #2021053 awarded to C.F., and a JHU PURA awarded to H.Z.

33.315 MAPPING VISUAL SEARCH ERRORS TO COVERT OPERATIONS WITH FRONTAL EYE FIELD NEUROPHYSIOLOGY AND DOUBLE FACTORIAL DESIGN

Wanyi Lyu[1] (wanyilyu@yorku.ca), Jeffrey Schall[1,2]; [1]York University, [2]Vanderbilt University

Behavior is the outcome of covert perceptual, cognitive, and motor operations that can be described by mathematical models and are produced by brain systems comprised of diverse neurons. Using the logic of selective influence, we previously distinguished the stages of processing supporting visual search (Lyu, Reppert, Schall, 2023). In that study, macaque monkeys searched for a color singleton among distractors. Two operations necessary for the task were independently manipulated. Singleton localizability was manipulated by varying the similarity between singleton and distractor colors. Stimulus-response mapping was manipulated by varying the discriminability of search array shape, signaling GO/NOGO response. The organization and termination rule of the two operations were determined using System Factorial Technology (SFT; Lowe et al, 2019). The necessary next step in this research is to account for performance errors in this difficult task, which influence the logic of the SFT diagnosis. Monkeys made two key errors: on GO trials, monkeys occasionally shifted gaze to a distractor due to unsuccessful localization (GO error). On NOGO trials, they failed to inhibit their saccade towards either the singleton or the distractors (NOGO errors). NOGO errors reflect failure in discrimination alone or both operations, respectively. We probed the neural sources of these error saccades using single-unit spiking in frontal eye field. Neurons representing stimulus salience were distinguished from neurons mediating saccade preparation. Our data suggest that GO errors occur when visual salience neurons misrepresenting the distractor as the singleton. NOGO errors to singleton arise from incorrect discrimination by saccade preparation neurons whereas NOGO errors to distractor arise from inaccurate response from both neuron types. The convergence of performance and neural results on error trials offer constraints to mathematical models and provide evidence so that distinct operations and their organization during visual search can be resolved.

Supported by NIH RO1-EY08890 and P30-EY008126

33.316 METACOGNITIVE CONTROL DRIVES BEHAVIOURAL EFFICIENCY IN DYNAMIC SENSORY ENVIRONMENTS

Tarryn Balsdon[1,2], Marios Philiastides[2]; [1]Ecole Normale Superieure and CNRS, [2]CCNi, University of Glasgow, UK

Metacognitive evaluations of decision confidence reflect insight into the quality and quantity of evidence supporting our decisions. Signatures of confidence emerge during decision-making, implying confidence may be of functional importance to decision processes themselves. We formulated an extension of sequential sampling models of decision-making in which confidence is used online to actively moderate both the quality and quantity of evidence accumulated for perceptual decisions. The benefit of this model is that it can respond to dynamic changes in sensory evidence quality. We highlighted this feature by designing a dynamic sensory environment where evidence quality can be smoothly adapted within the timeframe of a single decision. Observers made fine-grained motion discrimination decisions about random dot motion displays. Dot directions were sampled from a circular gaussian distribution where the mean and variance of the distribution were adapted frame to frame to create trials with increasing or decreasing decision-evidence quality. Observers made fast, accurate decisions with early high-quality evidence but slowed down when faced with early low-quality evidence (though trials were intermixed). Our model with confidence control offers a far superior description of this pattern of behaviour than can be obtained from traditional models without online control mechanisms. Using multivariate decoding of electroencephalography (EEG), we uncovered EEG correlates of the model's latent processes, and show stronger EEG-derived confidence control leads to faster, more accurate decisions within participants. These results support a neurobiologically plausible framework featuring confidence as an active control mechanism for driving efficient behaviour, that is, maximising precision given constraints of time and effort.

This work was supported by an European Research Council consolidator grant (865003) to MGP

Sunday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

Binocular Vision: Eye dominance and rivalry

33.317 INDEXING SENSORY EYE DOMINANCE

Qingzi Zheng[1], Dorita H. F. Chang[1]; [1]The University of Hong Kong

Sensory eye dominance occurs when one eye's input is weighted more preferably to that of the other. Currently, though there are different tests that are commonly used to index sensory eye dominance, they lack agreement in terms of their classification of the dominant eye. To systematically examine how sensory eye dominance might be influenced by stimulus and task demands, we designed tasks that we classified into two different categories: rivalrous tasks, which entail binocular competition and the eventual selection of one eye's data over the other; and integrative tasks, which allow for the combination of the two eyes' inputs in order to solve the task. For each category, we included tasks using a variety of different features. Specifically, for rivalrous tasks, we included letters with opposing polarity, opposite motion gratings, and masked oriented gratings of varying contrasts. For the integrative tasks, we included motion, glass pattern orientation, and gratings with phase shifts. Separately, we also included a depth perception task, which served as a test for one functional implication of eye dominance: stereovision. We found that sensory eye dominance classifications differ among tasks. When comparing across the categories of tasks, integrative tasks produced higher classification consistency than rivalrous tasks (e.g., the dominant eye classified by using integrative tasks is more likely to be the same eye). In terms of stereovision, we found no clear correlation between the magnitude of the sensory eye dominance and performance for the depth task. Our data suggest that sensory eye dominance is sensitive to task and feature demands and should therefore be indexed with caution. The apparent difference in outcomes derived from integrative- versus rivalrous-type tasks suggest they tap into independent mechanisms of sensory eye dominance.

33.318 VARIANCES IN SENSORY EYE DOMINANCE ACROSS THE VISUAL FIELD

Ka Yee Kam[1], Dorita H. F. Chang[1]; [1]Department of Psychology, The University of Hong Kong

Sensory eye dominance refers to the functional asymmetry between the two eyes, where the data from each eye is unequally weighted within the binocular neural network. In this study, we systematically measured sensory eye dominance at the fovea and at 16 isoeccentric locations in visually normal observers. Sensory eye dominance was measured using a dichoptic signal-in-noise motion (left-right) discrimination task. In this task, coherently moving signal dots and randomly moving noise dots were presented dichoptically to each eye. Observers were asked to make a two-alternative forced-choice judgement of the net motion direction of the dots. The magnitude of eye dominance was quantified as the ratio of the coherence thresholds (signal-to-noise ratio) obtained when signal dots were presented to one eye versus the other. Sensory eye dominance was measured at the fovea (0°), 3°, and 6° eccentricities. At each eccentricity, eight concentric locations were examined, corresponding to polar angles of 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°. Stimuli were scaled based on the cortical magnification factor. Results revealed that sensory eye dominance exhibited variation across the macular visual field. Specifically, individuals with relatively balanced eyes displayed variations in both the sign and magnitude of sensory eye dominance across the visual field. On the other hand, individuals with strong eye dominance tended to exhibit a consistent dominant eye, but with varying magnitudes of dominance observed across the visual field. Despite these variations, a significant correlation exists between sensory eye dominance in most tested locations, within observers. Our findings suggest that the degree of imbalance between the two eyes is not uniformly distributed across the visual field.

33.319 ASSESSING VARIATIONS IN EYE DOMINANCE ACROSS THE VISUAL FIELD

Chris Paffen[1] (c.l.e.paffen@uu.nl); [1]Utrecht University

Eye dominance (ED) in humans is not a unitary phenomenon: it depends on the method used (e.g. the Porta test, or sensory ED assessed by binocular rivalry), and it can vary over the visual field: both in the extent to which one eye (locally) dominates over the other, and in which eye dominates at a certain location. We used a variant of a new method, tracking Continuous Flash Suppression (tCFS; Alais et al, bioRxiv 2023), to rapidly assess ED across the visual field. TCFS is based on breaking Continuous Flash Suppression (bCFS). With the latter, a target gradually increasing in intensity is presented to one eye, while a dynamic mask is presented to the other. The observer responds when the target becomes visible. In our bCFS-experiments, a target-mask pair was presented at -10, 0 and 10 deg from fixation. In our variant of tCFS, a target-mask pair moved horizontally across the visual field (from -10 to 10 deg at 1 deg s-1), and the observer continuously increased and decreased the intensity of the target, respectively to make it visible and invisible. In Experiment 1A (bCFS, n=18) & 1B (tCFS, n=18), the mask contrast was 50%; in Experiment 2A (bCFS, n=18) & 2B (tCFS, n=18), it was 25%. ED varied over the visual field: observers had parts in the visual field at which one of the eyes was dominant, and parts where this was not the case. We even encountered observers with right ED for one location, and left ED for another. We also replicated the recently reported nasal visual field advantage for both methods (the nasal visual field of one eye dominates over the temporal hemifield of the other). Finally, we validated the new method by showing that ED assessed by tCFS correlated significantly with that assessed by bCFS.

33.320 INTERMITTENT THETA BURST STIMULATION (ITBS) OF PRIMARY VISUAL CORTEX REDUCES SENSORY EYE DOMINANCE.

Junyu Wang[1], Dorita H F Chang[1]; [1]The University of Hong Kong

Sensory eye dominance (SED) is characterized by an uneven weighting of visual inputs from the two eyes by the brain. It is well-speculated that SED is driven by mechanisms at primary visual cortex (V1). Here, we compared the effects of brain stimulation in the form of continuous theta burst stimulation (cTBS) and intermittent theta burst stimulation (iTBS) over V1 (and control site, Cz) on SED. We also investigated the effects of iTBS/cTBS stimulation on an important functional outcome of binocular vision: stereoacuity. Participants were tested in three phases: Pre-stimulation tests, stimulation delivery, and post-stimulation tests. During both the pre- and post-tests, participants completed two computer-based tasks: a dichoptic global motion task to index SED and a fine depth discrimination task to index stereoacuity. Depending on the group assigned to them, participants received stimulation in the form of cTBS (40s, 600 pulses) or iTBS (190s, 600 pulses) over neuronavigated V1 (localized via phase-encoded retinotopy), or Cz. We found that iTBS resulted in a significant decrease in SED. iTBS over Cz, nor cTBS over both V1 and Cz did not change SED between the pre- and post-tests. Furthermore, stimulation in both forms and all locations did not affect stereoacuity. Our data suggest that V1 is a key locus for driving SED. We speculate that iTBS may alter SED through counteracting interocular inhibition via its posited facilitatory effects. The ability to reduce SED has potential implications for clinical interventions aimed at ameliorating visual imbalances.

This work was supported by a General Research Fund [17610022] from the University Grants Committee (Hong Kong) to D.C.

33.321 ALTERING SENSORY EYE DOMINANCE USING MONOCULAR DEPRIVATION: DOES THE EYE DEPRIVED OR TASK MATTER?

Chenyi Qian[1], Dorita H.F. Chang[1]; [1]The University of Hong Kong

Recent work has shown that short-term monocular deprivation (MD) can alter sensory eye dominance (SED) among adults. Specifically, following deprivation, the deprived eye becomes more dominant. The challenge of unraveling the mechanism underlying deprivation-induced shifts of SED arises from the known task variability used to index it. Here, we aimed to examine the effects of MD on SED while contrasting the eye deprived (dominant vs. non-dominant) and the task used to index SED (phase combination and orientation rivalry tasks). Adults with normal vision underwent 2.5 hours of MD for each eye during two sessions separated by a 48-hour interval. We found that MD-induced shifts in SED were only evident when indexed by the phase combination task, and participants' SED decreased more evidently when the non-dominant eye was deprived. Subsequently, we investigated whether the observed effects generalized to different yet analogous SED metrics. New observers were tested on a dichoptic motion task, requiring binocular integration (akin to the phase combination task) and a letter polarity (rivalry) task invoking binocular competition (akin to the orientation rivalry task). The new comparison revealed that MD-induced SED shifts were only evident in the letter polarity task. As for Experiment 1, deprivation of the non-dominant eye produced a greater reduction in SED than deprivation of the dominant eye. Our data suggest that deprivation-induced SED shifts are independent of task category but depend on the eye deprived, with the deprivation of the non-dominant eye consistently leading to a more balanced outcome. We speculate that the changes in SED induced by MD of the non-dominant eye reflect a homeostatic response that transiently lifts the gain control exerted by the dominant eye. Conversely, we expect a weaker release of gain control when the dominant eye is deprived, as it experiences less suppression from the other eye.

33.322 SHORT-TERM MONOCULAR DEPRIVATION BIASES THE LOCATION OF THE VISUAL EGOCENTRE

Yalige Ba[1] (lpxby1@nottingham.ac.uk), Paul mcGraw[2], Timothy Ledgeway[2]; [1]University of Nottingham

Monocular visual deprivation, even for relatively brief periods, can bias sensory eye dominance in favour of the previously deprived eye. This effect has been measured using a range of tasks including binocular rivalry, binocular phase combination and continuous flash suppression. The mechanisms underpinning shifts in sensory eye dominance remain unresolved, but may involve unbalanced inter-ocular suppression or attentional shifts in eye selection. We investigated if short-term monocular deprivation also has consequences for judging the visual direction of objects in space, relative to the viewer (the visual egocentre). A binocular sighting task was used to establish the baseline location of the visual egocentre in visually-normal adults (N=9). Observers were asked to rotate the orientation of a rod in the horizontal plane until it pointed directly at them. Measurements were made along the horizontal azimuth for each of a range of eccentricities spanning ±30 deg. The mean point of intersection of the extensions of the rod's axis at each eccentricity, reconstructed the egocentre location for each individual. The egocentre was located on average 5.94 (±9.07) mm to the right of the median plane of the head, but varied greatly between individuals. One eye was then covered with an opaque patch and observers were free to engage in routine activities for 60 min. Upon patch removal the egocentre was remeasured repeatedly for 45 min. Short-term monocular deprivation induced a marked shift in egocentre location of 7.47 (±4.25) mm. Surprisingly, this was always towards the direction of the median plane of the head, rather than the eye that had been deprived. This effect was found for occlusion of dominant and non-dominant eyes and dissipated to baseline levels after 15 min. Unlike sensory measures of binocularity, short-term monocular deprivation appears to normalise the location of the visual egocentre to the middle of the head.

China Scholarship Council

33.323 HOMEOSTATIC AND HEBBIAN PLASTICITY ARE RELATED IN ADULT HUMANS

Antoine Prosper[1], Claudia Lunghi[1]; [1]Laboratoire des Systèmes Perceptifs, CNRS/PSL University, France

Introduction: Homeostatic and Hebbian plasticity co-operate during the critical period, refining neuronal circuits, however the interaction between these two forms of plasticity is still unclear, especially in adulthood. Here we investigate directly this issue in adult humans using two consolidated paradigms to elicit each form of plasticity: the LTP-like change of the visual evoked potential (Kirk et al 2021) and the shift of ocular dominance induced by short-term monocular deprivation (Baroncelli & Lunghi 2021). Methods: We tested three different conditions in a group (N=17) of adult volunteers (counterbalanced order). In the homeostatic condition, we measured ocular dominance (binocular rivalry between gratings, 2°, 2cpd, 64% contrast) before and after 1h of monocular deprivation (MD). In the Hebbian condition, we measured visual evoked potentials (VEP) in response to flashed (mean rate 1Hz) visual stimuli (checkerboard: 4°, 2 cpd, 64% contrast) presented binocularly in the upper visual field, before and after a 2-min high-frequency visual stimulation (HFS, checkerboard flickering at 9Hz). The mixed condition was similar to the homeostatic condition, except that HFS was delivered just before MD. Results: In the homeostatic and hebbian conditions we confirmed that the two experimental paradigms successfully induced each form of plasticity: ocular dominance shifted in favor of the deprived eye after MD ($t(16)=8.23$, $p<0.001$), and the amplitude of the N1b component of the VEP was reduced after HFS ($t(16)=-2.569$, $p=0.021$). Importantly, across participants, the effect of MD and the effect of HFS were correlated ($r=-0.538$, $p=0.026$). In the mixed condition, the effect of MD was comparable to the homeostatic condition ($t(16)=-0.138$, $p=0.892$). Conclusion: Homeostatic and Hebbian plasticity correlate in amplitude in adult humans, but they do not seem to interact when induced simultaneously. The relationship between these two forms of plasticity can explain the success of homeostatic plasticity based paradigms for the treatment of adult amblyopia.

This project has received funding from the European Research Council (ERC) under Horizon 2020 research and innovation program (No 948366—HOPLA)

33.324 TRANSCRANIAL MAGNETIC STIMULATION TO EARLY VISUAL CORTEX MODULATES BINOCULAR RIVALRY

Stefania S. Moro[1] (moros@yorku.ca), Remy Cohan[1], Jennifer K. E. Steeves[1]; [1]Centre for Integrative and Applied Neuroscience, Centre for Vision Research, and Department of Psychology, York University, Toronto, Canada

Neuromodulation techniques have been pivotal for mapping neural networks. Among these, repetitive transcranial magnetic stimulation (rTMS) is a non-invasive technique that has proven effective in promoting neuroplasticity in various brain regions. Continuous theta burst stimulation (cTBS) is a form of rTMS thought to promote inhibitory effects in primary motor cortex. Our study extends the exploration of cTBS to the visual domain, investigating its impact on binocular rivalry (BR) after administering 600 pulses at 80% of phosphene threshold (PT) to primary visual cortex (V1). Binocular rivalry is characterised by alternating periods of visual dominance when two different images are presented simultaneously to each eye. BR is orchestrated by a complex interplay across the visual pathway, from retinal input to lower and higher cortical processing (including attention). Notably, cortical columns in V1 exhibit monocular responses that underpin the perceptual alternations witnessed in BR, however, the exact mechanism is not well understood. In our sham-controlled study, 19 individuals underwent a BR task, observing orthogonal grey-scale gratings with fixed orientation (+/-45°) through a mirror stereoscope and reporting shifts in visual dominance. Following cTBS application to the left V1 using stereotaxic neuronavigation, participants reported the frequency of BR alternations. Our preliminary findings reveal an increased alternation rate in BR after active cTBS (and not sham), suggesting that cTBS can modulate perceptual dominance in visual processing. These insights enhance our comprehension of cTBS's neuromodulatory potential and underscore its promise as a tool for research and potential treatment of vision disorders rooted in neural dysfunction. Our research contributes to the expanding dialogue on the intersection of neuromodulation, visual perception, and cortical plasticity.

33.325 VISUAL UNCERTAINTY IN BINOCULAR RIVALRY

Zhangziyi zhou[1] (zzhou@ski.org), Chuan Hou[1]; [1]The Smith-Kettlewell Eye Research Institute

When different images are presented to the two eyes, the perception alternates between the eyes, known as binocular rivalry. The duration of perception between the eyes can be influenced by various factors, including the stimuli (size, contrast, spatial frequency, etc.) and the viewing eye. Binocular rivalry dynamics can also be modulated by attending to features on one of the rival stimuli. Aside from binocular rivalry during presenting the rival stimuli, there is also a third condition that reveals a mixture of two eyes' images and the observers are unsure of which image of the rival stimuli to report. The profile of this third condition (referred to as "binocular perception") from rival stimuli are less documented. In this study, we investigated binocular rivalry and binocular perception by presenting a pair of orthogonal gratings in the two eyes in normal-vision observers. The rival stimuli presented for 30s with equal contrast (40%) in each eye and varied in 1°, 2.5°, 5°, 8° and 18° stimulus sizes. In experiment 1, we measured percentage proportion over the 30 s stimulus duration for 3 perceptual conditions: dominant eye viewing (defined by hole-in-card test), non-dominant eye viewing and binocular perception. In experiment 2, we investigated how attention affects these perceptual conditions. Our results showed that the dominant eye had more proportion than the non-dominant eye and the stimulus size matters. Selectively attending to the stimuli of the non-dominant eye increased its proportion with the most increasing at 5° stimuli. The proportion of binocular perception is negatively correlated with the stimulus size but shown no effect on attention. Our results in binocular perception reveal that the visual uncertainty in binocular rivalry decreases with increasing the size of rival stimuli and suggests that the size of rival stimuli should be around 5° or beyond to get ideal binocular rivalry.

Support by NIH Grant R01- EY025018 and R01- EY035346 to C. H.

Sunday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

Color, Light and Materials: Neural mechanisms, models, disorders

33.326 PSYCHOPHYSICAL MEASURE OF THE IMPACT OF HEALTHY AGING ON RODS AND CONES OF THE RETINA

Geneviève Rodrigue[1] (genevieve.rodrigue.2@umontreal.ca), Julie Chambellant[1], Judith Renaud[1], Rémy Allard[1]; [1]Université de Montréal

Previous studies found that the efficiency of cones to detect light declines with healthy aging. These psychophysical studies used an equivalent input noise (EIN) paradigm based on the measurement of contrast thresholds with and without visual noise under specific conditions in which contrast threshold is limited by the variability in the amount of light detected by cones (i.e., photon noise). We recently adapted this paradigm to measure the efficiency of rods, which are the photoreceptors used under scotopic conditions (i.e., night vision). To assess the impact of healthy aging on the efficiency of rods and cones to detect light, the current study measured contrast thresholds in 22 young [20-30 years old, mean=25.7] and 22 older [65-77 years old, mean=70.9] healthy adults using a blue background under scotopic condition, and a red background under low photopic condition, respectively. The task consisted in discriminating the motion direction of an annulus grating centered on fixation with a radius of 10° of eccentricity in presence and absence of visual noise. A 2x2 mixed factorial ANOVA showed that older adults had significantly higher EIN, which suggests a lower amount of light detected by photoreceptors (rods and cones) compared to young adults ($F(1,42)=29.7$, $p<.001$). No significant interaction was found between the group and the display color, which suggests a similar age-related decline when using a blue and red display. We conclude that the amount of light detected by rods and cones declines similarly with healthy aging. These similar age-related declines in the amount of light detected by both rods and cones can be explained by similar age-related physiological alterations for the two types of photoreceptors, or a common cause affecting both types of photoreceptors such as a change in the orientation of the photoreceptors.

This research was supported by grants from NSERC, FRQNT, Réseau de recherche en santé de la vision and Fondation Antoine-Turmel to RA, and scholarships from École d'optométrie and Études supérieures et postdoctorales de l'Université de Montréal to GR.

33.327 SIGNALS FROM S-CONE-DRIVEN SINGLE-OPPONENT NEURONS IN THE HUMAN VISUAL CORTEX

Songlin Qiao[1] (qiaosl63@163.com), Karl Gegenfurtner[2], Jing Chen[1]; [1]Shanghai University of Sport, [2]Justus-Liebig-Universität Gießen

Recent studies of middle- and long-wavelength-sensitive cone responses have revealed neural activities from single-opponent neurons in the human visual cortex (Nunez et al., 2022; Qiao, Gegenfurtner & Chen, VSS 2022). However, it is still uncertain whether single-opponent activities driven by short-wavelength-sensitive cones (S-cones) can be observed in humans. Previous studies conducted on non-human primates have shown limited S-cone-driven single-opponent neurons (Lennie et al., 1990; Johnson et al., 2004; Solomon and Lennie, 2005). In human observers, Nunez et al. (2022) reported very weak or no S-cone-driven single-opponent signals using visual evoked potentials (VEPs), which may be due to a lack of single-opponent neurons in the human visual cortex or because VEP recordings are too noisy to detect weak signals. To address this issue, we conducted a study recording steady-state visual evoked potentials (SSVEPs) to S-cone stimuli flickering at 2 temporal frequencies (3Hz and 15Hz) and 6 spatial frequencies (0.2, 1, 2, 3, 4, and 8 c/deg) in 16 observers. Our results revealed that the response at 3 Hz has band-pass tuning (double-opponent) and the response at 15 Hz has low-pass tuning (single-opponent). We further investigated spatial tuning between 3Hz and 15Hz by measuring SSVEPs on 10 observers with S-cone stimuli at several temporal frequencies (3 Hz, 6 Hz, 10 Hz, and 15 Hz). The result indicates that spatial tuning is band-pass at low temporal frequencies (3, 6, 10 Hz) and low-pass at 15 Hz. Overall, our study suggests that there are S-cone-driven single-opponent neurons in the human visual cortex.

33.328 ALTERNATING ORIENTATION OF THE CHROMATIC PATTERN VISUAL EVOKED POTENTIAL IMPROVES SIGNAL, EVEN IN THE ABSENCE OF CONTRAST ADAPTATION.

Jawshan Ara[1], Alireza Tavakkoli[1], Michael A. Crognale[1]; [1]University of Nevada Reno

The visual evoked potential (VEP) to chromatic pattern reversal is greatly reduced compared to VEPs to pattern onsets. Chromatic pattern onsets produce large and stereotypical waveforms that reliably differ from standard achromatic pattern reversal VEP waveforms used in clinical applications. Rapid contrast adaptation for sustained chromatic but not transient achromatic mechanisms has been suggested as one explanation for these observations. Here we first examined changes in the magnitude of response during recordings to reversing and onset grating patterns that preferentially modulate the L-M, S, and achromatic pathways. VEPs were recorded for 60 s with 2 onsets/reversals per second using both fixed and alternating (horizontal/vertical) orientations. We hypothesized that contrast adaptation may be reduced by changing the orientation and/or phase of the pattern for each reversal or onset. However, responses for 6-second windows did not reveal evidence of adaptation for chromatic or achromatic onsets or reversal patterns over the 60-second recording period. Despite this, alternating pattern orientation and/or phase improved the signal amplitudes for chromatic onset and achromatic reversal conditions. Onset responses were larger for all conditions, even for achromatic stimuli. Additionally, alternating pattern orientation and/or phase produced shorter response latency for all onset conditions. The responses to achromatic and LM reversals and S onset conditions with orientation or phase were larger than responses to horizontal or vertical fixed patterns. In most cases, both orientation and phase changes improve the chromatic onset responses equivalently. Mechanisms other than contrast adaptation must be invoked to explain the results.

33.329 CHROMATIC CENTER-SURROUND ANTAGONISM REVEALED BY THE WESTHEIMER PARADIGM

Christopher S Wu[1], Daniel R Coates[1,2]; [1]University of Houston, College of Optometry, [2]University of Washington, Department of Ophthalmology

Visual sensitivity to an object is dependent on its surrounding context. The Westheimer paradigm quantifies the effects of surround pedestal size on sensitivity to a centrally located spot of light. The resultant curve is thought to reveal the center-surround interactions of underlying neural mechanisms. While the achromatic Westheimer function has been well characterized, chromatic mechanisms remain relatively unexplored. The goal of this study was to characterize spatial interactions for stimuli defined by the cardinal axes of DKL color space. Chromatic cone contrast thresholds to a spot of light were measured as a function of the size of a surrounding pedestal in four subjects. Stimuli were presented on persistent pedestals randomly to the left or right of fixation at 5° retinal eccentricity, with detection indicated by a button press corresponding to the side that contained the target. The target and pedestal were defined along the same direction in color space, with the pedestal set at a lower contrast. The achromatic background had a luminance of 10 cd/m2. For chromatic conditions, luminance noise was added to potential target locations to minimize non-chromatic cues. Results were well described by a ratio of gaussians (ROG) model. Pedestal-dependent amplitudes were consistent within visual mechanisms and between subjects. Perceptive field center size was defined by the peak of desensitization, and surround size by the asymptote where increasing pedestal size no longer affected thresholds. For achromatic conditions, mean center and surround sizes were 0.36° (±0.04) and 4.4° (±1.0) respectively, 0.98° (±0.16) and 3.8° (±0.36) for +LM direction, 1.2° (±0.10) and 5.5° (±0.71) for the -LM direction, 1.3° (±0.13) and 6.4° (±0.57) for the +S direction, and 1.49° (±0.12) and 6.4° (±1.0) for the -S direction. The large sizes suggest a cortical locus and provide insight into the spatial grain of low-level chromatic mechanisms.

33.330 CONSEQUENCES OF FIXATIONAL EYE MOVEMENTS FOR CHROMATIC SENSITIVITY

Alina Neverodska[1], Yuanhao H. Li[1], Eli Brenner[2], Michele Rucci[1]; [1]University of Rochester, [2]Vrije Universiteit Amsterdam

The human eyes are always in motion. Fixational eye movements (FEM) continually shift gaze even when attending to a single point, displacing stimuli on the retina over an area as large as the foveola. Previous research with achromatic stimuli has shown that the temporal luminance modulations from FEM enhance sensitivity to spatial luminance modulations in the scene, especially for high spatial frequencies. We here examine whether the same is true for chromatic modulations. It has been proposed that FEM enhance sensitivity to color, but the classical experiments that studied this by eliminating retinal image motion focused on the consequences of adaptation during unnaturally long periods of fixation. We examine how the retinal motion resulting from FEM influences chromatic sensitivity during stimulus exposures of durations that are comparable to natural fixation. In a forced-choice procedure, emmetropic observers (N=6) were asked to report the orientation (±45°) of an isoluminant red-green grating (1 or 10 cycles/deg) superimposed on a Brownian noise colored background. Subjects maintained fixation at the center of the display while their eye movements were recorded at high resolution via Dual Purkinje Imaging. We compared performance in the presence and absence of retinal image motion. The latter was achieved by moving the stimulus on the display to counteract eye movements by means of a custom-built system for gaze-contingent display. For each subject, we selected a contrast that yielded approximately 90% correct during normal fixational motion. Discrimination was greatly impaired under retinal stabilization, resulting in approximately 60% correct on average across subjects. Unlike with luminance stimuli, for which FEM consequences are clearest at high spatial frequency, the drop in performance with image stabilization was similar for chromatic stimuli at both high and low spatial frequencies. These results indicate that temporal modulations from FEM play an important role in the perception of color.

Research supported by Reality Labs and NIH P30 EY001319.

33.331 PARVO VERSUS MAGNO ISOLUMINANCE

George Sperling[1], Lingyu Gan[1]; [1]University of California, Irvine

Visual stimuli are isoluminant when they are uninterpretable by a system that is primarily sensitive to spatial variations in luminance but are interpretable by systems that are sensitive to variations in chromaticity or other visual variations that are invisible to the luminance system. Isoluminance stimuli are important for demonstrating that, while luminance information contributes to the representation of visual information in the brain, luminance is not the modality of representation. A common method for creating differently colored isoluminant patches is to alternately present them at a flicker frequency of 10 Hz, and to adjust the relative intensities to minimize perceived flicker (<1% luminance variation would be readily visible). This and other common methods create isoluminance only for the magno visual system--neurons that have large receptive fields and high temporal resolution. Parvo pathways involve neurons with small receptive fields, poor temporal resolution, but high spatial resolution for tasks like reading. Parvo isoluminance has been largely neglected. Procedures and results: To test parvo isoluminance, participants foveally viewed high spatial-frequency, yellow/red, and yellow/green gratings. The intensity of red and green was varied to find the points of minimal visibility. When the same colors were tested in alternating patches in a 15 Hz minimum flicker paradigm, the points of minimal visibility were profoundly different: Primarily, the utility of red light was much greater for spatial acuity than for flicker detection. These data demonstrate that isoluminance is a property that depends on the particular mixture of neurons involved, which in turn depends on the nature of the task (proportion of magno to parvo activation), the locations on the retina, and many other factors. Conclusion: Conditions for isoluminance vary greatly; it must be evaluated in circumstances as similar as possible to the tasks being studied.

None

33.332 MECHANISM OF POSITIVE COLOR AFTERIMAGE CAUSED BY DICHOPTICAL PRESENTED CONTOURS

Tan-Ni Yang[1], Chien-Chung Chen[1]; [1]National Taiwan University

The negative afterimage following an adaption to a specific color can be enhanced by a contour surrounding the adapting area in the test phase. However, if the contour is presented in a different eye from the adaptor, the observer would see a positive afterimage. To investigate the mechanism responsible for this positive aftereffect, we measured the target threshold versus pedestal contrast (TvC) functions under various adaptor and contour conditions. Each trial began with an adaptation phase in which adaptors (two red squares positioned above and below the fixation) were presented to either the left or right eye for 1s, followed by a test phase in which contours (two white square outlines matching the shape of the adaptors) were presented on either the same or the different eye from the adaptor, or not presented. Simultaneously, two pedestals (green squares) were displayed in the same position as the adaptors with a target square on top of one of the pedestals. Participants were to indicate the location of the target. Without the adaptor and the contour, the target threshold was flat at low and increased at high pedestal contrasts. The threshold increment caused by adaption was large at low pedestal contrast but weakened as the pedestal contrast increased. The same-eye contour decreased threshold at low pedestal contrasts and a masking effect at high pedestal contrasts, whereas the cross-eye contour increased threshold at low pedestal contrasts. These contour effects were weakened with adaptation. These data were fitted with a version of the divisive inhibition model with lateral sensitivity modulation, which showed that the same- and different-eye contours have an opposite effect on the visual mechanism's sensitivities to the target and that adaption changes the dynamic range of the response function. These effects combine to produce a positive color aftereffect.

NSTC (Taiwan) 109-2410-H-002 -086 -MY4

33.333 A RECIPE FOR A 4+ PRIMARIES DLP PROJECTOR

Lindsey Fraser[1] (lfraser@vpixx.com), Sophie Kenny[1], Jonathan Tong[1]; [1]VPixx Technologies

Human eyes sense light using five distinct receptor types: long, medium and short-wavelength cones; rods; and intrinsically photosensitive retinal ganglion cells (iPRGCs). The sensitivity distributions of these five receptors overlap substantially. To study their relative contribution to visual perception, it is necessary to develop specialized displays that can selectively isolate each receptor type. Here, we describe one such display, the PROPixx Multispectral DLP projector by VPixx Technologies. The PROPixx Multispectral is designed with red, green, blue and yellow LED primaries. We present technical details of this projection system, outline engineering challenges associated with building its light engine, and describe the spectral power distributions of each light source. The use of four primaries facilitates the method of silent substitution, an experimental paradigm in which light pairs are used to selectively modulate individual receptor types, while maintaining an identical level of activity in the others, thus "silencing" them. We discuss this technology's potential to advance vision research and briefly discuss how adding additional primary light sources (e.g., cyan, ultraviolet) can unlock additional advanced vision research paradigms.

33.334 BIOMIMETIC-INSPIRED RESILIENT LEARNING: IMPACT OF PROGRESSIVE CHROMATIC VARIATIONS ON THE FACE RECOGNITION PERFORMANCE

Joydeep Munshi[1] (joydeep.munshi@ge.com), Hojin Jang, Michal Fux, Suayb S. Arslan, Matt Groth, Walter Dixon, Pawan Sinha; [1]Lead Scientist, [2]Postdoctoral Associate

As is evident from the ease with which we can recognize grayscale images, humans are remarkably resilient to changes in chromatic content. Based on results from children with late-onset sight, we have hypothesized that this resilience may be based, in part, on the developmental progression of the color system, with initially poor chromaticity eventually maturing into rich color experiences (Vogelsang et. al., under review). Computational tests of this hypothesis have involved the creation of training regimens wherein the first half of training data is devoid of color, and the second half has full color information. The results so far indicate that such a training progression yields greater resilience to color shifts than one where all training data are in rich color. While encouraging, these investigations are still quite different from true developmental progressions which are characterized by gradual changes in the amount of chromatic content over time. The goal of this work is to examine the consequences of graded introductions of color information across the training epochs. We trained Facenet-512 from scratch following different training regimens: 1) end-to-end color, 2) end-to-end grayscale, 3) random augmentation of color and grayscale, 4) color to grayscale, 5) grayscale to color, and 6) grayscale to color through gradual enhancement of chromaticity. Our experiments revealed that the quasi-biomimetic regimen (#5) and the biomimetic one (#6), both significantly improved face recognition accuracy across a range of color shifts. Notably, strategy # 6 further enhances the accuracy relative to the abrupt grayscale to color change strategy (#5) suggesting that closer alignment with human developmental progression may be a useful computational training strategy.

This research is supported by ODNI, IARPA. The views are of the authors and shouldn't be interpreted as representing official policies of ODNI, IARPA, or the U.S. Gov., which is authorized to reproduce & distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

33.335 DEEP LEARNING MODELS FOR LIGHTNESS CONSTANCY CAN EXPLOIT BOTH NATURAL LIGHTING CUES AND RENDERING ARTIFACTS.

Alban Flachot[1] (flachot.alban@gmail.com), Jaykishan Patel[1], Marcus A. Brubaker[1], Thomas S.A. Wallis[2], David H. Brainard[3], Richard F. Murray[1]; [1]York University, [2]TU Darmstadt, [3]University of Pennsylvania

We previously showed that deep learning models that estimate intrinsic image components (albedo and illuminance) outperform classic models on lightness constancy tasks. Here, we examine what cues these models rely on. We considered two cue types: natural features such as shadows and shading, and artifacts of ray tracing softwares, which typically produce a residual rendering noise that varies with local illumination. We rendered training, validation and test sets via (1) ray tracing (Blender/Cycles) with 128 photons sampled per pixel (high residual noise); (2) same as (1) but 1024 photons sampled (low noise); (3) Blender's EEVEE renderer (rasterization engine, no noise). (Noise artifacts are also found in other ray tracing renderers, including Mitsuba.) Networks trained on EEVEE images showed similar performance on all three test sets (and performed much better than classic models), whereas networks trained on Cycles showed best performance with Cycles test images, and worst performance with EEVEE images. To assess dependence on naturalistic cues, we tested the networks on test images with various scene elements removed: (1) cast shadows on the floor; (2) shading; (3) all shadows and shading. In (3), no naturalistic lighting cues were available, and yet models trained on Cycles keep a partial, if low, constancy. These models were also almost unaffected by the removal of shadows and shading (less than 10% decrease in constancy). However, the model trained on EEVEE showed a 50% decrease in constancy when floor shadows were removed, and had lowest constancy in condition (3). These results show that widely used ray tracing methods typically produce artifacts that networks can exploit to achieve lightness constancy. When these artifacts are avoided, networks rely on more naturalistic lighting cues, and still exhibit human levels of constancy. Thus deep networks provide a promising starting point for image-computable models of human lightness and color perception.

Funded by a VISTA postdoctoral fellowship to Alban Flachot

33.336 UNVEILING THE TEMPORAL DYNAMICS OF DIURNAL AND CREPUSCULAR ILLUMINATION

Cehao Yu[1] (cy153@student.london.ac.uk), Sylvia Pont[2], Ruben Pastilha[1], Anya Hurlbert[1]; [1]Centre for Transformative Neuroscience, Biosciences Institute, Newcastle University, United Kingdom, [2]Perceptual Intelligence Laboratory, Faculty of Industrial Design Engineering, Delft University of Technology, The Netherlands

Natural diurnal illumination, daylight, consists of sunlight and skylight. Natural crepuscular illumination, twilight, is formed by scattered sunlight, moonlight, starlight, and airglow emissions. Outdoor illumination, encompassing both diurnal and crepuscular light, undergoes intensity and spectral changes throughout the day. These are characterized by spectral irradiance fluctuations across various temporal scales and spatial locations due to atmospheric dynamics. To facilitate future modelling of how these fluctuations influence visual perception and circadian rhythms, we analyzed spectral light-field data (ranging from 360 to 780 nm) collected from dawn to dusk on typical days. Our analysis combined data from two sources: the Delft light field database, covering one cloudy and one sunny day, and additional measurements from Newcastle for an overcast day. We decomposed global illumination into time-varying directional and diffuse components across three distinct weather conditions. This analysis uncovered a clear three-part pattern on the sunny day, marked by major chromaticity shifts from blue to yellow-orange in the early morning and yellow-orange to blue in late afternoon. Weather conditions and solar elevations substantially influenced illuminance, showing rapid transitions at both daybreak and dusk. On overcast days, the dominant illumination direction was upward, with chromaticity remaining stable. Conversely, cloudy days experienced sharp shifts in illuminance and chromaticity when direct light was interrupted by passing clouds, highlighting the variability of natural outdoor illumination. Psychophysical measurements suggest that many daylight variations, apart from those due to rapid cloud movements, will be undetectable. Such changes may be picked up by the slower non-visual light response pathway, potentially affecting circadian rhythms. Initial analyses revealed that directional components alter more swiftly than diffuse components in chromaticity and illuminance, underscoring their distinct roles in natural illumination. These findings highlight the need to differentiate between directional and diffuse effects in understanding natural illumination dynamics and in shaping the anthropogenic light environment.

33.337 IMAGE STATISTICS OF MELANOPSIN-MEDIATED SIGNALS

Pablo Barrionuevo[1,2] (pbarrionuevo@herrera.unt.edu.ar), Francisco Diaz Barrancas[1]; [1]Justus Liebig University, [2]ILAV, CONICET-UNT

Melanopsin-expressing intrinsically photosensitive retinal ganglion cells (ipRGCs) project to nervous areas involved in visual processing. It is thought that stable and long-term responses of these cells might provide unique information to visual perception complementary to the transient and adaptive rod and cone responses. There is a close match between natural statistics and processing of contrasts in the early visual system but most of the studies were carried out considering luminance (cone) statistics. This study aimed to determine the intensity and contrast statistics of melanopsin and ipRGC codification of photoresponses in natural and human-made scenes. We have computed intensity (melanopsin excitation and luminance) and within-image Michelson contrasts using the melanopsin (Mel) and cone (L, M, and S) CIE fundamentals on hyperspectral images with natural (21) or human-made elements (10) under natural illumination. Statistics were obtained considering 16 patches per image, mimicking receptive fields, each one weighted by a raised cosine window covering 2.4° (ipRGCs) and 1° (luminance-processing parasol cells). Also, statistics on inferred ipRGC photoreceptor codification (Mel+L+M-S) were computed. Our findings, based on median values, reveal that human-made environments exhibit significantly higher melanopsin (2.53x) and

ipRGC (15.3x) excitations compared to natural environments. This increment is higher than the luminance increment (1.57x). Regarding median contrasts, lower values for melanopsin (41.2% and 41.3%) and ipRGCs (42.3% and 42.6%) than luminance (49.1% and 45.9%) were found for both natural and human-made environments, respectively. These differences spread all over the contrast range. There was no apparent correlation between melanopsin contrasts and excitations. Our analyses showed that the melanopsin-mediated and ipRGC-mediated statistics differ from luminance-mediated statistics. Including human-made elements in the scene affects melanopsin excitations more than luminance, but melanopsin contrast was not affected. Melanopsin statistics found in natural environments might serve as a reference to discuss the implications of lab-conducted studies.

This study has received funding from the German Research Foundation (DFG) - project number 222641018 - SFB/TRR 135 TPs C2 and B2 and by the European Research Council Advanced Grant 'An object-oriented approach to color: Color 3.0.' -project number 884116.

33.338 EFFECT OR ARTIFACT? ASSESSING THE STABILITY OF COMPARISON-BASED SCALES

David-Elias Künstle[1] (david-elias.kuenstle@uni-tuebingen.de), Felix A. Wichmann[1]; [1]University of Tübingen

Measuring the subjective similarity of stimuli—for example, the visual impression of materials or the categories in object images—can be achieved through multidimensional scales. These scales represent each stimulus as a point, with inter-point distances reflecting similarity measurements from behavioral experiments. An intuitive task used in this context is the ordinal triplet comparison: "Is stimulus i more similar to stimulus j or k?". Modern ordinal embedding algorithms infer the (metric) scale from a subset of (ordinal) triplet comparisons, remaining robust to observer response noise. However, the unknown residual errors raise concerns about interpreting the scale's exact shape and whether additional data may be necessary or helpful. These observations demand an examination of the scale's stability. Here, we present an approach to visualize the variation of comparison-based scales via bootstrapping techniques and a probabilistic model. Simulation experiments demonstrate that variation is broadly captured by an ensemble of scales estimated from resampled trials. However, common methods to align the ensemble parts in size, rotation, and translation can distort the local variation. For example, standardization results in zero variation at some points but bloats variation at others, while Procrustes analysis leads to uniform "noise" distribution across all scale points. To address this, we propose a probabilistic model to identify the variation at individual scale points. In essence, we "wiggle" scale points to observe changes in triplet correspondence, indicating their stability level. These localized estimates are combined in the ensemble to provide a robust measure of variation. Simulations validate our approach, while case studies on behavioral datasets emphasize the practical relevance. In these case studies, we visualize perceptual estimates through regions instead of points and identify the most variable stimuli or trials. Beyond data analysis, our stability measures enable further downstream tasks like adaptive trial selection to expedite experiments.

Funded by EXC number 2064/1 – project number 390727645 and by German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A. Supported by the International Max Planck Research School for Intelligent Systems (IMPRS-IS).

33.339 COLOR CATEGORIES IN COLOR ANOMALOUS TRICHROMATS AND DICHROMATS

Aimee Martin[1] (aimee.martin@uqconnect.edu.au), Karl Gegenfurtner[1]; [1]Justus-Liebig-Universität Gießen

Observers with color vision deficiencies (CVD) tend to have similar color categories as trichromats, despite sometimes severe sensitivity limitations along the red-green. We investigated the similarity of color categories of anomalous trichromats and dichromats, compared to five color-normal trichromats. We also investigated whether EnChroma glasses could improve categorization for anomalous trichromats. Participants were asked to sort 450 Munsell chips (from the World Color Survey plus 40 desaturated chips) into 11 color categories. They then waited 2 hours wearing the glasses and sorted the chips again into the categories. They either did so with the glasses on while sorting the second time, or with the glasses removed after the adaptation phase. Five color-normal trichromats were measured for reference. We define 217 core chips as those where all trichromats agreed, and 444 majority chips as the ones where at least 3 out of the 5 agreed on the category. In line with previous findings, the color categories of the CVD observers matched the trichromats categories closer than compared to what simulations would predict. This was particularly the case for the core chips at the centre of the categories. Furthermore, for the anomalous trichromats, the color categories did not differ very much after wearing the EnChroma glasses for 2 hours, even when wearing the glasses during the second sorting task. The relatively high performance of the CVD observers may be due to the high saturation of most of the chips we used. It may also be supported by strategic sorting decisions, because the observers knew that the set of chips covered the whole color space.

DFG Sonderforschungsbereich SFB TRR 135 C2, project number 222641018

33.340 EVALUATION OF NOVEL TABLET-BASED COLOR VISION TESTS

Christabel Arthur[1] (christabela@nevada.unr.edu), Jingyi He[2], Peter J. Bex[3], Jan Skerswetat[3], Michael A. Crognale[1]; [1]University of Nevada, Reno, [2]University of California, Berkeley, [3]Northeastern University, Boston

Color vision testing has shown significant advancement throughout the years, with a recent shift to computerized and self-administered testing, including the application of adaptive psychophysical testing and automated scoring. Here, we report a preliminary evaluation of the AIM (Angular Indication Measurement) and FInD (Foraging Interactive D Prime) color vision tests, part of a tablet-based diagnostic suite offered by PerZeption (Boston, MA). We also ran a full battery of standardized color tests on the subjects and specifically compared results from the AIM and FInD with those from the "trivector" test of the Cambridge Colour Test (CCT), and the Mollon-Reffin test to determine their relative effectiveness in detection and classification of color vision deficiencies (CVDs). These tests were chosen for comparison because they each quantify discrimination along the individual cone axes. Initial diagnosis was determined with an anomaloscope. There were two AIM paradigms, a detection task in which participants identified the gap of Landolt C stimuli and a discrimination task in which the orientation of bi-partite stimuli was identified. For the FInD detection paradigm, participants chose those boxes out of an array of boxes in which they detected a color blob. For the FInD discrimination task, they chose boxes out of the array in which two blobs of color were discriminable in color. AIM and FInD color threshold estimates were recorded for observers with normal color vision and with CVD. As expected, congenital red-green color deficient subjects had significantly higher thresholds on L and M cone and L-M (AIM and FInD) tests than did the control group. Results of both of these tests correlated reasonably well with those from the CCT and the Mollon-Reffin tests. Our results suggest that AIM and FInD provide reliable CVD diagnostics.

33.341 ALTERATIONS IN RESTING-STATE FUNCTIONAL CONNECTIVITY IN CHARLES BONNET SYNDROME

Aysha N. Kinakool[1] (ayshak@yorku.ca), Stefania S. Moro[1], Remy Cohan[1], Sara A. Rafique[2], Jennifer K.E. Steeves[1]; [1]Centre for Vision Research and Department of Psychology, York University, Canada, [2]Defence, Science and Technology Laboratory, Department of Defence, UK

Charles Bonnet Syndrome (CBS) is a debilitating phenomenon where individuals experience complex visual hallucinations secondary to vision loss, e.g., age-related macular degeneration, and glaucoma. Altered resting-state networks may contribute to the visual hallucinatory manifestations of CBS. Additionally, CBS symptoms may arise from increased glutamatergic and decreased γ-aminobutyric acid (GABA) receptor activity in areas surrounding deafferented cells in visual cortex. The current study examined functional connectivity between resting-state networks and visual cortex GABA+ and glutamate (Glx) concentrations in CBS. A CBS participant was compared to healthy age-matched controls. A multi-band multi-echo resting-state fMRI sequence and seed-to-target analysis of network connectivity was performed. A MEscher-GArwood Point RESolved Spectroscopy (MEGA-PRESS) sequence was performed with a 25 mm3 voxel placed medially in the visual cortex to quantify GABA+ and Glx concentrations. The CBS participant showed changes in connectivity (both increases and decreases) within the salience network (SN), default mode network (DMN) and visual network (VN). For example, there were decreases in connectivity of the VN with the medial pre-frontal cortex in the DMN; decreases in connectivity with the precuneus in the VN; and decreases in connectivity with the superior temporal gyrus and an increase with the fusiform gyrus within the SN. There was no change in GABA+ or Glx concentrations in V1 between the participant with CBS and controls. Our findings of functional cortical changes but no neurometabolite changes in the CBS participant suggest network level alterations in CBS which could account for the experience of their visual hallucinations.

33.342 SYNESTHETIC COLOR MAPPING OF CHINESE CHARACTERS AND KANJI: COMPARATIVE ANALYSIS AMONG GRAPHEME-COLOR SYNESTHETES IN TAIWAN AND JAPAN

Chien-Chun Yang[1] (r06227142@ntu.edu.tw), Huan-Wei Lin[1], Daisuke Hamada[2], Jun Saiki[2], Su-Ling Yeh[1,3]; [1]National Taiwan University, [2]Kyoto University, [3]National Humanities Center

Grapheme-color synesthetes perceive colors in response to colorless characters, and these synesthetic colors are influenced by specific rule-based linguistic factors. While previous research predominantly focused on the alphabetic writing system, the effects of linguistic factors on synesthetic colors in logographic systems, such as Chinese and Kanji characters, remains unclear. In this study, Taiwanese and Japanese grapheme-color synesthetes were recruited to identify synesthetic colors for 179 shared characters using the Munsell color system. A total of 15,931 character pairs were generated, and relationships based on pronunciation, semantic radicals, and phonetic radicals were examined. Color Similarity Indexes were calculated using negative-transformed z-scores in the CIE Lab* space to quantify color similarity. The linear mixed effect model showed that, for both Taiwanese and Japanese synesthetes, synesthetic colors were more similar when characters shared (1) the same pronunciation, (2) the same semantic radicals, and (3) the same phonetic radicals, compared to when they lacked these shared characteristics. Notably, the regularity effect, wherein phonetic radicals share the same pronunciations as the characters they are embedded in, exerts a more pronounced influence on the synesthetic colors perceived by individuals in Japan as opposed to Taiwan. Our research underscores the intricate link between grapheme-color synesthesia and language processing in Chinese characters/Kanjis, revealing how language use context shapes the perception of both individual elements and their compositions, resulting in the distinctive colors experienced by synesthetes in different countries.

This work is supported by the National Science and Technology Council in Taiwan (MOST 110-2410-H-002-130-MY3 and MOST 111-2223-E-002-008)

33.343 A BLUE-LIGHT ABSORBING LENS IMPROVES VISUAL FUNCTION UNDER BRIGHT LIGHT CONDITIONS IN PSEUDOPHAKIC PATIENTS

Jacob B. Harth[1] (jbh46589@uga.edu), Lisa M. Renzi-Hammond[1], Billy R. Hammond[1]; [1]University of Georgia

Introduction: This study evaluated the visual performance of a blue-light filter (BLF). The following visual parameters were assessed: glare disability (GD); chromatic contrast (CC); two-point light thresholds; and dysphotopsia symptoms (halo and spoke diameter). Methods: Twenty-five pseudophakic patients previously implanted with a 'UV-only' intraocular lens (IOL) (cut off frequency at 397nm with no visible spectra filtering) were tested using a single-masked randomized cross-over design. In the control condition, a similar UV-only filter (not filtering any visible wavelength) was used. In the test condition, a BLF was used (matching the absorbance profile of a common BLF IOL; Acrysof Natural and Clareon; Alcon Research Ltd). The Intensity of broad-band xenon light (annulus) and a sky-blue background needed to veil a central grating target was used to measure GD and CC, respectively. The diameter of spokes/haloes was determined by measuring their lateral extent induced by a bright white point source. Two-point light thresholds were measured as the separation between two small relatively intense point sources of light (Rayleigh criterion). Results. All of the visual functions measured favored the BLF when compared to the UV-only filter ($p<0.001$). Conclusion: About 6% and 10% more light energy was needed to occlude the grating stimulus for GD and CC. The light spread measured using the two-point technique was about 26% less for the BLF. The BLF reduced haloing and spokes by about 24% and 19%. The BLF significantly improved function across a variety of visual indices.

Alcon Research Ltd sponsored this investigation.

Sunday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

Development: Natural experience and eye movements

33.344 GENERALIZATION OF IMPLIED MOTION TO REAL MOTION IN INFANCY

Riku Umekawa[1] (ichiyu034@gmail.com), So Kanazawa[2], Masami K. Yamaguchi[1]; [1]Chuo University, [2]Japan Women's University

We investigated whether adaptation from implied motion (IM) was transferred to real motion by using optokinetic nystagmus (OKN) in infancy. We examined whether viewing a series of images depicting motion altered the OKN response to Random dot kinematogram (RDK) in infants. The RDK were presented 10 times in pretest, followed by 10 trials of IM adaptation and test during which we measured infants' OKN responses. The coherence of RDK was 50% in Experiment 1 and 80% in Experiment 2. In the pretest, the signal dots of RDK moved left or right. In IM adaptation, 10 randomly selected leftward (or rightward) IM images were presented at the center of the monitor at a rate of 600 ms per image (with no ISI). During test, RDK was presented immediately following the last IM image. The test RDK and IM images were in the same directions. An observer, who was blind to the direction of motion, judged the direction of OKN. We calculated the number of matches of OKN responses of each RDK direction as a "match ratio of OKN." Forty from five to eight months infants in Experiment 1 and sixty from three to eight months infants in Experiment 2 participated. We conducted a two-way mixed ANOVA, with age group as between-participant factors and adaptation (pretest and test) as within-participant factors. In Experiment 1 we found that IM adaptation reduced the OKN response to RDK direction for 7-to-8-month-olds, but not for 5-to-6-month-olds. In Experiment 2 we found that IM adaptation reduced the OKN response to RDK direction for all infants. That is, we found that around these ages, OKN responses to RDK was altered by viewing a series of images depicting motion. These results suggest that 7- to 8-month-old infants were be able to detect the motion information in static IM images.

33.345 THE DEVELOPMENT OF INTERNAL NOISE

Daphné Silvestre[1] (daphne.silvestre@mcgill.ca), Clara Marty[1], Rémy Allard[2], Armando Bertone[1]; [1]McGill University, [2]Université de Montréal

Most behavioral visual development studies have focused on cortical processes, without concurrently investigating pre-cortical function. This study used a novel internal noise paradigm (Silvestre et al., 2018) to estimate calculation efficiency and equivalent input noise (EIN) due to either the amount of light detected by photoreceptors (i.e. photon noise) or internal noise occurring at a cortical level for both static and dynamic information at different developmental periods. Thirteen children (11-13 years, mean= 11.9), fifteen adolescents (14-17 years, mean= 15.6) and fourteen adults (19-39 years, mean= 25.5) participated in this study. All participants completed a 2AFC task to measure contrast thresholds to drifting (2, 7.5, 15 and 30 Hz) and static gratings (0.5 cpd) with and without external noise for different luminance intensities (5-519 Td). The EIN associated with cortical noise significantly differed between the children and adults for the dynamic detection task ($p<.05$), but not for the static detection task; photon noise did not significantly differ between the children and adults. Calculation efficiency significantly differed between the children and adults ($p<.05$) for both detection tasks. Regarding the EIN associated with the amount of light detected by photoreceptors, the photon noise reached adultlike levels for the children. However, the calculation efficiency for both detection tasks was significantly lower for the children compared to adults. These results suggest that the cortical noise limiting the processing of a detection task reached adult-like levels for the children for static stimuli, but not for dynamic stimuli, which reached maturity during adolescence and that the maturity of the visual system is reached earlier at the photoreceptor level than at the cortical level.

33.346 CHARACTERIZING THE STATISTICS OF NATURALISTIC VISUAL EXPERIENCE DURING HEAD-FREE FIXATIONS IN INFANCY

Zachary Petroff[1] (zpetroff@iu.edu), Mana Agrawal, Zoran Tiganj, Stephanie Biehn, Sarah Freeman, Kathryn Bonnen, T Rowan Candy, Linda Smith; [1]Indiana University

Purpose: Humans start to learn about and interact with their visual environment over the first months after birth. At the same time, this visual experience tunes the receptive fields of neurons across the visual field. The goals of this study were to determine the characteristics of fixations in natural viewing during infancy and the distribution of contrast around these fixation points in the visual field. Methods: Participants aged 2-12 months wore head-mounted scene and binocular eye-tracking cameras (a modified Pupil Labs Core system) while engaging in naturalistic behavior in an 8ftx8ft home-like lab environment. Binocular fixations ≥400ms were identified in the eye movement recordings using a dispersion algorithm, and then RMS contrast was averaged around these fixation points for each infant and age group (2-3(n=15), 5-6(n=31), 8-9(n=23) & 11-12(n=16) months). Results: The median fixation duration increased minimally with age from 0.51s at 2months to 0.55s at 11months ($p=0.025$). Additionally, average RMS contrast decreased monotonically with eccentricity from the point of fixation for all age groups, across the analysis radius of 20deg. Interestingly, within a local region of 3deg from the point of fixation, the slope of RMS contrast remained uniform for infants aged seven to twelve months but decreased for the younger age groups ($p\leq0.0001$). Conclusion: For younger infants, the fixation point is, on average, the point of highest local contrast. For older infants, the impetus behind fixation becomes more complex, leading the contrast levels around the point of fixation to be more evenly distributed over time. This result is consistent with the immature spatial vision of the youngest infants and highlights the qualitative change in their visual experience over the first postnatal months. Of note, the duration of fixations did not change dramatically with age.

NIH-NEI R01EY032897

33.347 INFANTS' USE OF EYE MOVEMENTS TO EXPLORE THEIR NATURAL ENVIRONMENT

T Rowan Candy[1] (rcandy@iu.edu), Stephanie Biehn[1], Sarah Freeman[1], Adam Dalessandro[1], Victoria Tellez[1], Bhagya Marella[1], Khushboo Singh[1], Zachary Petroff[1], Kathryn Bonnen[1], Linda Smith[1]; [1]Indiana University

Purpose: Human infants start to learn about and interact with their environment during their first postnatal months. Immaturities in their gross motor responses and spatial vision constrain their visual behavior during this period of rapid development. The goals of this study were to determine the distribution of their saccadic viewing behavior and fixations in depth in a head-free naturalistic setting. Methods: Participants aged 2-15 months wore head-mounted scene and binocular eye-tracking cameras (a modified Pupil Labs Core system) while engaging in naturalistic behavior in an 8ftx8ft home-like lab environment. Calibrated saccadic movements were identified in each recording (3 to 20 mins) using an approach based on Engbert & Mergenthaler (2006) and depths of fixations using an empirically calibrated MIDAS algorithm. Results: Recordings from infants aged 2-3(n=15), 5-6(n=31), 8-9(n=23), 11-12(n=16) and 14-15(n=11) months were analyzed. The distribution of eye positions relative to the head was tighter for the younger infants, particularly in the vertical direction, which was also reflected in the pattern of saccade amplitudes (Levene Test all $p<0.0001$). The distribution of depths being fixated by the infants in this environment was best described by 2 kernels (KDE) at each age with the probability at the closer distance (approx. 40cm) increasing between 2 & 6 months. Conclusions: The youngest infants have limited head and trunk control and exhibited the most restricted range of eye movements, suggesting they are not compensating for their limited mobility by shifting their gaze on the timescale of these recordings. This likely leads to less active sampling of visual information, slower rates of change in the input, and a tighter link between head- and eye-centered frames of reference for vision. Maturation of motor responses and vision appear to result in a greater range of ocular motor exploration over the first postnatal months.

NIH-NEI R01EY032897

33.348 LOCOMOTION THROUGH SURPRISING ENVIRONMENTS: AGE EFFECTS ON GAZE GUIDANCE AND OBJECT MEMORY

Sophie Meißner[1], Jochen Miksch[2], Lena Würbach[2], Sascha Feder[3], Sabine Grimm[2], Wolfgang Einhäuser[2], Jutta Billino[1,4]; [1]Justus Liebig University Giessen, Experimental Psychology, [2]Physics of Cognition Group, Chemnitz University of Technology, Chemnitz, Germany, [3]Cognitive Systems Lab, Chemnitz University of Technology, Chemnitz, Germany, [4]Center for Mind, Brain and Behavior, University of Marburg and Justus Liebig University Giessen

Gaze behaviour during locomotion must balance sampling relevant information and the need for a safe and stable gait. Aging might affect this balance due to an increased risk of falling as well as decreased information seeking. Here we investigated how violations of expectations affect gaze behaviour and information uptake across age groups. We studied well-characterized samples of younger (20-40 years, N=16) and older adults (60-80 years, N=12) using a virtual reality (VR) paradigm. Participants moved through a reconstruction of a real-world university hallway wearing a VR headset. Locomotion was performed without an additional task. We placed semantically congruent (e.g., chairs) and incongruent target objects (e.g., a tuba) in the hallway to investigate how gaze behaviour during locomotion is affected by violated expectations. After the locomotion task, the participants were asked to freely recall any objects from the virtual hallway. We analyzed dwell times on objects indicating gaze behavior and object memory performance indicating information uptake. Results showed a robust attraction of gaze by incongruent objects. Dwell times on incongruent objects were increased in both age groups, however, to a larger extent in older adults. While older adults showed lower dwell times on congruent objects in comparison to younger adults, dwell times on incongruent objects were similar across age groups. Consistently, memory performance was boosted for incongruent objects and this benefit was more pronounced in older adults. Our findings indicate that surprising information attracts attention during locomotion - particularly in older adults. This attentional priorization goes along with enhanced encoding of information which could be considered a particular advantage in old age when memory resources are challenged. However, during actual locomotion in the real world such a benefit in information processing might come at the cost of reduced gait safety if processing ressources are shifted away from maintaining a stable gait.

This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), Collaborative Research Centres SFB/TRR 135: Cardinal Mechanisms of Perception, project number 222641018 and SFB 1410: Hybrid Societies, project number 416228727.

33.349 THE POST-STIMULUS MODULATION OF THE SACCADE RATE REFLECTS INHIBITORY CONTROL AND TOP-DOWN EXPLORATION IN YOUNG AND AGING POPULATIONS

Orit Shdeour[1] (oritshdeour@gmail.com), Shlomit Yuval-Greenberg[1]; [1]Tel Aviv University

The growth in the world's aging population has made it increasingly important to understand the cognitive and perceptual changes associated with aging. Studies have suggested that exploratory behavior and motivation decline with age. Here we examine the hypothesis that the characteristics of changes in exploratory behavior can be revealed through examining the saccade rate modulation in response to a new stimulus. It is a replicated finding that following the abrupt presentation of a stimulus, saccades are inhibited for approximately 100-200 ms and then they rebound, before returning to baseline. The post-stimulus inhibition is modulated by low-level features of the stimulus such as contrast and rarity and was suggested to be linked to midbrain inhibitory processes. The rebound phase is hypothesized to represent the initiation of the drive to explore, following the release of inhibition. Young (age 18-35, n=13) and older (age >60, n=12) healthy participants were presented with images of familiar and unfamiliar objects. The objects were first presented during a study phase, and then again during a test phase. The analysis focused on two oculomotor metrics: duration of inhibition and amplitude of the rebound; the first, hypothesized to reflect inhibitory control, and the second to reflect a top-down drive to explore. Analyzing post-stimulus saccade rates revealed that age affected inhibition duration but not the rebound: inhibition duration was found to be shorter in older relative to younger participants, indicating a lower inhibitory control in the aging population. Object familiarity affected the rebound amplitude and not the inhibition, but only in younger participants. This suggests that high-level stimulus characteristics modulate the top-down drive to explore and that this effect deteriorates with aging. We conclude that exploring the post-stimulus modulation of the saccade rate could contribute to the understanding of the exploratory drive and other cognitive functions in different populations.

the study was funded by ISF grant 1960/19 to S-Y.G and by The Edmond De-Rotschild foundation (IL)

33.350 MODULATION OF SACCADE-RATE IN INFANTS DURING THEIR FIRST YEAR OF LIFE

Shlomit Yuval-Greenberg[1,2] (shlomitgr@tau.ac.il), sheer Wolf[1], Yael Ben-Shalom[1]; [1]School of Psychological Sciences, Tel Aviv University, [2]Sagol School of Neuroscience, Tel Aviv University

As predominantly visual creatures, humans rely on visual exploration to study their environment. This behavior undergoes a substantial development during the first year of life, shifting from a disorganized sampling process to an organized, top-down-controlled one. Research on the development of visual exploration focused mainly on its spatial, rather than temporal aspects: where rather than when the eyes move. In adults, it is known that visual stimulation leads to an inhibition of saccades, followed by a large rebound of saccade rate before returning to baseline. The characteristics of this saccade-rate modulation are affected by both top-down and bottom-up factors. Here, we examine, for the first time, the developmental course of the temporal dynamics of saccade rate. In one experiment, three groups of participants, 3, 6, and 12 months old (N=20 each), were presented with black-and-white checkerboards, shortly followed by a video of a face. Findings showed that saccade rate was modulated in response to the checkerboard stimulus, in all three groups of infants. In contrast, the saccade rate was modulated by filmed faces only in infants 6 months old and above. In a second experiment, we presented 3 and 12-month-old participants (N=12 in each group), with black-and-white drawings of meaningful and abstract objects. Findings showed a pronounced modulation of saccade rate for both types of images in infants of both groups. An effect of object meaningfulness (higher rebound for meaningful vs. abstract objects) was found only for older infants. We conclude that the temporal modulation of visual exploration is present in very early infancy but only when stimuli are simple and of high contrast. Saccade-rate modulation reflects top-down recognition of semantic contents in 12-month-old, but not younger, infants. We suggest that studying saccade-rate modulation is a promising novel approach for studying the development of perceptual and cognitive processes in infancy.

The study was funded by ISF grant 1960/19 to S-Y.G

33.351 DEVELOPMENTAL TRAJECTORY OF GAZE DURING NATURAL LOCOMOTION

Sara Schroer[1] (saraschroer@utexas.edu), Nathaniel Powell[1], Daniel Panfili[1], Mary Hayhoe[1]; [1]University of Texas at Austin

Although much is known about visual and motor development, little work has been done on how vision guides actions in young children. Recent work in adults uses novel technology that integrates gaze, body, and terrain data (Muller et al., 2023). Similar integrated systems are now possible with children. To study the development of visually guided locomotion, children (1-to-6-years-old) wear head-mounted eye trackers (Pupil Labs Neon) while walking on various terrains, including a sidewalk and loose pebbles, while simultaneously recording with a fixed ZED 3d-scene camera to estimate terrain structure and a skeleton of the walker in the terrain. Work in adults shows that walkers are sensitive to terrain, modifying gaze distribution, step speed, and step length (Matthis, 2018). Similarly, data from an adult, 2-, 3-, and 4-year-old found that where children looked was sensitive to terrain, though only the 3-year-old changed their speed. The 4-year-old and adult in the two terrains were similar in gaze distribution and speed to previous adult data in flat terrain (1.38±0.05m/s; Matthis, 2018), looking up while walking on the sidewalk (adult 61% of the time, 4-year-old 56%) and far ahead on the ground on the pebbles (adult 78%, 4-year-old 84%). The 3-year-old looked far ahead on the ground while on the sidewalk (61%) but slowed down and looked down near their body on the pebbles (66%; 0.96m/s to 0.78m/s). The two-year-old moved slowly on both terrains (0.36m/s). Although 2-year-old's gaze was distributed widely in the sidewalk condition, they mostly looked down near their body on the pebbles (54%). Additionally, the 2-year-old and 3-year-old showed awareness of the difficulty of the pebbles (e.g., asking for a hand). While these results are preliminary, they suggest that locomotion in young children, like adults, is controlled by complex decision mechanisms that take both costs and sensorimotor uncertainty into account.

NIH Grant EY05729

SUNDAY MORNING POSTERS IN PAVILION

Sunday, May 19, 8:30 am – 12:30 pm, Pavilion

Object Recognition: Neural mechanisms

33.401 BEHAVIORAL DETECTABILITY OF ELECTRICAL STIMULATION OF INFERIOR TEMPORAL NEURONS IS EASIER IN THE PRESENCE OF THEIR PREFERRED VISUAL STIMULI

Reza Azadi[1] (reza.azadi@nih.com), Timothy Ma[1], Emily Lopez[1], Josh Ebel[1], Arash Afraz[1]; [1]Laboratory of Neuropsychology, National Institute of Mental Health, NIH, Bethesda, MD 20892, USA

We have recently shown that detectability of cortical stimulation in high-level visual areas, such as inferior temporal (IT) cortex in monkeys, varies by the content of the visual input presented during stimulation. However, the neurophysiological mechanisms underlying this interaction remain unexplored. Since IT neurons respond differentially to various visual stimuli, it is possible to hypothesize a simple explanation: visual stimuli that maximally activate the neurons, result in a narrower dynamic range for the impact of artificial stimulation, making the behavioral detection of stimulation more difficult. This hypothesis predicts a negative correlation between image-induced neural activity and behavioral detectability of neural stimulation. To study this interaction, we implanted a multi-electrode Utah Array in central IT cortex of a macaque monkey, allowing recording and stimulation in the same sites across multiple sessions. The animal learned to detect and report IT cortical stimulation while fixating on different images. We also collected the neural responses while the animal passively fixated on a series of visual stimuli. The results revealed a strong positive correlation ($r = 0.81$, $p = 0.001$) between the neural responsiveness to visual stimuli and detectability of artificial cortical stimulation. Contrary to the simple hypothesis mentioned above, the behavioral performance for detection of cortical stimulation is better when the visual input induces higher neural activity in the simulated site. Additionally, we included trials in which multiple electrodes simultaneously delivered electrical stimulation. The results revealed a positive correlation between the average visual response across stimulated sites and the detectability of multi-electrode cortical stimulations. These findings suggest that the interaction between the neural activity evoked by visual input and cortical stimulation is complex, and that perturbability of a neural state might be dampened in the absence of its corresponding visual stimuli.

33.402 NETWORK MECHANISMS OF ONGOING BRAIN ACTIVITY'S INFLUENCE ON CONSCIOUS VISUAL PERCEPTION

Yuan-hao Wu[1], Ella Podvalny[1], Max Levinson[2], Biyu J. He[1,3,4,5]; [1]Neuroscience Institute, New York University Grossman School of Medicine, New York, NY 10016, U.S.A., [2]McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montréal, Québec, QC H3A 2B4, Canada, [3]Department of Neurology, New York University Grossman School of Medicine, New York, NY 10016, U.S.A., [4]Department of Neuroscience & Physiology, New York University Grossman School of Medicine, New York, NY 10016, U.S.A., [5]Department of Radiology, New York University Grossman School of Medicine, New York, NY 10016, U.S.A.

Spontaneous brain activity is energetically expensive and spatiotemporally organized in an intricate manner with important clinical implications. Yet, little is known regarding how spontaneous brain activity participates in online, task-oriented brain functions. While previous work has demonstrated that prestimulus ongoing activity can predict task performance from trial to trial, the underlying mechanisms remain elusive. Here, we systematically investigated prestimulus ongoing activity's influences on visual perceptual decision-making and conscious object recognition. We employed whole-brain 7 Tesla fMRI data acquired during a threshold-level visual object recognition task in 25 healthy human subjects. Our objective was to dissect the influences of prestimulus brain activity from distributed cortical and subcortical brain regions on multiple facets of perceptual behavior, including the sensitivity and criterion of conscious object recognition, and discrimination accuracy in a categorization task. To shed light on the mechanisms linking prestimulus ongoing activity and perceptual behavior, we further investigated how prestimulus activity modulates stimulus-related processing. Our findings reveal a diverse set of effects on perceptual behavior exerted by prestimulus ongoing activity originating from distributed brain regions. High prestimulus activity in the ventromedial prefrontal cortex enhances sensitivity and promotes a more conservative criterion in object recognition by reducing the trial-to-trial variability of distributed stimulus-triggered responses. Prestimulus activity in the cingulo-opercular and visual networks had opposite influences on recognition-related criterion and discrimination accuracy, with prestimulus visual network activity modulating the variability and stimulus encoding in sensory-evoked responses, and prestimulus cingulo-opercular network activity exerting a pattern of influences consistent with the modulation of tonic alertness. In sum, our study sheds light on the intricate contributions of spontaneous brain activity from distributed brain networks to perceptual decision-making and conscious visual perception. Our findings further illuminate how prestimulus activity from these distributed brain regions shapes multiple aspects of stimulus-related processing, providing concrete mechanistic insights into these behavioral effects.

This work is supported by the National Institute of Health (R01EY032085)

33.403 CONSERVATION OF CORTICAL CROWDING DISTANCE IN HUMAN V4: A REPLICATION AND EXTENSION

Jan W. Kurzawski[1] (jan.kurzawski@gmail.com), Brenda Qiu[2], Denis G. Pelli[1], Jonathan Winawer[1]; [1]New York University, [2]University of Washington

Crowding is the failure to recognize an object in the presence of other objects. Crowding distance is the minimum object spacing for recognition, and varies widely across healthy adults. The biological bases of this variation are unknown. Previously (VSS 2021), we reported that in a group of 26 participants, the square of crowding distance (i.e., "crowding area") correlates inversely with the surface area of hV4, but not V1, V2, or V3. Those results were exploratory and therefore potentially subject to selection bias. METHOD. Here, in a confirmatory study, we doubled the sample size and added reliability measures of crowding distance and map surface area. In 50 observers, we measured crowding on a letter recognition task and surface area from a retinotopic fMRI experiment. RESULTS. Reliability was high for both measures. Test-retest measures of the Bouma factor, b = (crowding distance)/eccentricity, correlated highly across two sessions ($r = 0.94$). Surface area estimates of visual maps, based on boundaries drawn independently by two researchers, were also highly correlated ($r = 0.94$, 0.88, 0.74, 0.73 for V1, V2, V3, hV4, respectively). These reliability checks confirm that we can accurately measure individual differences. We then replicated our prior analyses and confirmed that the square of crowding distance scales inversely with the surface area of hV4 ($R2 = 0.43$) but not V1, V2, V3 ($R2 < 0$). CONCLUSION. From the relationship between crowding distance and surface area, we estimate a recognition-threshold letter spacing of 1.5 mm on the hV4 map. That spacing is conserved across observers, despite the two-fold variations in crowding distance and visual map size. Our finding constrains the brain locus of crowding, a key step in a computational account of object recognition.

Support from NIH P30EY013079.

33.404 DIVERSE VISUAL FEATURE SELECTIVITY IS ENABLED THROUGH INHIBITORY FEATURE SURROUNDS IN DEEP NEURAL NETWORK MODELS

Christopher Hamblin[1] (chrishamblin@fas.harvard.edu), Talia Konkle[1], George Alvarez[1]; [1]Harvard University

Artificial neural networks are proving useful for understanding hierarchical visual processing, as we gain direct access to all of a model's internal workings, including not only the activation of every unit, but also the weights connecting these units. A new frontier of 'mechanistic interpretability' has emerged in machine learning, which seeks to understand not only what features are represented by the network, but how complex features are built from simpler ones through these connections. As in biological visual research, these approaches start by identifying interpretable features in the latent layers based on maximally activating images, such as 'edge-detectors', 'curve-detectors, 'object-parts' etc. Considering the weights between layers then yields a compositional explanation of feature construction —e.g. a rabbit detector is built from legs, ears and eye detectors; an eye from curves, a curve from line segments, etc. However, here we argue this 'composition-by-parts' account is fundamentally incomplete, because it fails to incorporate the role of inhibitory operations. Negative weights constitute half of the learned weights in a typical deep neural network. Where inhibition is conventionally treated as analogous to excitation —for example, by considering maximally inhibiting images — we show that the asymmetry introduced by the non-linear activation function (ReLU) necessitates distinct computational roles for excitation and inhibition. We put forward a theory of the 'inhibitory feature surround', in which inhibition enables the construction of diverse features with selective excitatory responses. We validate our account with a series of 'virtual lesioning' experiments on inhibitory connections. Lastly, we introduce a feature visualization technique designed to target a feature's inhibitory surround specifically, to help researchers understand the role of inhibition in particular cases. Broadly, these results provide clarity into the functional role of inhibition in deep neural network models, and offer a framework for empirical tests of inhibitory function in biological visual systems.

33.405 DISENTANGLING THE UNIQUE CONTRIBUTION OF HUMAN RETINOTOPIC REGIONS USING NEURAL CONTROL

Alessandro T. Gifford[1] (alessandro.gifford@gmail.com), Radoslaw M. Cichy[1]; [1]Freie Universität Berlin

Early- and mid-level retinotopic regions of the human ventral visual stream (V1 to V4) implement key stages of visual information processing. However, what aspects of the visual input each region uniquely encodes remains incompletely known. A major experimental roadblock in assessing each regions' unique role is that typically their activation profiles are highly correlated, hiding their respective contribution to information processing. Here we used a novel analytical approach to disentangle the unique contribution of each retinotopic region. We started by leveraging NSD, a large-scale fMRI dataset, to build encoding models of all retinotopic regions. With these models we predicted neural responses for >100k naturalistic images (coming from NSD/ImageNet). We then implemented two neural control algorithms to find images that maximally distinguished predicted responses between all pairwise region combinations, thus revealing their idiosyncratic computations. The first neural control algorithm determined images that maximally activated the univariate response of each region while maximally deactivating the univariate response of other regions. The second neural control algorithm used genetic optimization to select an imageset that decorrelated (r=0) the multivariate responses between regions, through representational similarity analysis. We cross-validated both algorithms across NSD subjects, resulting in quantitatively disentangled responses, particularly for non-adjacent regions. The controlling images showed consistent qualitative patterns such as texture frequency, color, and object presence. Finally, we collected EEG responses for the V1-V4 comparison controlling images. These images disentangled the univariate and multivariate EEG responses over time, showcasing the generalizability of the neural control solutions across neuroimaging modalities. In sum, our contributions are threefold: we provide new quantitative and qualitative findings on the unique computation of retinotopic regions; we propose novel neural control algorithms capable of disentangling univariate and multivariate representations within biological and artificial information processing systems; and we demonstrate how data-driven exploration promotes discovery in understudied regions of the brain.

A.T.G. is supported by a PhD fellowship of the Einstein Center for Neurosciences. R.M.C. is supported by German Research Council (DFG) Grant Nos. (CI 241/1-1, CI 241/3-1, CI 241/1-7) and the European Research Council (ERC) starting grant (ERC-StG-2018–803370).

33.407 TESTING THE POSSIBLE ORIGINS OF CATEGORY SELECTIVITY IN THE BRAIN WITH DNN MODELS

Bowen Zheng[1] (bowen27@mit.edu), Meenakshi Khosla[2], Nancy Kanwisher[1]; [1]Massachusetts Institute of Technology, [2]University of California San Diego

How do category-selective neural responses in the ventral visual stream arise? Maybe brains have specializations for faces, bodies, and places because these categories serve distinct post-perceptual processes like social cognition and navigation. Alternatively, simple exposure to natural visual input may suffice. The second hypothesis is hard to test in humans, but we can test its in-principle possibility by asking whether similar category selective responses arise in deep neural networks (DNNs) that have no domain-specific priors and know nothing about the meaning of these categories to humans (see also Prince et al., 2023). We trained DNN models unsupervised with contrastive embedding objectives on an ecologically representative dataset (Ecoset), and then used nonnegative matrix factorization to discover dominant components of the network's responses to natural images. These components included three with response profiles selective for places, faces, and food, respectively, as measured by correlations with a) responses of previously identified fMRI components from the ventral pathway (r = 0.6, 0.6, and 0.5, Khosla et al, 2022) and b) human ratings of the salience of places (r=.4), faces (.5), and food (.5) in the images. Thus, category selectivities could arise in brains from natural visual input statistics, without strong domain-specific priors. What properties of the training diet are critical for these category selectivities to emerge? To assess the role of color, we trained a DNN on grayscale Ecoset images, which largely retained selectivities to faces and places but not food. In contrast, a DNN trained on "cutout" images, with backgrounds removed, failed to develop robust selectivity for any of these categories, despite achieving similar performance on image categorization. These results suggest that while category selectivity could in principle emerge without domain specific priors, from mere exposure to natural visual environments, the presence of full scene context may play a crucial role.

We would like to thank NIH grant R01-EY033843 for their support of our research.

33.408 RELIABILITY OF FUNCTIONAL LOCALIZATION AND ACTIVATION PROFILES OF CATEGORY-SELECTIVE REGIONS USING FMRI

Olivia S. Cheung[1,2] (olivia.cheung@nyu.edu), Haiyang Jin[1], Lúcia Garrido[3]; [1]New York University Abu Dhabi, [2]Center for Brain and Health, NYUAD Research Insitute, [3]City, University of London

Regions of interest (ROI)-based approach is a popular method to study brain activity in cognitive neuroscience research. To what extent are functionally-defined ROIs reliable within individual participants? In a comprehensive, pre-registered fMRI study, we investigated reliability of ROI location and activation profiles across category-selective regions, including bilateral face-selective, body-selective, scene-selective, and object selective regions (FFA1, FFA2, OFA, ATL, pSTS, aSTS, FBA, EBA, PPA, OPA, RSC, and LOC). For each ROI, we examined the reliability of two localization measures: 1) coordinates of peak activation and 2) degrees of cluster overlap, and three activation measures: 1) response amplitudes, 2) selectivity amplitudes, and 3) response patterns. We compared the reliability measures of the ROIs defined within each participant based on either a p-value threshold ($p<0.001$) or an area threshold of contiguous vertices with the strongest selectivity (≤100 mm2). Because category-selective regions are commonly identified with either images or videos, where videos produce particularly robust activations compared with images in several ROIs, we also compared the reliability measures evoked by either format. All participants (N=30) completed two fMRI sessions viewing images or videos of faces, bodies, scenes, objects, and scrambled objects. Independent data sets were used to define the ROIs and to calculate reliability. Unsurprisingly, we found higher ROI reliability for all measures within- than between-sessions, and reliability was generally higher in the posterior than anterior regions. Notably, reliability was higher for all ROIs defined with images than with videos, with the exception of pSTS and aSTS. Moreover, reliability was consistently higher for ROIs defined with area thresholding than p-value thresholding methods. Together, these findings suggest that the reliability of the category-selective regions depends on specific cortical locations, ROI selection methods, and stimulus formats. Specifically, the use of image presentation and area thresholding to define the ROIs is recommended for most category-selective regions.

This work was funded by a New York University Abu Dhabi faculty grant (AD174) and a Tamkeen New York University Abu Dhabi Research Institute grant (CG012).

33.409 BEHAVIORAL DETECTABILITY OF OPTOGENETIC STIMULATION OF INFEROTEMPORAL CORTEX PREDICTS THE MAGNITUDE OF STIMULATION-INDUCED PERCEPTUAL EVENTS

Elia Shahbazi[1] (elia.shahbazi@nih.gov), Reza Azadi[2], Arash Afraz[3]; [1]National Institutes of Health, [2]National Institutes of Health, [3]National Institutes of Health

We have previously shown that the detectability of local cortical stimulation in the macaque inferior temporal (IT) cortex is markedly influenced by visual stimuli presented during stimulation. This suggests that the magnitude of perceptual distortions induced by cortical stimulation in a given site varies for various visual stimuli. To test this hypothesis, in this study we utilized our recently established technique, Perceptography, to develop photographic reconstructions of stimulation-induced perceptual events for various stimuli. Two adult macaque monkeys were trained to detect and report a brief optogenetic impulse delivered to their central IT cortex through an implanted LED array. Using Perceptography, we transformed seed images into "Perceptograms," specific images that trick the animal into reporting non-stimulated trials as stimulated (see previous work for details). First, we measured the animals' behavioral performance in detecting cortical stimulation while fixating on a set of seed images. Consistent with earlier results, we found that performance varied widely across different images. Subsequently, we calculated the degree of distortion in the Perceptograms acquired from each seed image. We found a strong correlation ($r=0.71$, $p<0.0001$) between the behavioral detectability of brain stimulation while viewing various seed images and the level of perceptual distortion caused by optogenetic stimulation of IT cortex. This result suggests that animals rely on visual perceptual distortions for behavioral detection of cortical stimulation. Moreover, these findings show that behavioral detectability of cortical stimulation while viewing a given image predicts the magnitude of the stimulation-induced perceptual event. This predictive power facilitates future exploration of the link between neural activity in the high-level visual cortex and visual perception.

National Institutes of Health NIMH ZIAMH002958

33.410 OBJECT SIZE AND DEPTH REPRESENTATIONS IN HUMAN VISUAL CORTEX

Mengxin Ran[1] (ran.55@osu.edu), Zitong Lu[2], Julie D. Golomb[3]; [1]The Ohio State University

One of the key abilities in human object perception is maintaining a reliable representation of an object’s real-world size across various distances and perspectives. Previous research has indicated that neural responses in the ventral temporal cortex reflect object real-world size representations. However, the stimuli used in these prior studies confounded two related properties: perceived real-world size and real-world depth (distance). Moreover, the stimuli did not include naturalistic backgrounds, preventing us from exploring visual mechanisms in more ecological conditions. Bridging this limitation, a recent study from our group conducted a model-based representational similarity analysis on EEG data from a large-scale dataset of subjects viewing natural images featuring objects of varying retinal sizes and depths. The EEG study successfully disentangled a distinct timeline of processing objects real-world size and real-world depth. To better understand object representations in human brain regions with better spatial resolution, our current study applies a similar analysis approach to fMRI data, aiming to explore how different parts of human visual cortex represent objects real-world size and depth information in natural images. Applying our model-based representational similarity analysis on the THINGS fMRI dataset, we isolated neural representations specific to real-world size, real-world depth, and retinal size across human visual cortex. We found the most robust real-world depth representations in scene-selective regions such as the Parahippocampal Place Area (PPA) and the Transverse Occipitial Sulcus (TOS), and the most robust real-world size representations in middle-level visual regions, such as the V4, V3A and V3B. Our study delineates how various regions in human visual cortex are involved in processing different object size and depth features via an advanced computational approach, which offers an insightful understanding of the human brain processing of object information within naturalistic images.

NIH R01-EY025648 (JG), NSF 1848939 (JG), Center for Cognitive & Behavioral Brain Imaging ADNiR scholar (MR)

33.411 CORTICAL DYNAMICS OF MATERIAL AND SHAPE PERCEPTION ACROSS ILLUMINATION AND VIEWPOINT

Alexandra C. Schmid[1], Hector O. Sanchez Melendez[1], Chris I. Baker[1]; [1]National Institutes of Health

Navigating our dynamic world, from traversing pebbly beaches to grasping slippery objects, demands an adept interpretation of visual information. For example, information about an object's material and shape are important for understanding its behavioral affordances. We hypothesize that cortical dynamics during object viewing will reflect visual regularities arising from the object's intrinsic properties, independent of viewpoint and illumination. To investigate this, we recorded magnetoencephalography (MEG) data while participants viewed 256 photorealistic object stimuli as part of an n-back task. These stimuli varied in material (e.g., rock, wool) and geometry (e.g. cuboid-, dome-shaped), and were presented in natural lighting conditions from different viewpoints. As part of these manipulations, the objects' surfaces varied in their complex mesostructure (medium-scale surface relief) and reflectance, leading to changes in in behaviorally relevant qualities, such as softness and roughness. Using the MEG sensor activation patterns and a linear classifier, we were able to decode both material and shape properties even across various viewpoints and lighting conditions. Subsequent analyses characterized links between distinct image information produced by these properties and the cortical dynamics observed. Our findings support the notion that cortical dynamics during object perception reflect affordance-related attributes inherent to an object's material and shape and enrich our understanding of how the perception of behaviorally relevant object properties unfolds.

This research was funded by a Walter Benjamin Fellowship from the German Research Foundation (DFG).

33.412 REASSESSING THE FOOD SELECTIVE COMPONENT IN HUMAN VISUAL CORTEX

Cyn Fang[1] (cxf@mit.edu), Meenakshi Khosla[1], Nancy Kanwisher[1]; [1]MIT

Three recent papers based on the Natural Scenes Dataset (Allen et al, 2022) reported that two bands of cortex in the ventral visual stream respond selectively to images of food (Khosla et al, 2022; Jain et al, 2022; Pennock et al, 2023). Khosla et al (2022) applied data-driven non-negative matrix factorization (NMF) to the NSD data and discovered a consistent component across participants, defined by a response profile over stimuli and a weight matrix over voxels, that correlated strongly with the salience of food in the image. Control analyses indicated that this "food component" responded more strongly to food than nonfood images matched for low-level visual features. Here we further tested the selectivity of this "food component" with new controlled stimuli in six new subjects. We developed a component localizer using a subset of 50 NSD images which minimize the variance of the inferred food component weights in new subjects. We used this localizer to infer the food component weights in new subjects, and then computed the response of this inferred food component to independent stimuli. Our findings replicated the broad anatomical localization of the food component, as well as the previously published findings of a greater response to food than nonfood in NSD images held out from the localizer. We further found that responses to these images were very similar when presented in greyscale. We also found a significantly higher response of this component to food than nonfood "reachspace" images (e.g., tabletops), though with a smaller effect size. However, we did not find a higher response to food than nonfood in "cutout" images, in which food and nonfood objects were isolated on a white background. These results suggest that the "food component" is not broadly selective to all food images, but its selectivity may depend on surrounding context.

Acknowledge NIH grant R01-EY033843

33.413 AGING DELAYS THE FORMATION OF OBJECT REPRESENTATIONS

Marleen Haupt[1], Douglas Garrett[2,3], Radoslaw Cichy[1]; [1]Freie Universität Berlin, Germany, [2]Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany, [3]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany

Aging commonly slows down behavioral task performance in humans. However, it remains unknown how aging affects the neural dynamics of processing visual information, required for most tasks. Here, we investigate age-related delays in the formation of object representations and the processing stages at which these delays manifest. For this, we measured EEG and fMRI in healthy younger (18-35 years, N=21) and older participants (60-75 years, N=22) while they viewed natural images. We applied three advanced multivariate techniques to analyze the formation of object representations. First, we decoded object identity in a time-resolved manner to quantify age-related processing delays. Our results reveal significant differences in decoding time courses between both age groups in an early time window (around 80-155 ms). Specifically, decoding time courses peaked significantly later (around 20 ms for different categorical comparisons) in older adults. These findings suggest delayed feedforward processing in aging. Second, we compared the representations underlying the observed delay in younger and older adults using RSA. Our results reveal that both age groups are characterized by time-shifted yet similar representations. Third, we linked the observed EEG patterns to fMRI signals in ventral visual stream regions through EEG-fMRI fusion. This allowed us to map temporal delays to cortical processing stages. We observed that age-related delays manifest in mid- and high-level regions (i.e., V3, LOC, FFA) rather than early visual areas (i.e., V1, V2). In summary, our findings reveal age-related changes in the temporal dynamics of object representations. By quantifying a temporal delay as well as linking it to mid-to-high level ventral visual cortex, our results provide a first nuanced view on the lifespan plasticity of visual object representations.

Attention: Features, objects 1

33.414 FLEXIBLE ALLOCATION OF FEATURE-BASED ATTENTION TO NARROW AND BROAD RANGES OF COLORS AS ASSESSED BY STEADY-STATE VISUAL EVOKED POTENTIALS

Mert Ozkan[1] (mert.ozkan.gr@dartmouth.edu), Angus F. Chapman[2], Viola S. Störmer[1]; [1]Dartmouth College, [2]Boston University

We can allocate our attentional resources to select a specific visual feature, such as the color red, and prioritize its processing. Can we also tune our attention to multiple feature values at the same time? This would be useful given that objects in the environment are often variable and include ranges of feature values (e.g., apples are not just a singular red, but often a mix of reddish colors). The current study examines how participants select narrow and broad ranges of colors, and how this selection affects early visual-cortical processing using EEG. Participants (N=24) continuously attended a set of target dots among distractor dots to detect a brief interval of coherent motion that occurred in a subset of targets (on 50% of trials) while ignoring distractors. The range of target colors was varied systematically to span narrower or broader parts of a perceptually uniform color wheel (20º vs 60º around a target color center), and the distractor dots were always chosen to be maximally distinct from the targets in that feature space (i.e., 180º away from the target center). To match task difficulty across conditions, motion speed thresholds were obtained separately for each individual and condition. We assessed early visual processing of the color arrays by flickering target and distractor dots at distinct frequencies (7.5 and 8.57 Hz) to elicit separable steady-state-visual evoked potentials (SSVEPs). Our results showed that the SSVEP power measured over the occipital cortex was enhanced for targets relative to distractors across both conditions; interestingly, this enhancement was not modulated by color range, suggesting that attention flexibly and efficiently increased color representations across narrow and broad color ranges at early stages of visual processing. Overall, these results indicate that feature–based attention can be adaptively tuned to varying feature values and enhance their neural processing efficiently.

This research was supported by National Institute of Mental Health (#1R01MH133689-01).

33.415 LOOKING FOR THE RED SHIRT: MEANINGFUL OBJECTS STRENGTHEN MEMORY AND ATTENTIONAL GUIDANCE

Yong Hoon Chung[1] (yong.hoon.chung.gr@dartmouth.edu), Jamal Williams[2], Viola Störmer[1]; [1]Dartmouth College, [2]University of California San Diego

The contents of visual working memory can incidentally guide attention towards matching features in the environment, and the magnitude of this effect scales with the strength of the memoranda (Williams et al., 2022). Other work has shown that working memory performance for colors is increased when they are part of meaningful objects (e.g., a blue shirt) compared to unrecognizable objects (Chung et al., 2023). Here, we test whether colors that are encoded as part of meaningful objects also guide attention towards a simple stimulus that matches this color, and whether the amount of guidance scales with memory strength. Participants were asked to remember colored silhouettes of real-world objects or scrambled versions of them and, after a short delay, indicate which color they saw using continuous reports. To assess attentional guidance, during the retention interval, a visual search display was presented and participants indicated the orientation of a single slanted line among vertical lines all embedded in separate colored circles. On half of the trials, all search colors were unrelated to the memory colors; on the remaining trials, one of the nontarget colors matched a color maintained in working memory. The difference in search response times between match and no-match conditions served as an index of attentional guidance. Across several experiments, we manipulated memory strength by changing encoding time, set size, and the recognizability (i.e., meaningfulness) of the stimulus. Consistent with previous research, we found that color memory was stronger with longer encoding, smaller set sizes, and for recognizable relative to unrecognizable objects. Importantly, the amount of guidance also scaled with memory strength. Broadly, this is consistent with recent studies suggesting that real-world contexts can enhance working memory for simple features and that interactions between working memory and attention depend on the representational fidelity of memories.

33.416 ON THE RELATIONSHIP BETWEEN TARGET-DISTRACTOR DISCRIMINABILITY AND SEARCH EFFICIENCY: THE CASE OF COLOR.

Alejandro Lleras[1] (alleras@illinois.edu), Zoe (Jing) Xu[1], Howard Tan[1], Yujie Shao[1], Simona Buetti[1]; [1]University of Illinois at Urbana-Champaign

Across six experiments, we studied the relationship between target-distractor discriminability and search efficiency, when the only differentiating feature between objects is color. We assessed target-distractor discriminability by measuring the distance in CIELab space between nearly equiluminant target and distractor stimuli, over a range where the distance in CIELab space measures the perceptual discriminability between two colors. We presented stimuli at either two or three different eccentricities, and we cortically magnified the stimuli to equate visual processing times. Finally, to prevent serial search, we selected color distances that were sufficiently large such that peripheral vision could accurately detect those color differences. Results from Experiment 1 demonstrated that there is an inverse relationship between logarithmic search efficiency and target-distractor color distance, as was proposed in Lleras et al. (2020). This relationship was confirmed in four additional experiments. Experiments 2-6 explored whether different stimuli conditions could break or modulate that relationship. Experiment 2 found no evidence that presenting distractor colors only to one side of the target's color on the color circle would improve search performance, as proposed by optimal tuning theories of feature-based attention and by relational tuning accounts. Performance was identical to when distractors were picked from both sides of the target color. In Experiments 3 and 6, blocking distractor colors resulted in a 50% improvement in search efficiency across the entire range of colors. In Experiment 4, target and distractor colors varied randomly from trial to trial. Search performance was no longer logarithmically related to set size and RTs were unusually elevated. Experiment 5 studied the improvements in search efficiency when the same color is repeated over the entire experiment. Across all experiments, we found no evidence that feature-based attention to color had altered the appearance of the stimuli, or that features had been boosted or suppressed to improve performance.

This material is based upon work supported by the National Science Foundation under Grant No BCS1921735 to SB.

33.417 ATTENTION INCREASES REPRESENTATIONAL DISTANCE NEAR TASK-RELEVANT ORIENTATIONS

Melissa Allouche[1] (melall@bu.edu), Angus Chapman[1], Rachel Denison[1]; [1]Boston University

Background: Attention can alter not just visual performance, but also the appearance of different stimulus features, such as contrast, spatial position, and color. Such perceptual effects can be characterized by their representational geometry: a framework that quantifies dissimilarities in the perceptual representations of different stimuli. Attention has been hypothesized to increase the representational distance (perceived dissimilarity between nearby orientations) around the attended stimulus, while decreasing distance around unattended stimuli. Goal: To date, studies have focused on how attention changes representations of specific target and non-target stimuli. Here we investigated how attention affects the entire representational space of orientation. Methods: On each trial, participants performed two interleaved tasks: a feature-based attention task and a triad similarity judgment task. In the attention task, participants attended to one of two spatially overlapping arrays of oriented bars (45° clockwise or counterclockwise from vertical) to detect small tilts in their orientations. In the triad task, participants judged the perceived similarity among an array of three Gabor patches that varied in orientation. To assess how attention affected the representation of orientation, we performed a multidimensional scaling analysis on the pairwise orientation distances estimated from the triad task under each attention condition. Results: Multidimensional scaling revealed a 1-dimensional manifold of orientation that was best represented in a 3-dimensional space, with two primary dimensions aligned to cardinal and intercardinal orientations and a third dimension that differentiated cardinals from intercardinals. Along this third dimension, attention significantly modulated orientation representations, with attention to -45° orientations expanding the representational space around that orientation. Attention to +45° led to a similar effect, though it was not significant. Conclusions: Our findings support the hypothesis that attention can distort representational geometries in perception. Attention altered participants' similarity judgments, exaggerating the distance between orientations near the attended angle.

Funding for this project was provided by Boston University startup funding to R.N.D.

33.418 OBJECT-BASED ATTENTION IS FLEXIBLE TO BOTH LOW- AND HIGH-LEVEL CHANGES IN REAL-WORLD OBJECTS.

Kelly McEvoy[1], Sarah Shomstein[1]; [1]Department of Psychological & Brain Sciences, The George Washington University

Attentional selection operates on an object-based representation using both simple geometric shapes (e.g., rectangles, trapezoids, letters, semicircles) as well as semantically meaningful objects in real-world scenes (e.g., a cup on a table). While previous work suggests that both low-level (boundaries) and high-level properties (object meaning) contribute to object-based attention, the relative contribution of these features remains unclear. Here, we characterize the relative contribution of a consistent object border (i.e., object outline) and consistent object semantic information (i.e., object identity) to object-based attentional selection. We used an adapted two-rectangle task in which following a brief exogenous cue (150 ms) and a delay (70 ms), the attended real-world object abruptly changed according to one of three conditions: an object with the same border but different semantic category, an object with a different border but same semantic category, or an 'outlier' object with neither the border nor category in common. We predicted that both objects with the same border or same category contribute to the perception of objecthood but to varying extents, while 'outlier' objects contribute to an overall smaller object-based effect compared to border and category. Robust object-based attentional effects were observed in conditions where border and semantic information were preserved. Additionally, object-based effects were also observed in the outlier condition, yielding a smaller magnitude compared to when the object's border or category was maintained. Overall, object-based attentional guidance was observed in all three conditions, but was modulated by the degree of low- and high-level changes in object properties. Our findings suggest that object-based attentional guidance persists despite altering the object properties in real-world objects, which can generalize to more complex, naturalistic environments.

NSF BCS 2022572 to SS

33.419 SEMANTIC RELATIONSHIPS BETWEEN SOUNDS AND IMAGES MODULATE ATTENTION EVEN WHEN THE STIMULI ARE TASK-IRRELEVANT

Kira Wegner-Clemens[1], Dwight J Kravitz[1], Sarah Shomstein[1]; [1]George Washington University

Semantic information plays an important but poorly understood role in guiding attention in naturalistic scenes. Semantic relationships among visual objects have been shown to modulate attentional priority, even in tasks where object identity is irrelevant. In an audiovisual context, semantically related sounds can improve search speeds for visual targets, with the benefit scaling with the degree of semantic relatedness. However, prior research almost exclusively focused on the targets defined by their identity, meaning the visual semantic information was task-relevant. Thus, it is unclear whether crossmodal semantic relationships influence attention only when they are task relevant, or whether those relationships play a more general role in attentional selection. In the present study, we investigate whether an audiovisual semantic benefit exists when both the image and sound's semantic information are task-irrelevant. Participants were presented with two images and a sound, then subsequently presented with two Gabor patches at the image locations and asked to identify whether the target Gabor was slanted clockwise or counterclockwise. On valid trials, when the sound matched the image where the target Gabor subsequently appeared, participants responded significantly faster than on invalid trials, when the sound matched the image at the location where a distractor appeared. The size of the validity benefit was modulated by the degree of semantic relationship between the sound and the other image. In a mixed effect model with relatedness and validity as fixed effects and subject, target location, and rotation direction as random effects, there was a significant interaction between semantic relatedness and validity, such that a stronger semantic relatedness between the unmatched image and sound results in a smaller validity effect. These results show that crossmodal semantic relationships guide attention even when task-irrelevant, suggesting that semantic relatedness plays a general role in guiding attentional selection.

NIH F31EY034030 to KWC; NSF BCS 1921415 to SSS; NSF BCS 2022572 to SSS & DJK

33.420 FEATURE SELECTIVITY IN THE FAR PERIPHERY

Ema Bidiwala[1] (ema.bidiwala@ttu.edu), Abigail Loddeke[1], Miranda Scolari[1]; [1]Texas Tech University

Humans can successfully recognize objects in the far periphery (30-80° of visual angle) despite limited acuity (Boucart et al., 2013; 2016; Baldwin et al., 2016; Rosenholtz, 2016). This ability may be supported by the prostriata, a brain region activated by stimuli at 60° periphery (Mikellidou et al., 2017), with retinotopic organization, and projections to the thalamus (Kurzawaski et al., 2020), similar to visual cortex. Notably, orientation selectivity generally serves as a building block for object recognition by critically supporting edge detection. However, it is unclear whether the same mechanisms are present in far peripheral sensory populations (e.g., akin to those in prostriata). Therefore we sought to determine if orientation selectivity was possible up to 90° eccentricity. Circular patches of moving or static square wave gratings were presented at 30, 60, and 90° at either cardinal (90° or 180°) or oblique (45° or 135°) orientations. Subjects (N = 30) reported whether motion was present followed by the orientation of the stimulus. Performance was above chance on motion trials across all eccentricities, and across all orientations presented below 90° eccentricity. In the orientation discrimination task, subjects performed better on cardinal orientation trials compared to oblique with the magnitude of the difference increasing with eccentricity. In a second experiment, we scaled stimuli to account for early visual cortical magnification (N = 31). This time, performance was above chance on all conditions, and interactions from Experiment 1 were no longer significant. However, we still observed a significant main effect of eccentricity and orientation. Our results suggest that despite degraded acuity in the far periphery, representations of visual orientation involved in edge detection are still maintained up to 90° periphery. These representations possibly serve as a building block for object recognition in the far periphery.

33.421 DECODING FEATURE-BASED ATTENTION IN VISUAL CORTEX

Rylee Faherty[1], Qiang Yang[1], Maeve R. Boylan[2], Andreas Keil[2], Mingzhou Ding[1]; [1]J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, [2]Department of Psychology, University of Florida

FMRI were recorded from participants viewing random dot kinematograms (RDKs) varying in either color or shape. In baseline conditions, participants attended a single feature array (gray circles, red circles, or gray squares) flickered at either 8.57 Hz or 12 Hz. In selection conditions, two superimposed RDKs differing in one feature dimension were flickered at two different frequencies (8.57 Hz vs 12 Hz), and the participant was instructed to attend color or shape. For all conditions, participants indicated whether they noticed a luminance change in the attended RDK array. We found that (1) for the baseline conditions, (a) comparing red circles against gray circles and gray circles against gray squares yielded no regions of significant differential activations and (b) MVPA decoding between red circles vs gray circles and between gray circles vs gray squares revealed above chance decoding accuracy in the entire visual hierarchy, suggesting that color and shape are broadly represented in the visual cortex in the form of linearly discriminable spatial patterns and (2) for the selection conditions, (a) comparing selecting red circles in the presence of gray circles against selecting gray circles in the presence of red circles and comparing selecting gray circles in the presence of gray squares against selecting gray squares in the presence of gray circles yielded no regions of significant differential activations and (b) MVPA decoding between selecting red circles vs selecting gray circles and between selecting gray circles vs selecting gray squares revealed above chance decoding accuracy in the entire visual hierarchy, suggesting that attention selection for color and for shape occurs broadly in the visual cortex in the form of linearly discriminable spatial patterns. Overall, our data support the idea that rather than modulating neural activities only in feature-selective visual areas, FBA modulates neural activity broadly in the visual cortex.

SUNDAY, MAY 19, 8:30 AM – 12:30 PM, PAVILION

Attention: Features, objects 2

33.422 PLEASE IGNORE THIS TITLE. SEARCHING FOR EVIDENCE OF EFFECTIVE ACTIVE SUPPRESSION USING A DOT FILTERING TASK

Andrew Rodriguez[1] (rodri818@msu.edu), Derrek T. Montalvo[1], Mark W. Becker[1]; [1]Michigan State University

There is strong evidence suggesting that feature-based attention can facilitate the selection of relevant items by feature enhancement. More recently a number of prominent papers have claimed that the suppression of irrelevant features also facilitates the selection of relevant items, however this evidence is less complete and more controversial. Here we present two experiments that fail to find evidence for active suppression. In both, participants performed a dot filtering task where they ignored one set of colored dots (distractor suppression) to detect the number of changes in luminance in a different set of colored dots (feature enhancement). On a random 25% of trials, the filtering task was interrupted by a visual search task. Search arrays consisted of clock face consisting of 12 letters, six of which appeared in a single color and the other six appeared in six different colors. Consistent with feature enhancement, our results show fast RTs when the target letter appeared in the filter task's attended color and slower RTs when half of the distractor letters appeared in the attended color. But we found no evidence for distractor suppression; neutral-colored targets were not detected more quickly when half the distractors matched the to-be-ignored color and RTs were not slower for targets that matched the ignore color than a neutral color. In Experiment 2, we altered the dot filtering task so that the to-be-ignored color remained constant while the to-be-attended color varied trial to trial. In addition, during the search task the attended color from the previous filter task never appeared as the target. These changes were done to reduce the benefits of target enhancement while maintaining the benefit created by suppression, thereby increasing the motivation to apply suppression. Nevertheless, our preliminary findings still show little evidence of effective active suppression.

33.423 THE TIME COURSE OF ACTIVATING, MAINTAINING, AND SWITCHING BETWEEN ATTENTIONAL TEMPLATES IN VISUAL SEARCH

Martin Eimer[1] (m.eimer@bbk.ac.uk), Gordon Dodwell[1], Rebecca Nako[1]; [1]Birkbeck, University of London

Research on task switching often focuses on stimulus-guided response selection and execution. In contrast, the processes associated with changing task settings involved in covert attentional control have been much less well studied. Here, we focus on the dynamics of activating and switching mental representations of target-defining features (attentional templates) during the preparation for visual search. We employed a new high-definition rapid serial probe presentation paradigm, where lateral "clouds" of multiple densely spaced dots in different colours are presented in rapid succession throughout the interval between target displays that contain a colour-defined target. By measuring N2pc components triggered by cloud probes that match a currently task-relevant colour, feature-specific search template activation processes can be tracked with very high temporal resolution. Participants prepared for and responded to a specific "early" target colour that appeared on a subset of trials, or switched to a different "late" target colour on other trials. In one experiment, the absence of the early target was the cue for switching to the late target colour. In a second experiment, an explicit stay/switch cue appeared when the early target was absent, indicating that the previous template should either be maintained or changed to a different colour. Results showed that target templates were activated and switched with remarkable speed and temporal precision, and in line with changes in task demands. They also provided new evidence for the simultaneous co-activation of multiple attentional templates.

This work was funded by a grant by the Economic and Social Research Council, UK (ES/V002708/1).

33.424 IS ATTENTION GONE WITH THE WIND: DOES MOTION WITHOUT CONTEXT CUE VISUOSPATIAL ATTENTION?

Genna Telschow[1] (genna.telschow@ucf.edu), Mark Neider[1]; [1]University of Central Florida

To account for cognitive processing delays in motion perception, visuospatial attention is shifted towards locations predicted by an object's motion and environment. Previous research demonstrates that a rotating object touching a surface shifts attention in the direction the object would move due to friction. In two studies using an attentional cuing paradigm we investigated whether motion information, without the environment context, would cue attention towards future target locations. For both experiments participants saw a rendered animation of a flag billowing (motion direction: leftward vs rightward) and after a stimulus onset asynchrony (SOA) (200ms Experiment 1; 300ms Experiment 2) a letter (H/N) appeared to the left or right of the animation. Participants were asked to press a key that corresponded with the letter's location. We predicted that response times would be faster when object motion was directed towards the letter (congruent) compared to away from the letter (incongruent). In Experiment 1 response times did not differ between congruent and incongruent trials for leftward ($t_{47} = .655$, $p = .52$) or rightward motion ($t_{47} = .764$, $p = .45$). Experiment 2 used an SOA of 300ms to determine if participants required more time to process potential predictive information from the within-object motion (non-translational motion). Again, response times did not differ between congruent and incongruent trials for leftward and rightward motion ($t_{50} = .233$, $p = .82$; $t_{50} = .427$, $p = .67$). Our results suggest that within-object motion, alone, is not predictive of future locations. This could be a result of differences in various objects' potential for movement. A rotating wheel, like a car tire, will typically move because of friction, whereas a billowing flag is usually secured to a pole and will not move. It's possible that motion must provide information regarding an object's potential for movement to shift attention.

33.425 EYE-TRACKING REVEALS ROBUST ATTENTIONAL FILTERING IN AN OBJECT-BASED ATTENTION TASK

Lasyapriya Pidaparthi[1] (lasyapriya.pidaparthi@vanderbilt.edu), Frank Tong[1]; [1]Vanderbilt University

We use object-based attention in our daily lives to process task-relevant objects, and sometimes we must ignore task-irrelevant objects, even if they are salient, dynamic, or situated in front of an attended object. We have previously shown the efficacy of eye-tracking at predicting the focus of object-based attention when participants must attend to one of two naturalistic objects (face, flower) that follow pseudorandom, minimally correlated trajectories while remaining partially overlapping (Pidaparthi & Tong, VSS 2023). Although smooth pursuit eye movements are considered by some to be strongly stimulus-driven, here we asked, how effectively can attention filter out the presence of a task-irrelevant object, as indexed by eye movements? To answer this question, we adapted our paradigm across two experiments. In Experiment 1, subjects were presented with either one or two moving objects and were instructed to respond whenever the task-relevant stimulus underwent brief spatial distortions (2 conditions: attend-face, attend-flower). We then evaluated the selectivity of object-based attentional filtering by using a sliding window correlation analysis to compare gaze trajectories with the attended stimulus trajectories. Notably, even with the overlapping sets of motion signals, pursuit eye movements were not perturbed by the irrelevant motion: observers could follow one attended object in the presence of the distractor object (mean r=0.581) just as accurately as a single object alone (mean r=0.579). In Experiment 2, we replaced the irrelevant object (e.g., flower during attend-face trials) with a moving Gabor stimulus that underwent random bursts of drifting motion (at 4 Hz for 500ms), and measured the extent to which this strong low-level motion signal influenced eye movements. In both instances, we found that observers can selectively attend to the task-relevant object such that gaze-following is unperturbed by extraneous motion signals, demonstrating the robustness of attentional filtering with respect to the eye movement system.

NIH R01EY029278 (FT) and NIH R01EY035157 (FT)

33.426 FMRI REVEALS A MODULATORY ROLE OF VISUAL FIELD MERIDIANS ON OBJECT-BASED SELECTIVE ATTENTION

David H. Hughes[1] (dhughes@mcw.edu), Adam J. Barnas[2], Adam S. Greenberg[1]; [1]Joint Department of Biomedical Engineering, Medical College of Wisconsin and Marquette University, [2]Department of Psychology, University of Florida

Previous behavioral results demonstrated a shift direction anisotropy (SDA) during object-based attentional selection that was modulated by visual field meridians (Barnas & Greenberg, 2016; 2019). We hypothesized that differential prioritization of horizontal and vertical target locations gave rise to the SDA due to attentional resources being divided differently by horizontal and vertical meridians. To test this, we measured BOLD signals during fMRI from visual cortex in 19 healthy adults. Subjects detected a target (letter T among distracting Ls) within an L-shaped object with two possible locations: vertex in upper-left (UL) or lower-right (LR) quadrant. Cue validity was 60% valid, 20 % invalid, 20% catch. Invalid trials were counterbalanced between shift orientation (horizontal and vertical) and whether target detection required a meridian crossing (crossing and non-crossing). The crossing SDA was significantly greater than the non-crossing SDA for the UL ($p = .039$, $d = 0.60$), but not the LR ($p = .454$, $d = 0.04$) object. A repeated measures ANOVA of cue-related activation revealed significant main effects of object location ($p < .001$) and retinotopic ROI ($p < .001$), as well as a significant interaction ($p = .007$). Post-hoc Tukey test for the UL object revealed no activation differences between valid and either non-crossing invalid-horizontal ($p = .963$, $d = 0.17$) or non-crossing invalid-vertical ($p = .969$, $d = 0.15$) locations, suggesting that attention pools within the cued quadrant. Compared to non-crossing, crossing activation was reduced (invalid-horizontal: $p < .001$, $d = 0.95$; invalid-vertical: $p < .001$, $d = 0.59$). This prioritization of each invalid location is consistent with the attentional spreading hypothesis (Richard et al., 2008). However, cue-related activation did not fully account for the SDA. Instead, the SDA could arise from differential spreading of attention resulting from stronger segregation of pools of resources by the vertical versus horizontal meridian.

This work was supported by grants from the US-Israel Binational Science Foundation (#2013400) and National Science Foundation (SBE 2122866).

33.427 DOES SELECTIVE ATTENTION UTILIZE PHYSICAL FEATURE VALUES OR PERCEPTUAL INTERPRETATIONS?

Vladislav Khvostov[1,2], Árni Gunnar Ásgeirsson[3], Árni Kristjánsson[1]; [1]Icelandic Vision lab, University of Iceland, [2]HSE University, Russia, [3]University of Akureyri, Iceland

Visual attention is thought to select task-relevant objects using attentional templates tuned to task-relevant features. When we look for strawberries, templates are tuned to red colors. But to what exactly are these templates tuned? Physical values (wavelength ~620-750 nm) or perceptual interpretations (what appears red due to color constancy mechanisms)? We addressed this using the priming of pop-out paradigm (Maljkovic & Nakayama, 1994). Observers (N=8) reported the tilt of an odd-one-out colored square (2AFC) presented along with two distractors. Participants performed mini-blocks (prime+test trials) within displays as seen through different color filters in a VR environment. We tested the nature of observers' representations on test trials by preserving perceptual color characteristics while changing the physical ones (via altering the filter color but keeping the object colors, “Perceptual” condition) or preserving physical values while changing the perceptual ones (via changing the filter and object colors to match the physical characteristics of the prime colors, “Physical” condition). Importantly, we rejected the assumption of perfect color constancy and introduced an adjustment task before the search task where participants reported the perceptual matches for all colors which we then used for the “True perceptual” condition. We calculated priming effects: RT differences between “Swap” test trials (target and distractor colors were swapped compared to the previous prime trial) and “Repeat” trials (target-distractors were repeated). All conditions produced significant priming but the magnitude was largest for the True perceptual condition (~70 vs. 40-50 ms). This shows that attention utilizes a color constancy mechanism - attentional templates are based on perceptual, not raw physical colors. This also highlights the importance of measuring true perceptual matches since color constancy is often far from perfect.

Supported by grant #228366-051 from the Icelandic Research Fund

33.428 SCAN PATTERN SIMILARITY PREDICTS THE SEMANTIC SIMILARITY OF SENTENCES ACROSS LANGUAGES ABOVE AND BEYOND THEIR SYNTACTIC STRUCTURES.

Moreno I. Coco[1] (moreno.cocoi@gmail.com), Eunice G. Fernandes[2], Manabu Arai[3], Frank Keller[4]; [1]Sapienza, University of Rome, [2]University of Adger, [3]Seijo University, [4]University of Edinburgh

Human cognition is a highly integrated system which synchronizes processes and representations across modalities. Our previous research on the synchronization between overt attention and human sentence production demonstrated that similar scene descriptions correspond to similar sequences of objects being attended to (i.e., scan patterns). Here, we generalise this finding from English to languages with a different word order. More specifically, we test whether synchronicity holds not just within a given language but across languages and examine the relative contribution of syntax and semantics. 74 participants (24 British English, 28 European Portuguese and 20 Japanese) were asked to describe an object (N = 24), either animate (e.g., man) or inanimate (e.g., suitcase), situated in a visual scene, and prompted with a cue word, while being eye-tracked. Across all participants, pair-wise similarities of sentences were computed using the Universal Sentence Encoder, which generates vector-based meaning representations across languages. Part-of-Speech (POS) sequences, which are a shallow representation of the syntax of sentences, were extracted using Spacy. Similarities between POS sequences and scan patterns (i.e., sequences of fixated objects) were computed using the Longest Common Subsequence algorithm. First, we successfully replicated that similar sentences are associated with similar scan patterns in all three languages. Moreover, we demonstrated for the first time that this relationship holds across languages: for instance if a Japanese and a Portuguese sentence are semantically similar, their associated scan patterns are also similar. In contrast to semantic similarity, we find that syntactic (POS) similarity is predicted by scan patterns only within the same scene, and only between languages with similar word order. This result not only

confirms that visual attention and language production are synchronized, but it theoretically points at a grammar of perception that is universal across languages, goes beyond (syntactic) surface realizations, and empirically manifests as oculomotor responses.

Progetti di Ricerca Medi; RM122181673FAF68 (Sapienza, University of Rome)

33.429 DIVIDED ATTENTION IN AMERICAN SIGN LANGUAGE PROCESSING

Dave Young[1] (dyoun200@gmail.com), Jasmine Awad[1], Ione Fine[1], Dina Popovkina[1]; [1]University of Washington

Previous work investigating simultaneous judgments of multiple stimuli in visual tasks has revealed a range of divided attention effects. There appears to be no capacity limit for simple tasks, such as judging the color of two stimuli simultaneously relative to judging one (White et al., 2018). In contrast, there are pronounced deficits in performance when judging two words simultaneously (White et al., 2018; Campbell and White, 2022), indicating a large capacity limit. Divided attention effects for the categorization of nameable objects are intermediate; less severe than for words, but greater than for color (Popovkina et al., 2021; 2023). Here, we investigated divided attention effects for American Sign Language (ASL) for signers and non-signers to test if capacity limits for processing ASL letters are more like written words than for objects, and whether expertise with ASL influences divided attention for signs. Our participants were either fluent with ASL (signers, n=6) or had no ASL knowledge (non-signers, n=7). Participants saw two simultaneously presented ASL letters, and responded whether a probe stimulus matched the stimulus in a previously cued location. In the single-task condition, only one of two stimuli was cued as relevant; in the dual-task condition, both stimuli were cued as relevant. The difference in performance between these two conditions is the dual-task deficit, which quantifies the cost of dividing attention. The dual-task deficit was similar in signers and non-signers (11.2% ± 1.92% vs. 10.72% ± 1.70%), with no significant difference between the groups ($t(11) = 0.19$, $p = 0.85$). The magnitude of the divided attention effect for ASL letters was smaller than those for written word judgments, but similar to those for nameable object judgments. These results suggest that ASL letter processing has limited capacity, but the underlying source of that limit may be different than for written words.

University of Washington Bolles Dissertation Fund, Mary Gates Endowment, Arc of Washington Trust Fund

33.430 AUTOMATED SYMBOLIC ORIENTING REVISITED: DO WORDS ELICIT THE SAME EFFECT?

Blaire J Weidler[1] (bweidler@towson.edu), Samuel W Bennett[2], Jared J McGinley[1], Elliott Fowble[1]; [1]Towson University, [2]Washington University in St. Louis

Traditionally, spatial attention was thought to be allocated either through endogenous (voluntary) orienting or exogenous (reflexive) orienting processes. However, Ristic and Kingstone (2012) suggested that a third form of orienting (automated symbolic orienting) that operates independently and in parallel with the other two forms of orienting could be elicited using nonpredictive arrow cues. Other prior research has found that directional word cues produce a similar effect as that of arrow cues. However, it is yet unknown whether the effect of nonpredictive words cues is due to automated symbolic orienting. Experiment 1 replicated Ristic and Kingstone (2012) but replaced the nonpredictive arrows with nonpredictive direction words. On each trial, participants detected the presence of a target as quickly as possible. Prior to the target's appearance, a traditional endogenous (predictive number) or exogenous (non-predictive border bolding) cue was presented simultaneously with a nonpredictive word cue. Words did not elicit any cueing effect, suggesting that the effect of nonpredictive words found in previous studies was not due to automated symbolic orienting. In Experiment 2, we performed a direct replication of Ristic and Kingstone (2012) using nonpredictive arrows, and the results supported the original findings for arrows eliciting automated symbolic orienting. Taken together, these findings suggest that directional word cues do not elicit automatic symbol orienting in the same way as arrows.

SUNDAY, MAY 19, 8:30 AM – 12:30 PM, PAVILION

Attention: Spatial selection 1

33.431 COMPETITION SHAPES SPATIAL CODING STRATEGY FOR SELECTIVE ATTENTION INSIDE VISUAL WORKING MEMORY: INSIGHTS FROM GAZE AND NEURAL MEASUREMENTS

Baiwei Liu[1] (b.liu@vu.nl), Zampeta-Sofia Alexopoulou[1], Anne Zonneveld[1], Freek van Ede[1]; [1]Vrije Universiteit Amsterdam

Selective attention enables us to prioritise task-relevant visual information, including among internal visual contents held in working memory. Studies targeting the foundational mechanisms of selective attention often consider situations where to-be-attended and to-be-ignored visual contents are presented or memorised in distinct directions from fixation – rendering direction sufficient for selection. Yet, in everyday life, direction alone is typically insufficient for selection, such as when multiple potential objects of visual attention compete along any given direction. To gain insight into the mechanisms of selective attention when direction is sufficient versus insufficient for selection, we cued participants to select memorised visual items that were encoded near or far from fixation while manipulating whether direction was sufficient (no competition along direction) or insufficient (competition along direction) for selection. Using fixational gaze behaviour and EEG-alpha lateralisation as read-outs of selective attention, we unveil the principle of 'efficient spatial coding', whereby selective attention in visual working memory considers just direction when sufficient and incorporates the distance of the attended memory target only when necessary.

This work was supported by an NWO Vidi Grant from the Dutch Research Council (14721) and by an ERC Starting Grant from the European Research Council (MEMTICIPATION, 850636) to F.v.E.

33.432 PRELIMINARY ANALYSIS OF THE CONTRAST RESPONSE FUNCTION AND ITS MODULATION BY SPATIAL ATTENTION USING MAGNETOENCEPHALOGRAPHY

Alfredo Sklar[1] (sklaral@upmc.edu), Jenay Kocsis[1], Dylan Seebold[1], Brian Coffman[1], Dean Salisbury[1]; [1]Clinical Neurophysiology Research Laboratory, UPMC Western Psychiatric Hospital, Department of Psychiatry, University of Pittsburgh School of Medicine

Introduction: The influence of attention on neural responses to gradations in visual contrast has been extensively studied in humans using electrophysiology and functional neuroimaging. However, these investigations have produced conflicting results regarding the mechanisms by which attention modulates the contrast response function (CRF), and both techniques possess limitations. The present study leveraged the spatiotemporal resolution of combined magnetoencephalography (MEG) and MRI to examine attentional modulation of the CRF within visual cortex. Methods: MEG was recorded from 12 participants during a 2-alternate forced-choice task. Participants identified the orientation of lateralized gabor patches during valid and neutral location cue conditions. Structural and resting-state fMRIs were acquired to obtain individual Human Connectome Project cortical parcellations. MEG sensor data was localized to 5 divisions of the visual cortex: V1, early (V2, V3, V4), dorsal (V3A, V3B, V6, V6A, V7), ventral (FFC, PIT, V8, VMV1/2/3, VVC), and MT+ (FST, LO1/2/3, MST, MT, PH, V3CD, V4t). Results: Participants responded more accurately ($p<.001$) and rapidly ($p<.001$) to higher contrast stimuli. Responses were faster on cued trials ($p=.005$) and to right visual field stimuli ($p<.001$). Across regions, increasing contrast ($p<.001$) and valid location cues ($p=.01$) enhanced cortical activity. The effect of contrast, however, differed across region with no significant effect observed within the ventral division ($p=.34$). Conclusion: Sensitivity to stimulus contrast and effects of covert attention were identified within V1 and early visual processing regions as well as some (dorsal division and MT+), but not all (ventral division), higher-order visual processing centers. The spatiotemporal resolution of combined MEG/MRI analysis may help replicate data obtained from single-unit recordings in animals and address discrepancies observed with fMRI analyses. Future work will model the CRF of each participant and derive its parameters to better characterize the response properties of the human visual cortex and assess the mechanisms of executive attentional modulation.

Supported by K23 MH127389 (Alfredo Sklar, MD, PhD PI).

33.433 UNRAVELLING THE INTERPLAY OF STATISTICAL LEARNING, TOP-DOWN, AND BOTTOM-UP MECHANISMS DURING TARGET SELECTION: INSIGHTS FROM BEHAVIOURAL AND EEG EXPERIMENTS

Carola Dolci[1], C. Nico Boehler[2], Einat Rashal[3], Sliann Ben-Hamed[4], Emiliano Macaluso[5], Leonardo Chelazzi[1], Elisa Santandrea[1]; [1]Department of Neuroscience, Biomedicine and Movement Sciences, University of Verona, Italy, [2]Department of Experimental Psychology, Ghent University, Ghent, Belgium, [3]School of Psychology, Keele University, United Kingdom, [4]Institut des Sciences Cognitives Marc-Jeannerod, Lyon, France, [5]Université Claude Bernard Lyon 1, CNRS, INSERM, Centre de Recherche en Neurosciences de Lyon (CRNL), U1028 UMR5292, IMPACT, F-69500, Bron, France

The natural environment exhibits consistent patterns, rendering it repetitive and partially predictable. Statistical learning (SL) enables us to discern these regularities from past experiences to then direct attention toward relevant elements for our objectives. Yet, it remains unclear whether SL collaborates or acts independently of other experience-independent attentional control mechanisms, specifically top-down and bottom-up control. In a series of interconnected experiments, we recorded behaviour and EEG activity during a visual search task to critically examine their interaction during target selection. In Experiment 1, we assessed the combined influence of top-down control (modulated via endogenous cueing - neutral/valid), bottom-up control (introduced by a salient item), and SL (induced by an imbalance in target probability - high/low - across locations) on behaviour. Additionally, Experiments 2 and 3 investigated how the N2pc, an EEG marker related to target selection, was impacted by the interplay of SL with top-down and bottom-up control, respectively. Simultaneously, SL and bottom-up control enhanced behavioural performance for targets at high (vs. low) frequency locations and for salient (vs. non-salient) targets. Crucially, an interaction revealed that the benefit of top-down control for validly cued targets could override the SL effect (Experiments 1-2). Moreover, EEG results indicated a greater N2pc for validly cued and salient targets, but only when at a low-frequency location. This finding suggests compensation for the lower attentional resources allocated to that location due to SL (Experiment 3). In summary, while top-down control and SL closely interact, the latter appears to be mostly independent from bottom-up signals.

33.434 ALLOCATION OF SPATIAL ATTENTION IN HUMAN VISUAL CORTEX AS A FUNCTION OF ENDOGENOUS CUE VALIDITY

William Narhi-Martinez[1] (narhi-martinez.1@osu.edu), Yong Min Choi[1], Blaire Dube[2], Julie D. Golomb[1]; [1]The Ohio State University, [2]Memorial University of Newfoundland

Certain areas of the brain are known to contain retinotopic maps of the visual field, and covert attentional guidance has been shown to result in spatially-specific increases of neural activity within certain cortical regions representing the attended locations. However, little research has been done to directly compare how attentional cues that carry differing levels of task-relevant spatial information will impact cue-modulated neural activity, particularly in terms of preparatory (i.e., pre-stimulus) attention. Here, we used fMRI to investigate how activity in area V4 would respond depending on if participants were cued with deterministic or probabilistic spatial information. Every trial began with a central arrow cue and subsequent pre-array delay, followed by a four-item memory array and subsequent post-array delay prior to the presentation of the memory probe for one of the array items. Critically, at the start of each run of trials, participants were informed that the arrow cues would indicate the to-be-probed location with either 100% validity (deterministic spatial cue) or 70% validity (probabilistic spatial cue). Our results revealed significantly higher cued versus noncued V4 quadrant activity for both probabilistic and deterministic cues prior to the onset of the memory array, but following the onset of the memory array, only significantly higher cued versus noncued V4 quadrant activity for deterministic cue trials. These findings reveal how cue validity alone can drive a differential allocation of neural resources across cued and noncued locations, and how this allocation can vary over time within a trial. Information providing certainty regarding the target's upcoming location appeared to bias attention both in anticipation of, and following, the presentation of task-relevant stimuli. In contrast, while information regarding where a target is most probable (but not guaranteed) to appear initially biased attention, this bias was more likely to spread or wane after the onset of task-relevant stimuli.

NIH R01-EY025648 (JG), NSF 1848939 (JG), and NSERC PDF (BD)

33.435 THE IMPACT OF PREDICTED AND UNPREDICTED EVENTS ON THE SPATIAL PRIORITY MAP: EVIDENCE FROM MULTIVARIATE PATTERN ANALYSES IN EEG

Jasper de Waard[1,2] (j.dewaard.jdw@gmail.com), Dirk van Moorselaar[1,2], Jan Theeuwes[1,2]; [1]Vrije Universiteit Amsterdam, [2]Institute of Brain and Behavior Amsterdam

Under the framework of predictive coding, perception is a constant interplay between the brain's prediction of upcoming events and incoming sensory information. One way that predictions shape perception is through the allocation of attention to locations that will likely contain relevant information. Spatial attention is assumed to be allocated according to the weights of an assumed integrated spatial priority map, but little is known about how this priority map changes in response to predictive processes. We used the high temporal resolution of electroencephalography (EEG) in combination with time-resolved multivariate pattern analyses to investigate the influence of predictions on the spatial priority map. We showed participants sequences of Gabor patches that appeared in a clockwise or counter-clockwise pattern of six locations, with varying spatial frequencies and orientations. Crucially, at random moments some Gabor patches were omitted from the sequence, while others violated expectations regarding the upcoming location of the subsequent Gabor patch. In combination with an independent localizer, this design allowed us to establish how and when the present shapes future representations in the context of established spatial associations. The EEG decoding results revealed that omitted anticipated locations were internally represented at their expected moments in time. Furthermore, we observed strong modulations in decoding accuracy for location violations. Together these findings advance our understanding of how we utilize recurrent patterns from past experiences to forecast future occurrences.

33.436 EFFECTS OF HEAD GAZE AND BODY POSITION ON ATTENTION IN REAL-WORLD SCENES.

Carmela Gottesman[1] (cvgottesman@sc.edu); [1]University of South Carolina Salkehathie

Orientation of people in scenes has been shown to affect viewers' attention distribution. Multiple studies show that gaze can cue attention, but other human features have also shown this effect. In previous studies we found that body orientation to the left of the right was sufficient to facilitate a search task in that direction, and gaze didn't have an added effect unless the search target was in the line of sight. In this study, we examined if the gaze has an added effect, if head position and gaze are moved toward the viewer, or to the side the body is oriented towards. Each trial showed a sequence of three images, all of the same person in the same scene. The body was oriented to the left or right of the scene, the same in the three pictures. In half the trials the head\gaze in the first picture was oriented toward the viewer and in half it was oriented to the same side their body was oriented towards. The second picture on each trials showed the opposite head orientation, resulting in apparent motion either towards or away from the viewer. After a brief mask, the second picture was repeated with a small cross superimposed either in the part of the scene towards which the body was oriented or in the opposite part. Viewers had to indicate if the cross was to the right or to the left of the person as quickly as possible. We found that viewers were slower to find the target when the person appeared to turn their gaze towards them, rather than away from them. There was an interaction of gaze with orientation direction so the left-right orientation didn't have an effect when the gaze turned towards the viewer but only if the gaze turned away.

33.437 INDIVIDUAL DIFFERENCES IN EYE MOVEMENTS AND PERCEPTUAL AVERAGING

William G. Hayward[1] (willhayward@ln.edu.hk), Herrick Fung[2], Janet H. Hsiao[3]; [1]Lingnan University, [2]Georgia Institute of Technology, [3]The University of Hong Kong

Numerous studies have documented our ability to perceive visual statistical averages over groups of similar objects. Nevertheless, the spatial attentional mechanisms and perceptual computations that constitute the basis of this ensemble representation remain unclear. Here, we explored these issues using a gaze-contingency paradigm. Participants performed a mean orientation estimation task under three conditions: (1) foveal vision only (obscuring the visual field further than 2.2° from fixation), (2) peripheral vision only (obscuring foveal information less than 2.2°), and (3) natural viewing. Both ensemble variance and stimulus duration were manipulated. Behaviorally, we found that participants could achieve similar accuracy as natural viewing when foveal vision was blocked but were impaired when peripheral vision was blocked. These results suggest that elements in the periphery are typically integrated into the ensemble representation, when available. Modeling analysis further revealed individual differences in integration mechanisms, with one group of participants better predicted by models assuming equal weighting of all items (which we liken to a Zoom-lens Model), as compared to another group which was better predicted by models weighted by the spatial distance between their fixations and the corresponding items (akin to a Spotlight Model). Using a Hidden Markov Model-based data-driven eye movement analysis, we found that participants could be clustered into two groups; one group tended to land their fixations on the center region of the ensemble whereas the other group distributed their fixations across the whole ensemble. Combining both analyses, we discovered that there was a weak correlation (r = 0.4) between the two groupings. Specifically, those better predicted by the zoom-lens model also tend to demonstrate a central fixation bias in their pattern of eye movements. Our findings argue against a universal mechanism in explaining perceptual averaging. Instead, individual strategic differences may play a role in determining the nature of perceptual averaging.

This research was supported by a Grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (17608519) to W.G.H.

33.438 THE INFLUENCE OF ATTENTION ON VISUAL ASYMMETRIES IN THE FOVEOLA

Samantha K. Jenks[1,3] (sjenks8@ur.rochester.edu), Martina Poletti[1,2,3]; [1]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA, [2]2 Department of Neuroscience, University of Rochester, Rochester, NY, USA, [3]3 Center for Visual Science, University of Rochester, Rochester, NY, USA

We have previously shown that fine visual discrimination varies over different directions in the foveola; in particular, it is better along the horizontal compared to the vertical meridian. Here we examine whether fine-tuning attention in the foveola could compensate for these asymmetries by enhancing perception to a greater degree at those locations characterized by lower performance. Participants (n= 5) performed a 2AFC orientation discrimination task while maintaining fixation on a central marker. Stimuli, small bars (7x2 arcminutes) titled ±45 deg, could be presented at four cardinal locations, 20 arcminutes from the preferred locus of fixation. Stimuli's contrast was changed adaptively using a PEST procedure. In half of the trials subjects were centrally cued to deploy their attention to one of the four locations where the target would briefly appear (100% cue validity). In the rest of the trials a neutral cue pointed in all four directions. The same task was replicated at 4.5 degrees eccentricity (n=3) with stimuli size adjusted to account for cortical magnification. Consistent with our previous work, when subjects were not cued to attend to a specific location, contrast thresholds along the horizontal meridian were more than double than along the vertical meridian in the foveola ($p<0.01$). However, when attention was engaged, contrast thresholds across the vertical (25%±9% contrast) and horizontal (21%±7% contrast) meridian did not differ statistically. Interestingly, a different trend was reported in the parafovea, where the horizontal-vertical meridian asymmetry was still present even when attention was engaged. These findings show that fine tuning of endogenous attention in the foveola can overcome visual asymmetries at this scale by enhancing visual discrimination to a greater extent in regions with lower baseline sensitivity.

This work was funded by NIH R01 EY029788-01 grant, NIH training grant T32EY007125, and EY001319 is the CVS core grant

33.439 FINE-TUNING EXOGENOUS ATTENTION AT THE FOVEAL SCALE: EFFECTS ON DIFFERENT SPATIAL FREQUENCIES

Yue Guzhang[1] (yue.zhang604@gmail.com), Martina Poletti[1]; [1]University of Rochester

We previously showed that exogenous attention can be selectively shifted within the 1-deg foveola, enhancing visual discrimination at the attended location only arcminutes away from distractors. In this study, we investigated how fine-tuning exogenous attention in the foveola affects contrast thresholds for different spatial frequencies. While observers (N = 5) maintained fixation on a central marker, an exogenous cue (100% validity) was presented at 0.75 deg eccentricity to the left/right of fixation. Shortly after the cue disappeared, two small gabor patches (0.5 deg visible area), tilted +/- 45 degrees, were shown (50 ms) on the left/right side at 0.5 deg eccentricity. The spatial frequency of the Gabors was either 4 or 12 CPD. Observers reported the orientation of the gabor that was presented at the location indicated by a response cue. Contrast thresholds for each spatial frequency were assessed in the presence (valid trials) and absence (neutral trials) of the attentional cue. Tight fixation was ensured by monitoring gaze position with a high-precision eye-tracker. In the neutral condition, contrast sensitivity peaked at 4CPD when tested over a wider range of spatial frequencies. In the attended condition, we observed a decrease in contrast thresholds with both 4 CPD (8%+/-2% v.s. 6%+/-2% contrast, P = 0.0082) and 12 CPD (24%+/-5% v.s. 20%+/-4% contrast, P = 0.0086) stimuli. The cueing benefit, the ratio between the contrast threshold in neutral and valid conditions, for high and low spatial frequencies was comparable. Further testing at intermediate spatial frequencies will determine whether these results are the outcome of either a broad attentional enhancement across spatial frequencies or a narrowly tuned attention enhancement peaking at frequencies slightly higher than the frequency at which contrast sensitivity peaks.

NIH R01 EY029788-01, P30 EY001319 & Meta, inc

33.440 PUPIL SIZE DURING VISUAL SEARCH: A MEASURE OF THE SPATIAL EXTENT OF ATTENTION

Emma Mueller[1] (emma.mueller@ttu.edu), Guangsheng Liang[1,2], Miranda Scolari[1]; [1]Texas Tech University, [2]Lingnan University

Given competing visual information, top-down attention flexibly prioritizes relevant over irrelevant input to accommodate behavioral goals. This is demonstrated by manipulating the validity of pre-cues that provide distinguishing information about an upcoming target. Participants may deploy space-based attention (SBA) following a location cue, where any visual object in the cued location is afforded an attentional benefit compared to uncued locations. Conversely, following a color cue, they may deploy feature-based attention (FBA) to objects depicted in the cued color regardless of location. Thus, although both SBA and FBA operate on sensory stimuli, they differ in how selection is distributed. Here, we investigate whether temporal changes in pupil size are associated with the distribution of selection. Participants were given either a location (N = 19) or color (N = 19) pre-cue (80% valid) before a visual search display containing one target and three distractors. Participants then made a two-alternative-forced-choice response about an orthogonal feature of the target. Importantly, all visual aspects were well-matched between conditions to allow us to draw conclusions about cognitive systems from pupillometry, and all analyses were restricted to correct trials. Among the location-cue group, pupillary responses were significantly greater on invalid compared to valid trials at 1,010-1,810 ms from search display onset; no significant differences were detected for the color-cue group. These findings likely reflect differences in attentional distribution: Successful shifts from one invalidly cued location to multiple uncued locations result in wider selected regions, whereas successful shifts from an invalidly cued to uncued colors result in no such spatial change, consistent with the global deployment of FBA. If this explanation is correct, manipulating the validity of a location pre-cue should modulate the effect, where an unreliable cue should elicit broad selection for both valid and invalid trials. A follow-up experiment was designed to confirm this hypothesis.

33.441 THE EFFICIENCY OF VISUAL PROCESSING ADAPTS TO THE "VIGOR" OF EYE MOVEMENTS: FROM WHAT BREAKS THROUGH INTO AWARENESS TO THE SPEED OF MEANING EXTRACTION

Joan Danielle K. Ongchoco[1] (joan.ongchoco@hu-berlin.de), Martin Rolfs[1]; [1]Humboldt-Universität zu Berlin

We can perform the same action with more or less "vigor" (e.g., walking to a good friend we haven't seen in a while versus walking after a long day) — and this phenomenon applies even to visual actions, i.e., eye movements. But as the speed of eye movements increases with higher vigor, sensory information also arrives at a faster rate, which might leave an inflexible system unable to catch up. Here we ask whether and how the visual system can flexibly adapt to the vigor of the actions that ultimately deliver its inputs. We manipulated the vigor of eye movements by varying the pace of saccade targets appearing and disappearing from view. Saccades of the same range of amplitudes were performed with higher velocities in high-pace than in low-pace conditions. We then assessed the consequences of this vigor induction across stages of visual processing. First, to test what breaks through into conscious awareness in the first place, observers pressed a key as soon as they saw a face in a repeated mask suppression experiment (e.g., Abir & Hassin, 2020). Observers showed a traditional face inversion effect (with faster breakthrough times for upright versus inverted faces) in high-vigor but not in low-vigor trials, suggesting that eye movement vigor may facilitate the effective prioritization of meaningful stimuli. Second, to test the speed of meaning extraction, observers reported whether they saw a particular concept (e.g., a dog) in rapid serial visual presentation (RSVP) streams (e.g., Potter et al., 1976; 2014). Detection rates increased with presentation durations (from 13ms to 80ms) — but this curve was shifted upwards and leftwards in high-vigor compared to low-vigor trials. Thus, the vigor of eye movements alters the efficiency of visual processing — from what enters awareness to how we extract meaning from incoming sensory information.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [865715 – VIS-A-VIS]).

Sunday, May 19, 8:30 am – 12:30 pm, Pavilion

Attention: Spatial selection 2

33.442 SPATIAL ATTENTION APPEARS MODULATED BY BEHAVIOURALLY RELEVANT CONTEXTS

Noah Britt[1] (brittn@mcmaster.ca), Jiali Song[2], Jackie Chau[1], Hong-jin Sun[1]; [1]McMaster University, [2]University of Toronto Mississauga

It is well-documented that visual spatial attention can be modulated by the visual features of objects in the environment if the features contain semantic information, especially when behaviourally relevant (e.g., emotional facial expressions). The current study demonstrated that observers could prioritize attention toward specific object features when, and only when, the object becomes relevant within a certain behaviourally relevant context. In the current study, using virtual 3-D technology, we presented to licensed drivers a modified cue-target paradigm where a peripheral cylinder cue was followed by a peripheral roadside pedestrian target. Participants discriminated the hand/arm position of the pedestrian with a button-press on a steering wheel. The pedestrian target could appear on the same or different side of the road as the cue. In addition, pedestrians could appear oriented toward the road or away from the road—but this feature remained irrelevant to the participants' responses. Through three experiments, we consistently found that, in the 3-D experimental condition where participants 'drive' within a virtual simulation, the cueing effect was significantly larger when pedestrians were facing towards the road compared to away from the road. This revealed enhanced attention towards targets—specifically those facing the road—in the cued location while driving. In contrast, this sensitivity for pedestrian orientation was not present in the three control conditions: 1) 3-D Stationary (non-driving), 2) 2-D Stationary (non-driving), and 3) another 3-D Driving scenario with an inanimate light-post target. These results suggest that drivers have heightened attention to pedestrians facing the road even though the pedestrian orientation was task-irrelevant. Licensed drivers likely demonstrated a preparatory mechanism to prioritize attention toward an event that may indicate a probability of impending collision. This novel phenomenon may be unique only to over-learned tasks such as driving (even simulated). These findings present additional evidence in favour of an embodied account of attention.

NSERC, Canada Foundation for Innovation

33.443 SEMANTICALLY RELATED OBJECTS ACT AS SPATIAL PREDICTORS DURING VISUAL SEARCH

Makayla Souza-Wiggins[1,] (mhsouza@ucdavis.edu), Joy Geng[2]; [1]University of California, Davis

In our daily lives, we navigate complex environments in search of target objects with remarkable efficiency. Successful detection relies on strategic attention allocation informed by semantic and structural information, but how this occurs remains unclear. In two studies, we tested the hypothesis that target search is facilitated by associated objects acting as semantic primes and spatial predictors. In Study 1 (N=43), we used an online RSVP task. Each trial featured five displays with two lateralized objects. The target (e.g., "toothbrush") appeared, preceded by a local or thematic "prime." The local prime was unrelated but consistently appeared before the target (e.g., "refrigerator"). The thematic prime was a semantically related object (e.g., "sink"). Results revealed quicker target detection with the thematic prime. Study 2 (N=114) replicated the previous study, placing the prime in a naturalistic spatial position relative to the target. We aimed to investigate if the thematic prime improved performance solely as a semantic cue or also functioned as a spatial predictor. Object pairs were extracted from Unity scenes for naturalistic object placement. The target appeared in four conditions: (a) thematic prime in a congruent spatial location (e.g., toothbrush on the sink), (b) local prime in a congruent spatial location (e.g., toothbrush next to the refrigerator), (c) thematic prime in an incongruent position (e.g., toothbrush on the floor by the sink), or (d) local prime in an incongruent position (e.g., toothbrush on top of the refrigerator). Participants exhibited a larger congruency effect for targets with a thematically related prime compared to a local prime, indicating semantic primes not only captured attention but also created a spatial prediction.

Overall, these studies deepen our understanding of efficient attention allocation in a semantically rich world.

33.444 MODELING ATTENTIONAL DEPLOYMENT IN AN IMMERSIVE ENVIRONMENT USING A FORAGING TASK IN VR

Chloe Callahan-Flintoft[1] (ccallahanflintoft@gmail.com), Joyce Tam[2]; [1]United States Army Research Laboratory, [2]The Pennsylvania State University

There is a wealth of information about how the brain decides on and selects locations of the visual field for enhanced processing. However the bulk of the knowledge acquired about visual attention is derived from paradigms using two-dimensional desktop displays. These experimental paradigms oftentimes involve restrictions of the head or eyes, where information is passively presented to the participant's visual system. However, real world vision occurs in an immersive and interactive environment. One consequence of this is processing differences between actionable versus non actionable locations in the visual field. This has been seen both in speeded processing times of graspable locations (Reed et al., 2010) as well as biases in distance judgements of reachable locations, an effect shown to extend with tool use (Brockmole et al., 2013). Therefore in order to properly understand how mechanisms of visual attention operate in real world vision, it is essential to account for effects that may only be elicited in such immersive, interactive environments. Toward this goal, the current work had participants perform a conjunction, foraging task using a head mounted display in virtual reality. Participants shot targets amongst distractors using an airsoft rifle integrated into the virtual space. To probe how the presence of a trajectory tool (the rifle) affected the deployment of attention, a cue (a red disk) was presented on each trial, either at the location of the participant's gaze, the weapon's aim point, or at a neutral location. Differences in response times to the cue, particularly slowing when the cue was presented at aim point, as well as selection biases for targets near gaze or aim point are used to characterize the effect of the rifle on the attentional landscape in a 3-dimensional environment.

33.445 THE ROLE OF STATISTICAL LEARNING IN ATTENTIONAL GUIDANCE DURING SEARCH THROUGH NATURALISTIC SCENES

Justin Frandsen[1] (justin.frandsen@tamu.edu), Brian Anderson[1]; [1]Texas A&M University

Associations between objects and locations can guide attention in visual search. Previous research has found that relationships between items and their location (e.g., cup on the counter, or sponge near the sink) could be used to guide the control of attention in naturalistic scenes. These associations are built over an individual's lifespan based on their own personal experiences, so it is difficult to understand the nature of how these associations have been formed. Previous research has found that different forms of memory, including knowledge concerning semantic relationships between objects and implicit learning, play a role in guiding attention. The present study looks to more closely examine how learning-dependent attentional guidance can occur in naturalistic scenes. Participants were assigned four target shapes, each of which was associated with one of three locations in a scene (floor, counter, or wall). While being eyetracked, on each trial participants were cued with one of the four target shapes and instructed to search for it in the following scene, responding with a keypress based on the orientation of a letter appearing within the target. The target shape appeared in its associated location within the scene on 75% of all trails and in an unassociated location on the remaining 25%, along with three distractor shapes. To require use of the cue, on half of all trials, one of the three non-target distractors was an uncued shape from the target set. Results show that response time was faster on valid trials than on invalid trials, and the first shape fixated was more likely to be the target on valid trials. These results suggest that associations between an arbitrary shape and a location can be used to guide search, providing evidence for the role of statistical learning in attentional guidance during search through naturalistic scenes.

33.446 THE DYNAMIC NATURE OF MEMORY-GUIDED ATTENTION

Oryah Lancry-Dayan[1,2] (oryah.lancry@mail.huji.ac.il), Tal Nahari[2], Gershon Ben-Shakhar[2], Adi Mizrachi[3], Galia Avidan[3], Yoni Pertzov[2]; [1]Brigham and Women Hospital, Harvard Medical School, [2]The Hebrew University of Jerusalem, [3]Ben Gurion University of the Negev

How does the cognitive system effectively navigate its limited capacity amidst the overwhelming details of the external world? Accumulating evidence suggests that addressing this longstanding question involves understanding the role of various memory systems in directing visual attention. While prior research has illuminated the enhancement of attention through semantic and working memory, our focus is on how attention adjusts to familiar stimuli. We conducted two tasks exposing participants to displays featuring four items—one familiar and three unfamiliar. The relevance of familiar items for task performance varied between the tasks. In the first task, participants detected a dot on one of the items, making familiarity orthogonal to the task. In the second task, participants memorized the set of four items. Here, familiarity could potentially facilitate item encoding, encouraging the cognitive system to allocate attentional resources to the unfamiliar items. We utilized eye tracking to characterize memory-guided attention patterns in both tasks. Despite familiarity being irrelevant to performance in the dot detection task, participants still exhibited a preference for looking at the familiar item. This preference was also evident in the initial phase of the memory task but shifted towards a preference for unfamiliar items as the trial progressed. Intriguingly, participants could voluntarily attenuate the attentional preference for familiar items but could not eliminate the task-related preference for unfamiliar items. Exploring neurodivergent populations, such as those with congenital prosopagnosia, revealed that the preference for familiar items required explicit knowledge, whereas a preference for unfamiliar items persisted even when these items were not explicitly recognized. Collectively, these results provide compelling evidence that memory-guided attention is a dynamic process that adjusts adaptively to prevailing circumstances and facilitates efficient deployment of visual attention. Understanding these attentional patterns also holds potential practical value, particularly in the context of detecting concealed information in forensic scenarios.

33.447 GENERAL AND SPECIFIC EFFECTS OF MEDITATION ON GAZE CUEING OF ATTENTION

Maria Falikman[1] (maria.falikman@gmail.com), Samuel Kearley[1], Robert Strickland[1], Oleksii Pidlypenets[1]; [1]The University of The South

Different types of meditation have been demonstrated to selectively modulate performance on different visual attention tasks. To our knowledge, there have been no studies of the effects of meditation on the gaze cueing of attention, or orienting one's visual attention to the direction of another person's gaze (Friesen, Kingstone, 1998). We assumed that spatial gaze cueing might be sensitive to the so-called 'loving-kindness meditation' (LKM), which is known to increase social connectedness (Hutcherson et al., 2008). More specifically, our hypothesis was that LKM would increase the gaze cueing effect in human observers. In our experiment, two groups of participants performed a speeded peripheral target detection following a valid, invalid, or neutral gaze cue from a cartoon-like face in the center of the display. The experiment was divided into two sessions, 128 trials each, separated by a pre-recorded ten-minutes LKM in the experimental group and a ten-minutes recording of a short story in the control group. Response times to the peripheral target were analyzed. For all conditions, we obtained a robust gaze cueing effect, which did not differ significantly between groups before the intervention. Contrary to our expectations, the interaction between the condition, session, and cue type for absolute RTs was not significant. However, the analysis of RT costs and benefits before and after the intervention revealed that LKM led to a significant 14 ms increase in the valid-cue benefit with no effect on the invalid-cue cost, whereas listening to a story did not have any effect either on costs or on benefits, in line with our hypothesis. Besides, for the experimental but not for the control group, we observed a significant training effect across all cue types, with shorter RTs after the meditation. Our results imply that LKM might have both general and specific effects on gaze cueing of attention.

33.448 LARGE-SCALE EXAMINATION OF THE BENEFIT AND COST OF SPATIAL ATTENTION AND THEIR VARIABILITY

Yaffa Yeshurun[1] (yeshurun@research.haifa.ac.il), Felipe Luzardo[1]; [1]University of Haifa

Spatial attention—the ability to prioritize relevant regions in our environment—is crucial for human cognition and has accordingly been studied extensively. However, large population samples have seldom been used, leaving inter-individual variability largely unexamined. Moreover, the relatively scarce studies that did consider individual differences in spatial attention did not include a baseline condition in which neutral cues do not guide attention to a particular location. This is important because a neutral condition is necessary for the distinction between attentional benefit (improved performance when attending the correct location) and attentional cost (impaired performance when attending the wrong location), and this distinction is critical for differentiating facilitatory from inhibitory mechanisms. Here, we recruited a sizable participant pool (N=438) across three experiments. The task was an acuity task with a Landolt square target that could appear in one of four possible locations. We included valid, invalid, and neutral conditions, in which precues indicated the right, wrong, or no location, respectively. We also performed a direct comparison of two types of neutral cues (a central neutral cue—a small bar presented at the center of the screen and a multi-cue with which four small bars cued all possible locations simultaneously). We measured both accuracy and RT and analyzed both the averaged attentional benefits and costs as well as their inter-individual variability. We found robust attentional benefits with high levels of inter-individual variability. In contrast, the cost averages and inter-individual distributions depended on the performance measurement (accuracy vs. RT), the cue-target distance, and the type of neutral cue. These findings provide a more refined understanding of the attentional mechanisms. Additionally, the direct comparison of the different types of neutral cue marked the 'multi-cue' as an optimal baseline for assessing both facilitatory and inhibitory aspects of spatial attention.

33.449 AN ATTENTIONAL SERIAL REACTION TIME TASK

Margaret L. Duffy[1] (duffy.390@buckeyemail.osu.edu), M. Fiona Molloy[2], David E. Osher[1]; [1]The Ohio State University, [2]University of Michigan

The brain has the ability to filter and prioritize a cacophony of often extraneous visual stimuli via a mechanism called attention. Traditionally, attention is thought to be guided by either a goal-driven mechanism (top-down, supported by fronto-parietal brain networks) or a stimulus-driven mechanism (bottom-up, supported by visual networks). However, recent work has highlighted a type of attention that does not fit within this dichotomy: one guided by memory. There are multiple forms of memory, including habitual memory - inflexible, insensitive to outcome devaluation, and automatic. Contextual cueing has arguably demonstrated habit-like attention, but effect sizes are very small, especially for individual subjects. Within the domain of habit research, serial reaction time (SRT) tasks have demonstrated habitual properties; however, this type of task typically is used for motor habits. For the present study, we adapted the SRT with spatial-temporal demands to probe habit-like attention. Participants were tasked with making decisions about stimuli that appeared in blocks alternating between learned and novel locations. We assessed participants' knowledge of the learned sequence with a short post-test that asked them to predict the next location of the image in this sequence. Finally, we altered the learned sequence to test perseveration errors of the attentional habit. This task produces very large reaction time and accuracy effects for the learned spatiotemporal attentional sequence, demonstrable in individual subjects. Participants appeared to possess explicit knowledge of the learned sequence, however, challenging the inclusion of implicitness as a criterion for attentional habits. Further, the altered sequence demonstrated a difference in reaction time between critical (altered) and non-critical trials in the learned sequence, and a trending effect for accuracy. Future directions include the implementation of the task using a more complex learned sequence and a more highly powered reversal task. This study is part of an undergraduate honors thesis.

Alkire Research Fund, Undergraduate Research Scholarship (Ohio State)

33.450 INFLUENCE OF AGING ON VISUAL ATTENTION AND PERIPHERAL PERCEPTION

Anne-Sophie Laurin[1] (anne-sophie.laurin@umontreal.ca), Noémie Redureau[1], Christine Gao[1], Julie Ouerfelli-Éthier[1], Daria Balan[1], Amine Rafai[1], Laure Pisella[2], Aarlenne Khan[1]; [1]Université de Montréal, [2]Université Claude Bernard Lyon 1

When searching for an item among distractors, we distribute our attention to a certain degree within our peripheral visual field. It has been suggested that in aging, there is a reduced attentional distribution, leading to visual and attentional declines. To investigate attentional distribution, 27 younger and 16 older adults performed a pop-out visual search task. With gaze-contingent methods, we presented different visible window sizes around participants' fixation during the task. We extracted the size of each participant's attentional window based on their search times for the different visible window sizes. To test whether performance in the visual search task was related to peripheral visual function, participants performed a contrast detection task and two motion detection tasks (local and global motion perception). In these tasks, stimuli were presented at two different peripheral eccentricities (5° and 10° distant from fixation point). Overall, we observed that older adults took longer to report the target's presence compared to younger adults, $t(41) = 6.31$, $p<.001$. Compared to younger participants, they also had a significantly smaller attentional window, $t(41)=2.16$, $p=.036$. In addition, older adults had higher contrast detection thresholds, $F(1, 41) = 23.23$, $p < .001$, and higher thresholds in local, $F(1,41) = 4.53$, $p=.039$, and global motion perception, $F(1,41) = 27.58$, $p<.001$, contrary to the idea that motion perception is preserved with aging. Overall, we observed a reduced attentional distribution as well as lowered contrast thresholds and lowered motion perception in aging. These findings, specific to the peripheral visual field, are compatible with the hypothesis of a functional decline of the visual dorsal stream in normal aging.

This research is funded by the Canadian Institutes of Health Research (CIHR) and by the Fonds de recherche du Québec en santé (FRQS).

33.451 PREPARATIONS TO BREAK THE TYRANNY OF FILM: DESIGNING STIMULI THAT CHANGE COMPREHENSION AND MOMENT-TO-MOMENT CONTENT IMPORTANCE IN VIDEO

Prasanth Chandran[1] (prasanthp@ksu.edu), Alyssa Smith[2], Ashton Lofing[2], Lindsey Wilson[2], Brayden Wallace[2], William Mass[2], Joseph P. Magliano[3], Lester C. Loschky[4]; [1]Doctoral Student, Kansas State University, [2]Undergraduate Student, Kansas State University, [3]Professor, Georgia State University, [4]Professor, Kansas State University

We know that people's visual attentional selection is strongly influenced both by top-down tasks, and bottom-up salience. But, how does viewers' understanding of what they see, namely their event model, influence their attention? We have tested this in narrative film, through context manipulations that influence viewer's event models. However, in prior studies, large differences in viewers' event models produced small differences in their attentional selection, termed the Tyranny of Film. We hypothesize this is because film stimuli have powerful stimulus features to guide viewers' attention. Based on the Scene Perception & Event Comprehension Theory (SPECT), we propose two necessary conditions for strongly attenuating the Tyranny of Film: 1) There are predetermined differences in viewers' event models (specified via event indices), focusing on a target of attention. 2) The target of attention has relatively low visual salience. Method: We used 10 pairs of silent commercial video clips that could satisfy the first necessary condition. We created a context manipulation for each clip, and a common viewing period that would be watched 1) without having seen the context, and 2) after having seen the context. After each viewing, participants were asked to watch the clip again and use their mouse to 1) track the most important person/place/thing at each moment, and 2) click at particularly important moments. Results: Participants hovered their mouse longer, and clicked more on the targets of attention when in the Context condition. Thus, the context manipulation influenced viewers' moment-to-moment judgments of importance on the screen, further confirming that our selected videos meet the first criterion for breaking the Tyranny of Film. To test the second necessary condition, we will analyze the common viewing period video clips using the DeepGaze-3 saliency algorithm, and determine whether the targets of attention have relatively low saliency. Future studies will measure eye movements.

Research reported in this abstract was partially supported by the Cognitive and Neurobiological Approaches to Plasticity (CNAP) Center of Biomedical Research Excellence (COBRE) of the National Institutes of Health under grant number P20GM113109.

SUNDAY, MAY 19, 8:30 AM – 12:30 PM, PAVILION

Eye Movements: Fixational eye movements

33.452 A VIRTUAL TARGET CONTROLS FIXATION BETTER THAN A REMEMBERED TARGET

Scott Watamaniuk[1,2] (scott.watamaniuk@wright.edu), Devashish Singh[2], Arvind Chandna[2], Stephen Heinen[2]; [1]Wright State University, [2]The Smith-Kettlewell Eye Research Institute

Previously we investigated how an eye is controlled during fixation when it is covered and unable to view a visual target that its fellow eye sees. During monocular fixation the two eyes moved differently - the covered eye traversed a larger region (assessed with bivariate contour ellipse area or BCEA) and drifted faster which we attributed to poorer fixation control of the covered eye using a "virtual" target constructed by the brain rather than the visual target. Here we ask how fixation is controlled when neither eye sees the target. Binocular eye movements were recorded simultaneously in a darkened room at 1000 Hz with an EyeLink. Observers fixated a small spot for 500 msec that disappeared while fixation of the remembered target was maintained for 19 sec. Compared to an occluded eye's behavior during monocular fixation, here when the target's position had to be remembered, the two eyes continued to move differently. However, relative to the occluded eye in the monocular condition, BCEAs in both eyes were considerably larger, drift was slower, saccade frequency was lower, and saccades were larger. However, as during monocular viewing, saccades remained conjugate. This suggests that the oculomotor strategies for following a remembered target are different from those used to follow a virtual target, and result in poorer control. Our new model of binocular control includes a virtual target to provide position error feedback to an occluded eye. Without the virtual target the eyes wander farther and produce less frequent saccades that apparently do not well correct for the eye drift. The differences between the movements of the two eyes provides further evidence that drift during fixation is controlled independently while microsaccades remain conjugate, as posited in our model.

NIH 1R01EY034626-01

33.453 ENHANCEMENTS AND IMPAIRMENTS IN VISUAL DISCRIMINATION OF DIFFERENT SPATIAL FREQUENCIES BEFORE THE ONSET OF SPONTANEOUS MICROSACCADES

Zoe Stearns[1,2] (zstearn2@ur.rochester.edu), Krishnamachari Prahalad[1,2], Martina Poletti[1,2]; [1]Center for Visual Science, [2]University of Rochester

Highly localized perceptual enhancements at the microsaccade goal have been reported to occur before the onset of these tiny gaze shifts. Yet, these modulations have been examined only in the unnatural context of instructed microsaccades. It remains unknown whether they unfold with the same magnitude for spontaneous microsaccades and whether they are broadly or selectively tuned for spatial frequency (SF). Here we investigate how naturally occurring microsaccades impact discrimination of foveal stimuli of different SFs. Eye movements were recorded with a high-resolution digital Dual-Purkinje Image eye-tracker while subjects (N=7) performed a 2AFC discrimination task. Two vertically oriented gabor patches (0.5deg diameter) at different SFs (4,8,12 and 16 cpd) were presented 0.5 deg to the left and right of the display center. After a variable delay, one of the two gabor patches briefly (50 ms) tilted ±45 deg and observers reported the direction of the orientation change. Contrast thresholds for each SF were assessed in microsaccade-free trials (baseline condition), stimulus contrast was then maintained at threshold throughout the task. Only trials in which the gaze was maintained close to the center of the display before microsaccade onset were examined. Our results show that, before the onset of a spontaneous microsaccade, visual discrimination is improved with respect to baseline when microsaccades are directed toward the stimulus that changed orientation ($\Delta d'=0.73\pm0.46$, $p<0.05$). This modulation appears to be stronger for lower SF ($\Delta d'=1.1\pm0.56$ vs. $\Delta d'=0.25\pm0.71$, $p< 0.05$). On the other hand, performance was impaired when microsaccades were directed away from the stimulus that changed orientation. These findings show that in everyday settings, when observers view complex foveal stimuli, discrimination for a wide range of SFs across the central fovea is systematically modulated prior to an impending microsaccade.

funded by Meta and EY001319

33.454 HOW DO PUPILLARY LIGHT RESPONSES AND MICROSACCADES ALLUDE TO VOLUNTARY AND INVOLUNTARY AUDITORY ATTENTION?

Hsin-I Liao[1] (hsini.liao@ntt.com), Yung-Hao Yang[1,2], Shimpei Yamagishi[1], Yuta Suzuki[1], Shigeto Furukawa[1,3]; [1]NTT Communication Science Laboratories, NTT Corporation, [2]Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, [3]Shizuoka Graduate University of Public Health, Shizuoka General Hospital

Pupillary light responses (PLR) and microsaccades allude to voluntary and involuntary visual attention, presumably reflecting the function of the frontal eye field and superior colliculus. In line with such allusion to visual attention, we recently found that PLR corresponds to the luminance condition of the voluntarily attended auditory space and object (Liao et al., 2023). This suggests a common audio-visual spatial attention mechanism underlying the neural circuits controlling eye movements. The current study further examines how PLR and microsaccades allude to voluntary and involuntary auditory attention using the same stimuli. Data from Liao et al. (2023) were reanalyzed to investigate the effect of voluntary attention on microsaccades and compare the PLR result with that induced by involuntary attention. To probe involuntary attention, we applied the Posner cueing paradigm. A task-irrelevant yet salient noise burst was presented to the left or right ear (or location) via headphones (or loudspeakers) before the target sound. Participants identified the target sound as quickly as possible while ignoring the noise distractor. Behavioral results confirmed an involuntary attentional shift to the distractor, as evidenced by faster reaction times when the target was presented on the same side as the distractor than when it was presented on the opposite side. Pupillary response, in contrast to the finding relating to voluntary auditory attention in Liao et al. (2023), showed little to no significant difference corresponding to the luminance condition where the distractor was presented. Microsaccades were found to occur in the same direction as the voluntarily attended sound and involuntarily attended distractor at different timings. Opposite microsaccadic directional bias was found only to the distractor, suggesting inhibition of return. These findings suggest an interaction between the top-down modulation and sound-driven attentional orienting underlying the microsaccade-generation system. The PLR reflects voluntary auditory attention but not involuntary auditory attention.

33.455 INVOLUNTARY EYE-MOVEMENT SIGNATURES DIFFER FOR RECOGNITION OF ONESELF, FAMILIAR AND UNFAMILIAR FACES

Lisa Schwetlick[1] (lisa.schwetlick@uni-potsdam.de), Hendrik Graupner[2], Olaf Dimigen[3], Ralf Engbert[1]; [1]University of Potsdam, [2]Bundesdruckerei GmbH, [3]University of Groningen

Miniature, involuntary eye movements occur continuously during fixations. Two important examples of these involuntary movements are pupil dilation, which responds primarily to luminance changes, and fixational eye movements, specifically microsaccades, which perturb the gaze during fixations at a rate of about 1/s. Both movements are primarily functionally-driven but are also subject to modulation by cognitive and emotional influences. Face recognition is a process which is tightly coupled to visual, cognitive, and affective processing. Here, we show that the signatures of both microsaccades and pupil size are different when a face is recognized compared to unknown. We conducted a highly-powered (n=116) visual fixation experiment, comprising two distinct groups. Participants knew the faces of the people in their own group, but not of the other group. We then compared self-recognition (i.e., seeing one's own face), other-recognition (i.e., seeing a familiar face), no recognition (i.e., a stranger's face). The pupil typically constricts in response to new stimuli, followed by a dilation. We show that constriction is attenuated and the dilation increased during recognition compared to unknown faces. This effect is strongest for self-recognition. The microsaccade rate is typically inhibited by stimulus onsets. We find that inhibition is strongest during self-recognition, followed by recognition of familiar faces, and weakest in response to unknown faces. Our results are consistent with findings from the Oddball Paradigm that show stronger pupil dilation and stronger microsaccade inhibition effects in response to the infrequent target stimuli, related to increased attention. The difference in our task was that familiar and unfamiliar faces occurred with the same frequency. Thus, effects can be clearly attributed to recognition processes. Our results shed light on the process of face recognition and indicate a potential use of oculomotor measures in revealing hidden knowledge or in biometric identification procedures.

This work was funded by Bundesdruckerei GmbH, Berlin, Germany.

33.456 SPATIAL DYNAMICS OF MICROSACCADES: INVESTIGATING THE INFLUENCE OF FIXATIONAL PROXIMITY TO VISUAL STIMULI ON SACCADE-INDUCED EEG MODULATION

Christopher Turner[1] (christopher.turner@glasgow.ac.uk), Aleksandra Vuckovic[1], Gemma Learmonth[2], Alessio Fracasso[1]; [1]University of Glasgow, [2]University of Stirling

Microsaccades have emerged as a valuable tool for examining EEG activity during tasks requiring eye fixation. While it is established that microsaccades generate a distinct evoked potential and EEG topography (Dimigen et al., 2009), it is unclear to what extent this modulation is due to motor processing, sensory input, or a combination of both. This is particularly relevant as there is significant spatial variability of visual stimuli across experiments, representing different sensory input. This study aims to determine if the post-saccade proximity of the fovea to high-contrast visual stimuli influences saccade-locked EEG activity. Examining data from an experiment involving a Posner reaction time task before and after neurofeedback and tACS stimulation, we focused on microsaccades during the 2.5-second cue period of the Posner task. Microsaccades were categorized into leftward and rightward directions, further stratified into 'close' and 'far' bins based on their landing proximity to lateralized visual stimuli. Employing a between-conditions cluster-based permutation test, we statistically assessed the evoked response of 'close' and 'far' saccades for both directions. The results revealed a saccade-locked evoked potential with timing and topography matching that which has been found in previous studies (Dimigen et al., 2009; Meyberg et al., 2015). This evoked potential was characterized by a biphasic muscle spike potential at saccade onset, followed by a positive peak at 106ms and a negative peak at 162ms in occipital electrodes. However, statistical analysis demonstrated no significant differences between 'close' and 'far' conditions for left or right saccades. The replication of saccade-locked ERPs and the absence of significant differences between ‘close’ and ‘far’ saccades suggest that fixational proximity to visual stimuli minimally influences the saccade-induced modulation of the EEG signal. Moreover, the saccade-locked evoked potential appears to reflect purely motor processes, rather than changes induced by the change of visual input on the retina.

Engineering and Physical Sciences Research Council (EPSRC)

33.457 TEMPORAL ATTENTION AND EXPECTATION JOINTLY MODULATE MICROSACCADES

Aysun Duyar[1] (aysun@nyu.edu), Marisa Carrasco[1]; [1]New York University

Background. Temporal attention is deployed at specific moments to prioritize relevant information, and temporal expectation is developed based on the probabilities of when events occur. Microsaccades, small fixational eye movements, serve as an online measure for cognitive processes. Microsaccades impair perception of brief stimuli, are suppressed with expectation, and are suppressed more so with temporal attention. Previous studies either focused solely on expectation or attention effects on microsaccades, although these two mechanisms interact at the behavioral and neural levels. Goal. To investigate how temporal attention and expectation jointly modulate oculomotor control dynamics to improve perception. Methods. Observers maintained fixation while performing a 2AFC orientation discrimination task. Two Gabors were presented sequentially at the fovea with a variable onset. A precue instructed participants to attend to the first, second, or both stimuli. A response cue at the end of the trial indicated the target observers had to discriminate. We manipulated temporal expectation via precision (within-session onset variability), and hazard rate (within-trial onset delay). The probability of stimulus onset peaked in the middle of the temporal window, therefore, stimuli could appear earlier, at, or later than the expected moment. Results. Microsaccade rates decreased before the stimulus onset, and more so with hazard rate. Temporal attention amplified this effect. Temporal attention pulled microsaccade inhibition and rebound to an earlier time point, regardless of precision. Although optimal pre-stimulus attentional inhibition was not possible under low precision, post-stimulus microsaccade rebound occurred sooner with temporal attention. Despite similar inhibition, rebound microsaccades had smaller amplitude when temporal precision was high, indicating a decoupling between inhibition and rebound, and suggesting expectations modulate oculomotor control even after stimulus offset. Conclusions. Temporal attention and expectation are distinct mechanisms that jointly modulate microsaccade timing to optimize visual processing.

NIH R01-EY019693 to M.C.

33.458 TIME COURSE OF MICROSACCADES DIRECTIONALITY DURING AN ENDOGENOUS ATTENTION TASK

Riccardo Brandolani[1,2], Claudio Galletti[1], Patrizia Fattori[1], Rossella Breveglieri[1], Martina Poletti[3]; [1]Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy, [2]University of Camerino, Center for Neuroscience, 62032 Camerino (Italy), [3]Department of Brain & Cognitive Sciences, University of Rochester, (NY), USA

Several studies have shown that microsaccades index the direction of covert attention shifts. Yet, recently, this correlation has been questioned. In this study, we investigated this topic in a modified Posner task while recording eye movements with a high-precision eye-tracker. Participants (n=7, 13500 trials) were instructed to maintain fixation on a central square. Two differently colored peripheral circles (1-deg diam) simultaneously appeared 5 degrees to the right and to the left of the fixation point for 500ms. Then, the circles turned gray, while the central square took on one of the two colors, thereby serving as a cue for directing covert attention to the corresponding side (Cue window, 700-1000ms). The cue was valid in 80% of trials. A low-contrast target then appeared in one of the two circles, and participants had to release a button as soon as the target was detected. As expected, reaction times were modulated by trial validity (valid: 412ms vs invalid: 482ms; $p<0.001$). A fine grain temporal analysis on microsaccade directionality showed that, even though each participant was characterized by an idiosyncratic bias before cue onset, microsaccade direction on average followed the cue from 100 to

400ms after cue onset. However, directionality of microsaccades in this window did not affect reaction times, raising questions on the effectiveness of microsaccades as an index of covert attention.

EU Marie Sklodowska-Curie grant agreement No 101086206 - PLACES; MIUR (PRIN 20208RB4N9); R01 EY029788: Vision Attention and eye movements at the scale of the foveola - National Institutes of Health.

33.459 THE RELATIONSHIP BETWEEN THE DECLINE IN CONE DENSITY AND ACUITY ASSESSED IN NORMAL VIEWING CONDITIONS ACROSS THE CENTRAL FOVEA

Ashley M. Clark[1] (aclark43@ur.rochester.edu), Krish Prahalad[1], Benjamin Moon[1], Austin Roorda[2], Pavan Tiruvee[2], Wolf Harmening[3], Aleksandr Gutnikov[3], Samantha K. Jenks[1], Krishnaveni Nagarajan[1], Sanjana Kapisthalam[1], Michele Rucci[1], Jannick P. Rolland[1], Martina Poletti[1]; [1]University of Rochester, [2]University of California Berkeley, [3]University of Bonn

Even if cone density declines with increasing eccentricity within the 1-degree foveola, acuity across this area is believed to be uniformly constrained by eye optics. Acuity aligns closely with the Nyquist sampling frequency of the cone mosaic only when optical aberrations are corrected. Although the impact of optics across the foveola is uniform, vision, assessed in the presence of physiological optical aberrations, is not uniform within this region. Here we examined the relationship between the rate at which cone density declines and changes in normal acuity across the foveola. Acuity thresholds were determined using an adaptive PEST procedure and a 4AFC visual discrimination task, while subjects' gaze position (N=8) was recorded with a high-precision dDPI eye-tracker to ensure fixation remained within a 5-arcmin region around the center of the display throughout the task. Acuity thresholds were assessed at eccentricities of 0, 10, 15, and 25-arcmin from the fixated point. Subjects' retinae were imaged at high-resolution with an Adaptive Optics Scanning Laser Ophthalmoscope, and cone density was determined as a function of eccentricity with respect to the preferred retinal locus. We report considerable individual variations in the rate of decline of both acuity and cone density. At all eccentricities, acuity was below the corresponding Nyquist frequency of the cone mosaic. Yet, all subjects showed a substantial drop in acuity (from 20/17 to 20/22 on the Snellen chart) from 0 to 25-arcmin eccentricity, leading to a 33% increase in thresholds. Despite a slight decrease in cone density, for most subjects, visual acuity remained relatively constant from 0 to 10-arcmin. On the other hand, the decline in acuity from 0 to 25-arcmin closely mirrored the rate of cone decline at these eccentricities, suggesting that factors beyond optics play a role in limiting acuity at larger eccentricities in the central fovea.

NIH EY029788, EY018363, EY001319, EY023591, German Research Foundation (DFG), Ha5323/6-1, Ha5323/8-1

33.460 ULTRA-FINE KNOWLEDGE OF GAZE POSITION IN SACCADE PLANNING

Yuanhao Howard Li[1] (yli162@u.rochester.edu), Martina Poletti[1], Michele Rucci[1]; [1]University of Rochester

Sensorimotor integration is an important component of spatial representations, as retinotopic information needs to be combined with extra-retinal knowledge about eye movements to properly locate objects in space. Recent research has shown that extra-retinal oculomotor knowledge extends to fixational drift, the persistent wandering of the eye in between saccades: human observers are capable of inferring geometrical configurations even in the absence of spatial information on the retina, purely based on motor knowledge of eye drift (Zhao et al, Nature Communications, 2023). Here we examine whether extraretinal information about fixational drift is also used to control eye movements. Specifically, we study whether saccade planning takes into account drift-induced displacements of the line of sight from the intended fixation location. Emmetropic observers (N=5) attempted to maintain fixation on the location P0 of a previously briefly displayed marker (a 10' dot at maximum contrast) in complete darkness. After an interval of 2s, the eye had drifted to a new location PE, and a saccade cue (another 10' dot displayed for 50 ms) instructed subjects to perform a 4 degree horizontal saccade. The task of the subject was to perform a return saccade to the remembered location of P0. Our data show that the return saccade compensated for the previous drift displacement. This compensation accounted for 55% of the angular correction needed to fixate back to the original point. Similar results were found when the task was replaced by a 2AFC where subjects performed a saccade toward one of two cues placed at positions P0 and PE. In both conditions, performance was higher than when subjects visually selected P0 or PE with a button press instead of a return saccade, suggesting that the uncertainty in extraretinal drift estimation is smaller in the motor modality than for visual judgments.

Research supported by NIH grants EY018363 and P30 EY001319

SUNDAY AFTERNOON POSTERS IN BANYAN BREEZEWAY

SUNDAY, MAY 19, 2:45 – 6:45 PM, BANYAN BREEZEWAY

Object Recognition: Reading

36.301 FINETUNING PRIMATE VISUAL REPRESENTATIONS WITH WORD RECOGNITION

Aakash Agrawal[1] (aakash@alum.iisc.ac.in), Shubhankar Saha[2], Surbhi Munda[2], S. P. Arun[2], Stanislas Dehaene[1,3]; [1]CEA, [2]Indian Institute of Science, [3]College de France

Reading accurately requires precise encoding of letters and their relative positions, especially to distinguish similar words like TRIAL and TRAIL. While reading expertise leads to the formation of the Visual Word Form Area (VWFA), the neuronal changes enabling invariant word recognition remain unknown. In this study, we trained two macaque primates to process orthographic inputs using a same-different task. The neural responses to a large set of 4-letter strings were wirelessly recorded from the Inferior Temporal (IT) cortex before and after 5 days of training. A subset of these stimuli was designed as “words,” while others were used as “nonwords.” During each training session, the monkeys performed at least 360 task trials. The sample stimuli were always words, and their positions varied to test for invariant word recognition. The test stimuli, which could be either words or nonwords, were always centered. Neural data was also recorded during the training sessions. As expected, the monkeys' performance (response time and error rate) improved over the five days of training, suggesting that word recognition became increasingly position-invariant. Furthermore, they became sensitive to the orthographic structures of words, and their performance on non-matching conditions degraded with an increase in similarity to words. Neural recordings revealed that many cells were tuned to specific letters. Using a letter x position-based model, we observed that most neurons encoded retinotopic position in the contralateral visual field, but a few neurons also encoded the ordinal position of letters. With training, the representation space expanded, and receptive field size of few neurons broadened, eventually leading to enhanced position invariance. Overall, we observed finer discriminability between letter strings with training, which could potentially be a consequence of a shift from retinotopic to ordinal position encoding among a few units.

36.302 TASK-BASED MODULATION OF HIGHER-ORDER LEXICAL STATISTICS IN THE VENTRAL AND DORSAL VISUAL STREAMS

Oscar Woolnough[1] (oscar.woolnough@uth.tmc.edu), Xavier Scherschligt[1], Ellery Wheeler[1], Stanislas Dehaene[2], Nitin Tandon[1]; [1]UTHealth Houston, [2]INSERM

Words appear everywhere in our everyday environment. But how does our brain process these words differently while we are actively reading them compared to when they are in the background? Prior work clearly demonstrates enhancement of activation to attended stimuli, but does attention also modulate sensitivity to higher-order word-level statistics within visual cortex? To probe this, we used direct intracranial recordings in six humans as they each performed four different visual-word-based tasks in four separate blocks: (i) Fixation, (ii) One-back, (iii) Semantic decision, and (iv) Lexical decision. Words and pseudowords were presented on screen for 500ms each and participants were asked to push a button when target stimuli for that task appeared on screen. Using linear mixed-effects modelling of broadband gamma activity (70-150Hz), we mapped how these task demands modulated spatiotemporal representations of higher-order lexical statistics, such as lexical frequency and lexicality. We targeted three primary hubs of the reading network in mid-fusiform cortex, inferior parietal sulcus and inferior frontal gyrus. Within individual electrodes we clearly detected task-based modulation of overall activation and sensitivity to both lexicality and lexical frequency. At group level, when performing a lexical decision task, both sensitivities were enhanced across the entire reading network, as compared to fixation. Modulation of lexicality sensitivity manifested as increased activation in response to pseudowords with little modulation of overall known word activation. In contrast, while performing a semantic decision task we observed selective enhancement of lexical frequency sensitivity only in inferior frontal gyrus. This was paired with selective increases in bidirectional Granger causality between mid-fusiform cortex and inferior frontal gyrus, relative to fixation, specifically for known words. This work demonstrates that both local processing of higher-order lexical statistics and inter-regional connectivity throughout the reading network are highly task dependent and are sensitive to the specific demands of a given task.

This work was supported by the National Institute of Neurological Disorders and Stroke U01NS128921.

36.303 QUANTIFICATION OF READING CIRCUITS IN THE VENTRAL OCCIPITOTEMPORAL CORTEX

Yongning Lei[1], David Linhardt[2], Pedro M. Paz-Alonso[1,3], Garikoitz Lerma-Usabiaga[1,3]; [1]Basque Center on Cognition, Brain and Language, San Sebastián, Gipuzkoa, Spain, [2]High Field MR Center, Medical University of Vienna, Austria, [3]Ikerbasque, Basque Foundation for Science, Bilbao, Bizkaia, Spain

Substantial evidence shows that one of the circuits supporting reading is located in the ventral occipito-temporal cortex (VOTC), a region that was first functionally identified and denominated visual word form area (VWFA). Recent evidence combining functional, structural, and quantitative MRI suggests that the VOTC can be segregated into at least two sub-regions: one involved in visual feature extraction in the posterior occipito-temporal sulcus (pOTS), and one implicated in integrating with the language network in the middle occipito-temporal sulcus (mOTS). Due to the heterogeneity of functional localizers, it is usually difficult to know if the same cortical region is being located across studies and if the results of the analyses are comparable. The aim of the present work was to develop a single-subject multimodal VOTC word recognition localizer, replicable across different labs and studies. Our experiment followed a dense-sampling (10 repetitions) strategy with 2 participants. We acquired 2 different fMRI reading region localizers, classical 8-bar retinotopy scans with different contents inside the bars (checkers, words, pseudowords, and false fonts), and structural MRI scans (quantitative and diffusion MRI, T1w, and T2w images). Our results revealed that by combining multimodal MRI measurements we can segregate two different reading regions at the individual subject level, and that with 10 repetitions, we can measure the variability of the two localizations. Furthermore, we developed and tested in different settings a shorter localizer capable of capturing similar individual variability, that will be made available to the scientific community. In sum, here we propose a new protocol to harmonize inter-lab and inter-study visual word recognition, which will be critical to advance our understanding of the role of the VOTC in neurobiology of reading.

This research is supported by the Basque Government through the BERC 2022-2025 program and Funded by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation CEX2020-001010/AEI/10.13039/501100011033 and through project PIBA_2022_1_0014 funded by Basque Government

36.304 CORTICAL ACTIVATIONS FOR SYMMETRY EFFECT ON VISUAL WORD FORM PERCEPTION IN DEVELOPMENTAL DYSLEXIA

Shao-Chin Hung[1,2], Hsiao-Lan Sharon Wang[3], Chien-Chung Chen[1]; [1]National Taiwan University, [2]Academia Sinica, [3]National Taiwan Normal University

Developmental reading difficulties may be associated with visual-spatial deficits. Poor readers often make mirror errors along the vertical axis and cannot discriminate between symmetrical patterns, such as d and b as different letters. These mirror errors, observed both in the alphabetical and orthographic languages, may contribute to the difficulties in analyzing the spatial relationships among strokes within a character. Here we investigated the effect of symmetry on visual word form perception in developmental dyslexic (DD) and chronological aged (CA)-matched typical juvenile (aged 12-17 years). We employed real-, non-, jiagu- and scrambled characters in hanzi with symmetric or asymmetric compositions as stimuli. In the behavioral experiment, observers had to judge whether the two stimuli presented on both sides of the central fixation were identical. Both groups responded more accurately in the symmetric compositions than the asymmetric ones. This symmetry effect was found for all character types except for the real-characters. In a block-design functional magnetic resonance imaging (fMRI) experiment, we measured blood oxygenation level dependent (BOLD) activations when observers were asked to match the presented character with that of the previous trial. The visual word form area (VWFA) in the left fusiform gyrus showed significantly differential cortical activations to real- versus scrambled characters for the CA group but not for the DD group. The CA group showed stronger BOLD activations in the lingual and fusiform areas while processing non-characters versus real-characters compared to the DD group. While analyzing symmetric versus asymmetric compositions, the DD group had a greater activation in the bilateral fusiform gyri which are sensitive to global configurations of a visual word form. Taken together, the behavioral and neural evidence revealed intact symmetry processing in the dyslexic readers, suggesting independent processes for spatial configuration and orthographic visual word forms.

36.305 MEASURE LETTER RECOGNITION PERFORMANCE: A SUBJECTIVE EVALUATION METHOD

Deyue Yu[1] (deyueyu@gmail.com), Emily Watson[1]; [1]Ohio State University

Yu & Watson (2022) found that letter recognition performance can be estimated using a guided subjective self-report, and there was a high correlation between the subjective estimation and the performance-based (objective) measure obtained independently with psychophysical method. However, in their study, the objective and the subjective data were collected from different observers, and the correlation between measurements was only assessed at the group level. Here we are examining whether similar results can be observed at the individual level. Nine normally sighted young adults completed five testing blocks (four objective blocks and one subjective block). Target stimuli were single lowercase letters presented at 10° eccentricity left and right of the fixation. In the objective blocks, observers performed a letter recognition task following each letter presentation (10 trials/letter/block). In the subjective block, observers viewed each target letter ten times in a row, and then reported all the letters that they perceived the target to be and assigned a percentage value for each reported letter (1 trial per letter). Consistent with the findings from Yu & Watson (2022), there was a good agreement between subjective (81%) and objective measurement (85%) for overall letter recognition accuracy, and the correlation between subjective and objective measure was significant for recognition accuracy of individual letter ($r=0.74$, $p<0.001$) and distribution of letter confusion ($r=0.78$, $p<0.001$). When examining individual observers, we found significant correlations between subjective and objective measure for overall letter recognition accuracy ($r=0.65$, $p=0.03$), recognition accuracy of individual letter ($r=0.36$ to 0.72, $ps\leq0.03$), and distribution of letter confusion (except one observer, $r=0.24$ to 0.65, $ps\leq0.04$). All tests were one-tailed tests. Our results showed that the subjective evaluation method can provide estimation on letter recognition performance with much fewer trials and has potential for serving as a viable alternative to performance-based evaluation.

Neiderhauser AMD Grant

36.306 PARALLEL PROCESSING OF WRITTEN WORDS AS A FUNCTION OF VISUAL FIELD POSITION

Kimya Firoozan[1] (krf2141@barnard.edu), Amritha Anupindi[2], Alex White[1]; [1]Barnard College, [2]Northwestern University

Reading places intense demands on the visual system. Although many words on a page are visible at once, it is debated whether readers fully process multiple words in parallel, or focus attention on one word at a time. We have used a dual-task paradigm to investigate whether it is even possible to recognize two words simultaneously. In these experiments, participants view pairs of English words that are flashed briefly and replaced by post-masks at an interval set to each participant's threshold. This gives them just enough time to recognize one word with focused attention. The question is: can they recognize two words with divided attention in that same amount of time? We previously found that the answer was no: accuracy in the divided attention condition was so impaired that it supported a serial model. In those studies, the words were positioned either just above and below fixation, or to the left and right, centered at 2.75º (~4 letter spaces). Here, we report multiple experiments in which we investigate whether parallel processing of two words is possible when they are arranged more like in natural reading. By varying the words' eccentricities and lengths, we have found that for many observers (but not all), accuracy rises above the serial model's prediction when the words do not extend more than ~6 letter spaces from the fovea. Performance is most consistent with parallel processing when one word is directly fixated and the other is placed to the right. Altogether, this study supports the hypothesis that two words can be recognized in parallel as long as they are arranged horizontally and fit within a narrow window around the point of gaze fixation. Further research will examine the nature of individual differences in this processing capacity.

Supported by National Eye Institute grant R00 EY029366, and the Arnold and Mabel Beckman Foundation

36.307 REDUNDANT TARGET EFFECTS REVEAL CAPACITY LIMITS FOR RECOGNIZING WORDS AS A FUNCTION OF VISUAL FIELD POSITION

Jannat Hossain[1] (jh4328@barnard.edu), John Palmer[2], Alex White[1]; [1]Barnard College, [2]University of Washington

The visual system can encode many stimuli across the visual field simultaneously, but the number of objects that can be fully identified in parallel is limited. At the extreme, some stimuli must be identified serially. One useful tool for distinguishing parallel from serial processing is the redundant target paradigm, which compares responses to displays containing one target to displays containing two targets. Parallel models generally predict a redundancy gain: faster responses to two targets. The standard serial model (improved to incorporate response accuracy as well as speed) predicts either no effect or slower responses to two-target than to one-target displays. We conducted several experiments that measured performance for three different judgments of written words: color detection (detecting colored letters), lexical decision (detecting real words among pseudowords), and semantic categorization (detecting nouns that refer to living things). In one set of experiments, the words were above and below fixation. In another set, the words were to the left and right of fixation to better match natural reading. The color detection tasks yielded a strong and positive redundancy gain, while the lexical and semantic tasks yielded zero or negative effects. These results are consistent with low-level features (color) for two words being processed in parallel, while the meanings of two words are processed serially. Altogether, this study informs models of reading and suggests opportunities for the redundant target paradigm to investigate other aspects of vision.

Supported by National Eye Institute grant R00 EY029366

36.308 THE PROCESSING OF SPATIAL FREQUENCIES THROUGH TIME IN VISUAL WORD RECOGNITION

Clémence Bertrand Pilon[1,2] (clemencebertrandpilonn@gmail.com), Martin Arguin[1,2,3,4]; [1]Department of Psychology, Université de Montréal, Montreal, Quebec, Canada, [2]Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Montreal, Quebec, Canada, [3]Centre interdisciplinaire de recherche sur le cerveau et l'apprentissage (CIRCA), Department of Psychology, Université de Montréal, Montreal, Quebec, Canada

The spatial frequencies (SF) optimal for word recognition are well established. Studies with other classes of stimuli however, demonstrate a rapid temporal evolution of the SFs most useful for visual recognition in a coarse-to fine order. The present study is the first to examine the time course of SF processing in a visual word recognition task. Word images were filtered according to four non-overlapping SF bandpass conditions (center frequencies of 1.2, 2.4, 4.8, and 9.6 cycles per degree). The stimulus presented on each trial was an additive combination of the target image (i.e. signal) and of a white noise field. The signal-to-noise ratio (SNR) varied randomly throughout the 200 ms exposure duration. Performance was maintained at 50% correct by adjusting the target contrast on every trial. Accuracy-based classification images (CI) of processing efficiency as a function of time were calculated to demonstrate the temporal progression of SF processing. These time-domain CIs show that the highest spatial frequency range dominates early processing, with a shift toward lower spatial frequencies later during stimulus exposure. This pattern interacted in complex ways with the temporal frequency content of signal-to-noise oscillations, as demonstrated by time-frequency CIs. A machine learning algorithm (support vector machine with leave-one-out cross validation) was exposed to the data patterns of individual participants with the task of deciding the SF band it corresponds to. The maximum classification accuracy of 90.6% correct (over 25% chance performance) was achieved when the classifier was exposed to the Fourier transformed time-frequency CIs. This level of accuracy is about twice that obtained from the classifier with raw or Fourier transformed time-domain CIs or raw time-frequency CIs. These findings suggest that the temporal progression of SF processing in visual word recognition is best understood if the time dimension is actually recast into its Fourier transform.

Supported by grants from the Fonds de Recherche Québec—Nature et Technologie and the Natural Sciences and Engineering Research Council of Canada (NSERC) to Martin Arguin and a research scholarship to Clémence Bertrand Pilon from the Conseil de recherches en sciences naturelles et en génie (CRSNG).

36.309 A PSYCHOPHYSICAL APPROACH FOR INVESTIGATING FORMAT READABILITY ONLINE

Kurtuluş Mert Küçük[1] (kurtulusmertkucuk@gmail.com), Rashid Md Mamunur[2], Veronica Penkova[3], Amy Giroux[4], Nilsu Atılgan[5], Shaun Wallace[6], Sam Berlow[7], Anna Kosovicheva[8], Stephanie Day[9], Ben D. Sawyer[10]; [1]University of Central Florida, [2]University of Toronto, [3]University of Rhode Island

We introduce a scientific tool designed for online reading performance studies. Tool assesses optimum reading format for individuals by allowing experimenters to manipulate various text parameters. Developed using psychophysical research, the tool utilizes online testing via Pavlovia and Psychopy, enabling large-scale participant testing with reduced environmental noise and increased external validity. Our tool's primary function is to assess reading performance across various typefaces, font parameters (e.g., weight, width, etc.), letter spacings by ranking comprehension scores and reading speed. The tool focuses on paragraph reading (approximately 150-word paragraphs), though it can also evaluate other forms of reading such as single word recognition and sentence reading. Stimuli are presented as .jpg images of texts with `modified fonts or spacings. Using images of texts instead of directly rendering using the browser, prevents potential incompatibility problems across different monitors while manipulating letter spacing and axes of variable fonts. We outline the methodology, emphasizing the tool's reliance on automatic randomization and counterbalancing, and the creation of stimulus sets. We provide a pilot study as an example to explain the configuration of tool's settings and how counterbalancing functions. Example also outlines how behavioral performance measures such as comprehension scores, reading speed calculations (as words per minutes), and experimental conditions are registered in the data file. Overall, we provide an overview of the tool's design, functionality, and potential to expand the capabilities of online readability studies.

36.310 TYPEFACE MATTERS: PSYCHOPHYSICAL INSIGHTS INTO READABILITY ACROSS DIFFERENT READING TASKS

Nilsu Atilgan[1] (nilsuatilgan@gmail.com), Jonathan Dobres[1], Md Mamunur Rashid[1], Sam M Berlow[2], Ben D Sawyer[1]; [1]University of Central Florida, [2]Typography for Good

Although reading is a complex activity, as skilled readers, we often do not realize how often we naturally engage in this task in our everyday lives. The field of reading research is vast, encompassing diverse perspectives, from single letter recognition in the periphery to evaluating comprehension and fatigue in reading longer texts. While our visual system operates similarly in various reading tasks, the specific underlying visual mechanisms for each task may differ. To gain a complete understanding of the factors affecting reading, it is crucial to assess and compare their impact across different tasks. Here, we investigated the impact of typefaces, by using eight different fonts, across three different reading modes: i) interlude (150-word long paragraphs), ii) sentence (followed by true/false questions), and iii) glance (single word/pseudoword lexical decision task). Fifty participants completed the study and their reading speed / reaction time and accuracy were measured during each task. Preliminary

results demonstrated that the optimum typeface, associated with the best performance, showed correlations across reading tasks, suggesting shared underlying mechanisms. The highest correlation was found between glance and sentence reading ($r = .52$, $p < .001$), with moderate correlations observed between glance and interlude tasks ($r = .44$, $p < .001$), as well as sentence and interlude tasks ($r = .36$, $p < .001$). However, there were also observed differences in the optimum typeface across different tasks. Specifically, while Merriweather was found to be the optimum font for interlude and glance reading, Source Serif Pro yielded the best performance in sentence reading. Current findings suggest that the complex processes of reading may lead to the development of diverse reading strategies, with external factors potentially influencing the process variably. This suggests that individuals' preferences for optimal reading parameters, especially fonts in this study, may vary depending on the task, despite correlations.

36.311 LIGHT OR BOLD? NAVIGATING FONT WEIGHTS AND GRADES FOR ENHANCED READABILITY

Md Mamunur Rashid[1,4] (mdmamunur.rashid2@ucf.edu), Nilsu Atilgan[1,4], Hilary Palmén[2], Jonathan Dobres[4], Sam M Berlow[3,4], Mert Küçük[4], Ben D. Sawyer[1,4]; [1]University of Central Florida, [2]Google Fonts, Google, [3]Typography for Good, [4]Virtual Readability Lab

Several visual factors impact an individual's reading performance, with fonts being recognized as one of the most pivotal elements. Before the advent of variable font technology, assessing subtle, systematic changes in font designs across continuously varying parameters was challenging. Variable fonts now allow precise adjustments to specific parametric axes, facilitating detailed examinations of their impact on readability. Here, we investigated two parameters: weight and grade, as they are widely used in digital environments. The weight of a font allows adjusting the style from lighter to bolder in typographic color, by varying stroke weights, spacing and kerning, and other aspects of the type. The grade of a font allows changing the style from lighter to bolder without any changes to the overall width, line breaks, or page layout. Separately examining these parameters helps make inferences about the underlying visual mechanisms. In this study, we examined three variable fonts (i.e., Google Sans, Roboto Flex, and Roboto Serif), each including three levels of weight and grade.179 participants (Mean age= 45.7) were recruited through an online platform and were asked to complete a paragraph reading test. Their reading speed and comprehension levels were measured. While there was little to no effect of typeface, different light/bold levels of grade and weight revealed an effect on reading speed. The results showed that lighter grades and weights were read slightly faster than the bolder ones (X2 (2, N = 141) = 5.4, $p = .063$). This finding is in line with the previous research, potentially suggesting the effect of reduced crowding. As expected, a strong impact of age was also observed on the degree to which these parameters influencing readability, showing a greater impact on younger readers (X2 (1, N = 141) = 12.9, $p < .001$), supporting the importance of individualization of font parameters for individuals.

36.312 THE EFFECTS OF VARIABLE FONTS ON SENTENCE-LEVEL READING

Silvia Guidi[1] (s.guidi@mail.utoronto.ca), Zainab Haseeb[1], Anna Kosovicheva[1], Benjamin Wolfe[1]; [1]University of Toronto Mississauga

The visual appearance of text on a screen can have a large impact on how efficiently we extract information from it. While many studies have examined visual factors in reading with sentences and passages, these stimuli are often challenging to use in standard psychophysical paradigms (e.g., forced-choice tasks). Single-word reading lends itself more easily to these techniques but does not always reflect how we read in real life. To facilitate the study of visual factors in reading, we developed a sentence classification task that uses true or false sentences and varies the presentation duration of these stimuli to determine a duration threshold. For this, we developed a human-validated sentence corpus of sentences taken from GenericsKB, a repository of internet-derived sentences. The sentences were filtered based on word and character length, truthfulness, and word frequency. We validated the database by having participants rate the truthfulness of each sentence, which showed high inter-rater agreement (mean ICC of 0.98, n=79). We have used this corpus tool to examine visual factors in reading using variable fonts, which are fonts where each parameter, such as slant and width, can be customized on a continuous axis. We measured participants' (n=8) duration thresholds for five parameters (thin stroke, thick stroke, width, weight, and slant) at five different settings on each axis by varying the duration of the displayed sentences. Thin stroke and thick stroke showed a U-shaped trend where extreme settings had longer duration thresholds and the middle setting had the fastest duration threshold. Our sentence corpus and paradigm allow researchers to use forced-choice psychophysical methods to study reading based on naturally-occurring sentences. By understanding how font settings affect reading performance, this work supports the goal of determining what font helps each reader.

This work was supported by a SSHRC Insight Grant to AK and BW.

36.313 LANGUAGE-UNIVERSAL AND SCRIPT-SPECIFIC FACTORS IN THE RECOGNITION OF LETTERS IN VISUAL CROWDING: THE EFFECTS OF LEXICALITY, HEMIFIELD, AND TRANSITIONAL PROBABILITIES IN A RIGHT-TO-LEFT SCRIPT

Amit Yashar[1] (amit.yashar@edu.haifa.ac.il), Adi Shechter[2], Sivan Medina[3], David L. Share1[4]; [1]University of Haifa

Background: Visual crowding – the failure to identify an object in clutter, imposes significant constraints on reading and has been linked to reading difficulties and developmental dyslexia. Previous studies in alphabetic scripts have demonstrated that letter recognition within a trigram string is more accurate when the string forms a word compared to a pseudoword (the well-known "lexicality" effect). This effect occurs both in the fovea and the parafovea. However, words and pseudowords differ not only in their lexical properties, such as print frequency, but also sublexically in the transitional probabilities of their letters (n-grams). These probabilities which capture the likelihood of the occurrence of a specific letter given its neighboring letters, play a crucial role in reading. The precise mechanism through which transitional probabilities facilitate reading, however, remains unclear. Objective: We investigated the effects of transitional probability (bigram/trigram frequency), lexicality (words vs. pseudowords) and visual hemifield on crowded letter recognition among skilled readers in Hebrew, a right-to-left script. Method: In two experiments (N = 27), we measured font-width threshold in three conditions: uncrowded (an isolated letter), crowded word, and crowded pseudoword. In Experiment 2, observers also performed several blocks of crowded word and pseudoword tasks at threshold level. We used logistic regression analysis to determine the contribution of transitional probability to performance. Results: We revealed two language-universal effects: a lexicality effect and a right hemifield (left hemisphere) advantage, as well as a strong language-specific effect – a left bigram advantage stemming from the right-to-left reading direction of Hebrew. This finding suggests that transitional probabilities are a significant factor in parafoveal letter recognition. Conclusions: These results shed light on the visual system's processing of crowded stimuli in general and in printed words in particular. They reveal that script-specific contextual information, such as letter combination probabilities, influences letter recognition in crowded displays.

This research was supported by the Israel Science Foundation (Grant No. 1980/18 to AY)

SUNDAY, MAY 19, 2:45 – 6:45 PM, BANYAN BREEZEWAY

Object Recognition: Acquisition of categories

36.314 CONTRASTING LEARNING DYNAMICS: IMMEDIATE GENERALISATION IN HUMANS AND GENERALISATION LAG IN DEEP NEURAL NETWORKS

Lukas S. Huber[1,2] (lukas.s.huber@unibe.ch), Fred W. Mast[1], Felix A. Wichmann[2]; [1]Cognition, Perception and Research Methods, Department of Psychology, University of Bern, [2]Neural Information Processing Group, Department of Computer Science, University of Tübingen

Behavioral comparisons of human and deep neural network (DNN) models of object recognition help to benchmark and improve DNN models but also might help to illuminate the intricacies of human visual perception. However, machine-to-human comparisons are often fraught with difficulty: Unlike DNNs, which typically learn from scratch using static, uni-modal data, humans process continuous, multi-modal information and leverage prior knowledge. Additionally, while DNNs are predominantly trained in a supervised manner, human learning heavily relies on interactions with unlabeled data. We address these disparities by attempting to align the learning processes and examining not only the outcomes but also the dynamics of representation learning in humans and DNNs. We engaged humans and DNNs in a task to learn representations of three novel 3D object classes. Participants completed six epochs of an image classification task—reflecting the train-test iteration process common in machine learning—with feedback provided only during training phases. To align the starting point of learning we utilized pre-trained DNNs. This experimental design ensured that both humans and models learn new representations from the same static, uni-modal inputs in a supervised learning environment. We collected ~6,300 trials from human participants in the laboratory and compared the observed dynamics with various DNNs. While DNNs exhibit learning dynamics with fast training progress but lagging generalization, human learners often display a simultaneous increase in train and test performance, showcasing immediate generalization. However, when solely focusing on test performance, DNNs show good alignment with the human generalization trajectory. By synchronizing the learning environment and examining the full scope of the learning process, the present study offers a refined comparison of representation learning. Collected data reveals both similarities and differences between human and DNN learning dynamics. This disparity emphasizes that global assessments of DNNs as models of human visual perception seem problematic without considering specific modeling objectives.

This research was funded by the Swiss National Science Foundation (214659 to LSH). FAW is a member of the Machine Learning Cluster of Excellence, funded by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy—EXC number 2064/1—Project number 390727645.

36.315 A NEURAL NETWORK MODEL OF HOW CATEGORY LEARNING ALTERS PERCEPTUAL SIMILARITY

Luke Rosedahl[1] (luke.rosedahl@gmail.com), Thomas Serre[1], Takeo Watanabe[1]; [1]Brown University

Decades of research have shown that how we categorize objects changes how we perceive them. For example, category learning can increase the perceptual similarity of within-category items and decrease the similarity of between-category items (Goldstone and Hendrickson, 2010). However, the mechanism for this phenomenon (commonly called Categorical Perception or CP) is under debate. Based on previous work showing that CP is reduced when attention is directed away from the stimulus (Brouwer and Heeger, 2013), here we postulate that categorical perception occurs through category-learning steered Feature-Based Attention (FBA). In this theory, category learning alters perception by steering FBA to change the gain of visual cortex neurons responding to specific feature values. We test this hypothesis utilizing two category structures: Rule-Based (RB) and Information-Integration (II). In RB structures, the optimal strategy is to make binary decisions along feature dimensions and performance is increased if attention is targeted to specific feature values. In II structures, information from multiple feature dimensions must be combined before a decision can be made and performance is decreased if attention is targeted to specific feature values. The theory predicts that RB structures will cause greater changes in perception than II structures. Subjects (n=20) were divided evenly between RB and II conditions and underwent same-different testing before and after category learning. We found that the RB condition caused significantly more changes in same-different performance than the II condition. We then implement a neural network model that learns to apply feature-specific feedback (gain) modulation during category learning. We show that the feedback connections in the model, which was previously developed to explain how category learning induces transfer of visual perceptual learning, enable the network to show the same behavior patterns as human participants. This work provides strong behavioral and computational evidence for feature-based attention being the mechanism for categorical perception.

This work was supported by the National Eye Institute of the National Institutes of Health under award numbers [R01EY019466, R01EY027841, and R01EY031705], NSF-BSF under award number [BCS2241417], and the Office of Naval Research under award number [N00014-24-1-2026].

36.316 DECODING CONTEXTUAL EFFECTS IN VISION: A CROSS-SPECIES BEHAVIORAL APPROACH

Anaa Zafer[1] (anaazaf@yorku.ca), Sara Djambazovska[1], Hamidreza Ramezanpour[1], Gabriel Kreiman[2], Kohitij Kar[1]; [1]Department of Biology, York University, [2]Harvard Medical School

The significance of context in visual perception is undeniable. Our understanding of the natural world is shaped not just by the foveated visual objects but also by the surrounding scene and prior experiences. While the influence of context on vision has been demonstrated psychophysically, the underlying mechanisms integrating objects and surrounding information during scene comprehension are not fully understood. Studies have extensively examined "low-level" contextual effects, such as extra-classical receptive fields and surround suppression, yet gaps remain in comprehending how context affects "higher-level" visual recognition tasks. To elucidate these neural processes, a detailed examination of the neural networks involved is essential. Rhesus macaques, with their visual processing circuits akin to humans, present an ideal model for this purpose. In our study, we assessed the behavior of 90 human participants via Amazon Mechanical Turk in a binary match to sample object discrimination task, using images with varied contexts (full, incongruent, no context, etc.). The results revealed a significant alteration in human performance due to contextual changes, exhibiting a consistent behavioral pattern across context categories (trial-split reliability of ~0.8). This finding was crucial for comparison with macaques. After training monkeys (n=2) to achieve ≥80% accuracy in object categorization with full-context images, we exposed them to the same contextually manipulated images. The behavioral variance shared between humans and monkeys was significant (~31%), and not attributable to low-level image factors such as object size or contrast. Interestingly, naive macaque inferior temporal (IT) neural responses did not fully account for the observed human-monkey shared variance (13% of image-level explained shared variance), suggesting that the effects are likely driven more by learning processes and feedback mechanisms than by innate IT response statistics. This research paves the way for future investigations into the neural mechanisms of contextual modulation in primate vision.

Canada Research Chair Program, Google Research, CFREF, Brain Canada, SFARI

36.317 ACCELERATION OF VISUAL OBJECT CATEGORIZATION IN THE FIRST YEAR OF LIFE

Celine Spriet[1,2] (cel.spr@gmail.com), Emilie Serraille[1,2], Liuba Papeo[1,2], Jean-Rémy Hochmann[1,2]; [1]Institut des Sciences Cognitives Marc Jeannerod, CNRS, UMR5229, [2]University Lyon1

A remarkable acceleration of visuo-attentional processes has been described in the first months of life up until adulthood. In parallel, throughout the first years of life, infants seem to integrate more and more visual features together leading to the ability to represent more and finer-grained visual categories. Here, we show that these two processes are related: as infants grow older, visual processing gets faster, facilitating the integration of visual features to form categorical object representations. Using frequency-tagging electroencephalography (EEG), we targeted a response associated with the perceptual distinction between animate and inanimate objects in adults, in 4- and 9-month-old infants. Images (640 animate and inanimate objects) from one category (animate) were presented at a regular, base frequency (Fb), interleaved with images from the other category (inanimate) presented at the regular target frequency Ft=Fb/5. Visual categorization was tested at increasingly faster stimulation frequency (for adults (n=36): Fb=6, 12 or 30 Hz; for 4-month-olds (n=64): Fb=4 or 6 Hz; for 9-month-olds (n=64): Fb=6 or 12 Hz). Results revealed that the baseline-corrected response amplitude at Ft (and harmonics), used as a measure of categorization, decreased as the stimulation frequency increased. In 4-month-olds, the categorization response was only observed at 4 Hz; while it was found with a stimulation frequency as fast as 12 Hz in 9-month-olds, and 30 Hz in adults. The 4-month-olds' response at 4 Hz was equivalent to the 9-month-olds' response at 12 Hz, suggesting that 9-month-olds process categories ~3 times faster than 4-month-olds. These results demonstrate a dramatic acceleration of visual categorization, a process relying on feature integration, in the first year of life, which continues into adulthood. We will propose a model where, as the processing speed accelerates, more and more visual features activated by visual stimulation can be integrated, yielding more efficient and richer categorization.

This research was funded by the European Research Council Starting Grant to Liuba Papeo (Grant number: THEMPO-758473), the Labex Cortex grant to Jean-Rémy Hochmann and the Fondation pour la Recherche Médicale - Fin De Thèse Grant (FDT202304016547) to Céline Spriet.

36.318 COMPARISON TRAINING IMPROVES PERCEPTUAL LEARNING OF SKIN CANCER DIAGNOSES

Victoria L. Jacoby[1] (vjacoby@ucla.edu), Christine Massey[1], Philip Kellman; [1]University of California, Los Angeles

Perceptual learning of complex classifications has largely used trials in which a learner makes a categorization response to a single presented instance. Some prior research, however, suggests simultaneous comparisons of items from different categories can enhance processes of perceptual learning that facilitate successful categorization. Little work has directly compared how the concurrent learning of multiple perceptual classifications may differ when training emphasizes active comparison of items as opposed to individual item classification. We tested the effectiveness of paired comparisons for learning the differential diagnosis of 10 categories of benign and cancerous skin lesions. A Paired Comparisons condition, in which a learner was presented on each trial with a category label and required to choose between instances from two different categories, was compared to Single-Classification and Dual-Classification conditions, where instances of one or two categories were presented for classification on each trial. Novice participants were given a 40-min learning period before completing immediate and one-week delayed posttests on the classification of novel exemplars. Importantly, these assessments used the standard single-classification trial format regardless of the learning condition that preceded it. Results: In the learning phase, participants in the Paired Comparisons condition were able to complete significantly more trials than those in the classification conditions (large effect sizes), suggesting the relative ease of this trial type. In the posttests, classification accuracy reliably favored the Paired Comparisons condition, both immediately following learning, as well as after a one-week delay (all medium effect sizes). The Single- and Dual-Classification conditions did not reliably differ from each other on any learning or assessment measures. These results indicate that training based on paired comparisons can promote more efficient learning and generalization of complex categories than the more standard classification-based approaches.

This research was supported by National Institutes of Health award R01CA236791 to PK and CM. Any opinions, findings, and conclusions are those of the authors and do not necessarily represent the views of the NIH.

36.319 SHIFTING PERCEPTIONS: THE EFFECTS OF SUBORDINATE LEVEL TRAINING ON CATEGORY RESTRUCTURING

Anna K. Lawrance[1] (annalawrance@uvic.ca), Johannes Schultz-Coulon[2], Brett D. Roads[3], James W. Tanaka[1]; [1]University of Victoria, [2]Maastricht University, [3]University College London

Experts identify objects in their domain of expertise faster, more accurately, and at a more specific level of abstraction than novices (Tanaka & Taylor, 1993). Whereas a novice sees the yellow bird flitting in the bush, the expert instantly recognizes this object as a Cape May Warbler. Although substantial research has explored the behavioral and neural correlates of the expert's downward shift in recognition, less is known about how their mental structure mediates such speeded identification. In our experiment, 75 participants were trained to identify ten images of Cape May, Magnolia, Prairie, and Townsend warblers to a criterion of 90% accuracy. Before and after training, category structure was assessed with PsiZ. PsiZ (https://psiz.readthedocs.io) is a machine learning package that generates a multi-dimensional category representation (i.e., psychological embedding) based on the participant's judgments of image similarity. The key finding was that training produced profound changes in category structure. Specifically, warbler images belonging to different species became significantly more differentiated, while warbler images of the same species became more compact; hence, training produced between-category expansion and within-category compression. What is the relationship between category structure and category performance? Once participants completed their post-training PsiZ judgments, participants were given a recognition test where they were asked to identify the species of novel Warbler images and images used in training. Based on their recognition accuracy, the group of top 25% and bottom 25% performers were identified. The psychological embeddings were then inferred for each group and compared. The PsiZ results revealed significant differentiation between species, particularly among the lower quartile participants following training. Moreover, after training, top-performers showed denser within-species clusters than lower performers. Collectively, subordinate-level training produced significant category restructuring. Further, the quality of this reorganization appears to play a functional role in one's expert recognition performance.

36.320 THE INFLUENCE OF EXPERTISE AND INDIVIDUAL DIFFERENCES ON PSYCHOLOGICAL EMBEDDINGS

Eric Mah[1] (ericmah@uvic.ca), James Tanaka[2], Brett Roads[3]; [1]University of Victoria, [2]University of Victoria, [3]University College London

How does knowledge and expertise affect our perceptions and representations? How do individuals differ in the ways in which they represent and judge the similarity of concepts and percepts? And how best to measure psychological representations? In two experiments, we attempted to provide insight into these questions using PsiZ, a novel method for obtaining psychological embeddings–rich multi-dimensional representations of psychological similarity spaces inferred from behavioural similarity judgments. Specifically, we investigated whether psychological embeddings could be used to measure individual differences in the use of conceptual versus perceptual judgement strategies in domain experts and novices. In the first experiment, we presented two basketball experts, 12 basketball fans, and 16 novices unfamiliar with basketball with arrays of faces of famous basketball players from four NBA teams and asked them to make similarity judgments. We predicted that experts, and fans would show embeddings characteristic of a conceptual strategy (i.e., organising faces by team), whereas novices would show embeddings characteristic of a perceptual strategy (i.e., organising faces by featural similarity). As predicted, expert embeddings were more compatible with a conceptual judgement strategy, although fans and novices had embeddings more compatible with a perceptual judgement strategy. Importantly, embeddings aligned with participants' self-reported strategies. In the second experiment, we presented 13 native Japanese speakers and 24 non-Japanese speakers with arrays of Japanese kanji characters from four semantic categories. We predicted that Japanese speakers (experts) would show more conceptually-structured embeddings while non-Japanese speakers (novices) would show more perceptually-structured embeddings. While novice embeddings and self-reported strategies were consistent with perceptual judgments, the embeddings and self-reported strategies of Japanese speakers were consistent with both conceptual and perceptual strategies. Crucially, self-reported strategy use was highly related with embedding structure. Overall, we provide evidence for the viability of using psychological embeddings to measure individual differences in perception and representation.

This research was supported by a National Sciences and Engineering Research Council grant held by the second author

36.321 UNVEILING THE ORIGIN OF THE WORD-SPECIFIC AREA WITH THE OBJECT SPACE MODEL

Jia Yang[1] (jiayang@pku.edu.cn), Yipeng Li[1], Haoxuan Yao[1], Jingqiu Luo[1], Hongyu Li[1], Xiaoya Chen[2], Shiming Tang[1], Pinglei Bao[1]; [1]Peking University, [2]Vanderbilt university

Words constitute a unique, experience-dependent category within the representational space of the human ventral pathway. The prevailing view holds that learning to read repurposes a pre-existing region in the ventral occipitotemporal cortex for the recognition of written words. However, the initial function of this prototypical region (visual word form area, VWFA) remains elusive. In this study, by leveraging deep learning neural networks, we initially show that considerable word discrimination capacity can be derived from general non-word object recognition training. We find that objects similar to words in the network's 'object space' share more features that help in word recognition. This is mirrored in the human brain: our fMRI studies show that the objects that are closer to the words in the object space elicit higher responses in the VWFA. More importantly, such an effect is even true in the inferotemporal (IT) cortex of macaques, which are presumed to be naïve to words, both in terms of evolutionary history and visual experience. By utilizing fMRI techniques, we have successfully identified word-selective areas in the macaques' IT cortex. Furthermore, we measured the responses to 1000 words and

1000 objects within the anterior IT cortex of two macaques, using wide-field imaging techniques. The results align with the findings in the human brain: the word area responds more to objects close to words in the object space. Through integrating findings from CNN, human fMRI, and fMRI and wide-field imaging in macaques, our work highlights the possibility that VWFA may have initially evolved to represent features of non-word objects that are closely related to words in the object space and shed light on the general principles governing the genesis of category-specific areas in the IT cortex.

This work was supported by National Science and Technology Innovation 2030-Major Project 2022ZD0204803 to P. B., Natural Science Foundation of China Grant 32200857, China Postdoctoral Science Foundation Grant 2023M740125, 2022T150021, 2021M700004 to J.Y.

36.322 CROSS-SPECIES AND CROSS-MODALITY STUDIES OF FOOD-SPECIFIC BRAIN REGIONS

Baoqi GONG[1] (g1470645805@outlook.com), Haoxuan YAO[2], Yipeng LI[1], Wanru LI[1], Shiming TANG[123], Pinglei BAO[134]; [1]Peking-Tsinghua Center for Life Sciences, Beijing 100871, China, [2]Peking University School of Life Sciences, Beijing 100871, China, [3]IDG/McGovern Institute for Brain Research at Peking University, Beijing 100871, China, [4]School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China

Research in both humans and non-human primates has revealed that the inferotemporal (IT) cortex contains specialized areas responsive to vital object categories like faces, bodies, and natural scenes, which are essential for survival and daily activities. Recent studies revealed a new category-specific area, the food area, by analyzing 7T fMRI data when subjects viewed the Natural Scenes Dataset (NSD). However, these findings are somewhat constrained as they are based on a narrow range of datasets and exclusively human participants. Our study addressed these constraints by employing human fMRI to measure brain responses to two kinds of food stimuli: Chinese (e.g., Baozi) and Western foods (e.g., Pizza), juxtaposed with non-food stimuli in a blocked design format. This method identified specific subregions in IT cortex showing a preference for food images. Interestingly, there was a stronger activation for Chinese foods, reflecting the dominant dietary exposure of the subjects. In expanding our research to non-human subjects, we scanned two macaques with two stimulus sets, each containing food and non-food images. One set came from NSD as natural stimuli, and the other comprised single objects. Both sets, despite their different low-level properties, indicated a consistent food-responsive network in the IT cortex with three distinct patches from posterior to anterior. Complementary to this, widefield calcium imaging conducted on the macaques identified a food-preferred area in the anterior IT cortex. Intriguingly, unlike human subjects, the area showed a marked preference for fruit images. These outcomes not only affirm the existence of a dedicated food area within the IT cortex across different species but also underscore the significant impact of visual experience in the formation and specialization of these food areas.

National Science and Technology Innovation 2030-Major Project 2022ZD0204803, Natural Science Foundation of China Grant NSFC 32271081,32230043 to P. B. Natural Science Foundation of China Grant NSFC32200857, China Postdoctoral Science Foundation Grant 2023M740125, 2022T150021, 2021M700004 to J.Y.

SUNDAY, MAY 19, 2:45 – 6:45 PM, BANYAN BREEZEWAY

Scene Perception: Virtual environments, intuitive physics

36.323 CAN PEOPLE DETERMINE OBJECT DISTANCE FROM ITS VISUAL SIZE AND POSITION IN A CORRECTLY SCALED 2D SCENE DISPLAYED ON A LARGE SCREEN WITH ALIGNED GROUND PLANE?

Jong-Jin Kim[1] (johnk84@yorku.ca), Laurence Harris[1,2]; [1]Center for Vision Research, York University, [2]Department of Psychology, York University

INTRODUCTION: In our previous experiment (J Vis. 2022; 22(14):3310), we demonstrated that people cannot infer distance of an object reliably from its visual size and position in a 2D scene when viewed on a computer screen. Here we repeated the experiment with the visual stimuli projected onto a wall, fully displaying a simulated scene at natural size with the ground plane aligned with the floor. Might such a realistically scaled environment improve people's distance perception in a 2D scene? METHOD: Participants sat on a chair and viewed a hallway scene projected on a wall 2m away with their chin on a chinrest. They were asked to imagine they were looking into an actual hallway. Participants held a reference object (a cereal box) in their hands and compared its size to the image of a corresponding object in the scene. Participants adjusted either its SIZE based on its position in the scene (position-to-size task) or its POSITION based on its visual size (size-to-position task) to match the reference size. They did the tasks either binocularly or monocularly. RESULTS: In general, the adjusted target position was consistent with the object's size. The adjusted size, however, was consistently larger than geometrically correct when viewed binocularly but not monocularly. DISCUSSION: These results suggest that participants were more accurate at judging size based on position when conflicting stereo cues were removed despite reporting that using one eye was more difficult. Stereoscopic vision conflicts with reality when viewing a 2D simulation of a 3D scene and may underly the errors we see when judging the size of an object. The fact that we observed errors only when determining size and not position support the idea that size and distance perceptions use different mechanisms as suggested in our previous paper (Kim & Harris, 2022, Vision 6, 25).

NSERC, VISTA

36.324 ENHANCING WAYFINDING IN SIMULATED PROSTHETIC VISION THROUGH SEMANTIC SEGMENTATION AND RASTERING

Tori N. LeVier[1] (tori.levier@gmail.com), Justin Kasowski[1], Michael Beyeler[1]; [1]University of California, Santa Barbara

Introduction: Prosthetic vision offers the possibility of rudimentary vision restoration for blind individuals. Due to the limitations of current

devices, such as low resolution and perceptual distortions, simplifying the visual scene is crucial; for instance, by segmenting it into objects belonging to different semantic categories and then rendering either their outlines ("smart edge detection") or one category at a time ("smart rastering"). Here we evaluate these scene simplification strategies for wayfinding using simulated prosthetic vision in immersive virtual reality. Methods: We engaged 24 sighted participants (14 females, 10 males; ages 18-40) to navigate a virtual town square as "virtual patients." They used one of three rendering modes: naive edge detection, smart edge detection (outlining people, bikes, and buildings), or smart rastering (displaying these outlines category-wise). Each participant had 45 seconds to traverse the town square, avoiding stationary obstacles and moving cyclists. Performance metrics included path tracking, collision count, and success rate. After each session, participants rated the difficulty of each rendering mode. Results: Success rates improved from 39% with naive edge detection to 41% with smart rastering and 47% with smart edges. The smart modes reduced collisions, mainly with stationary objects (linear mixed-effects model, p<.01), but did not enhance safety on the bike path. Participants found the smart modes easier than the naive method. Conclusion: Smart edge detection and rastering improved wayfinding success rates and reduced collisions in this immersive task. However, less than half of the trials were successful, indicating a need for better scene simplification strategies. Future research should aim at enhancing judgment of moving objects' speed, direction, and approach time.

This work was supported by the National Library of Medicine of the National Institutes of Health (DP2-LM014268 to MB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

36.325 FLYINGOBJECTS: TESTING AND ALIGNING HUMANS AND MACHINES IN GAMIFIED OBJECT VISION TASKS

Benjamin Peters[1] (benjamin.peters@glasgow.ac.uk), Eivinas Butkus[2], Matthew H. Retchin[3], Nikolaus Kriegeskorte[2,3,4]; [1]School of Psychology & Neuroscience, University of Glasgow, UK, [2]Department of Psychology, Columbia University, [3]Zuckerman Mind Brain Behavior Institute, Columbia University, [4]Department of Neuroscience, Columbia University

Tasks lend direction to modeling and drive progress in both cognitive computational neuroscience and AI. While these disciplines have some shared goals, they have traditionally navigated the space of possible tasks with different intentions in mind, leading to vastly different types of tasks. Cognitive scientists and neuroscientists often prioritize experimental control leading them to use abstract tasks that remove many of the complexities of real-world experience, which are considered unrelated to the question at hand. AI engineers, by contrast, often directly engage the complex structure and dynamism of the real world, trading explainability for performance under natural conditions. However, AI engineers, too, are interested in gaining an abstract understanding of their models and cognitive computational neuroscientists ultimately want to model cognition under real-world conditions. If science and engineering are to provide useful constraints for each other in this area, it will be essential that they engage a shared set of tasks. Here we attempt to bridge the divide for dynamic object vision. We present a conceptual framework and a practical software toolbox called "FlyingObjects" that enables the construction of task generative models that span a vast space of degrees of naturalism, interactive dynamism, and generalization challenge. Task generators enable procedural sampling of interactive experiences ad infinitum, scaling between abstracted toy tasks and real-world appearance of objects and complex dynamics, access to and control over the task-generative variables, and sampling of atypical and out-of-distribution experiences. FlyingObjects connects science and engineering and enables researchers to acquire large-scale human behavioral data through smartphones, web-browsers, or in the lab, and to evaluate the alignment of humans and machines in dynamic object vision.

B.P. has received funding from the EU Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 841578.

36.326 PREDICTIVE PROCESSING OF UPCOMING SCENE VIEWS IN IMMERSIVE ENVIRONMENTS: EVIDENCE FROM CONTINUOUS FLASH SUPPRESSION

Anna Mynick[1] (anna.r.mynick.gr@dartmouth.edu), Michael A. Cohen[2], Adithi Jayaraman[1], Kala Goyal[1], Caroline E. Robertson[1]; [1]Dartmouth College, [2]Amherst College

Although our visual environment is immersive, we explore it in discrete and fleeting glimpses. How do we overcome our limited field of view to attain our continuous sense of visual space? Previous studies show that memory for a visual stimulus can speed perceptual awareness (Jiang et al., 2007). Here, we used virtual reality (VR) to test whether memory for immersive environments likewise facilitates perceptual awareness of upcoming scene views across head turns. Participants (N=29) first studied immersive, real world scenes drawn from the local campus in head mounted VR (Study Phase). In each trial of a subsequent priming task, a studied scene (prime) was first presented, then fully occluded. Participants then head-turned (left/right) toward a target image. The target, presented to the non-dominant eye, was initially masked by a dynamic Mondrian presented to the dominant eye (CFS). Participants' task was to detect the target, which either contained a spatially congruent view of the prime (e.g. the left view following a left head turn) or a spatially incongruent view (e.g. the right view following a left head turn). To ensure true target detection, only half of the target was displayed (a semi-circle) and participants indicated which side of the circle the target was on (left/right). Participants detected incongruent scene views faster than congruent ones (t(28) = 2.27, p = .031), suggesting that memory-based predictions affect the speed of perceptual awareness for scene information across head turns, favoring unexpected over expected input. This interpretation dovetails with a predictive processing account of visual processing, wherein top-down predictions suppress responses to expected sensory input, allowing deviations from expected input to be represented (Walsh et al., 2020). More broadly, this work underscores the possibility that memory-based predictions support efficient processing of visual input across fields-of-view as we move our eyes, heads, and bodies in space.

36.327 SCENE SEMANTIC AND GAZE EFFECTS ON ALLOCENTRIC CODING IN NATURALISTIC (VIRTUAL) ENVIRONMENTS

Bianca Baltaretu[1], Immo Schuetz[1], Melissa Vo[2], Katja Fiehler[1]; [1]Justus Liebig University, Giessen, Germany, [2]Goethe University Frankfurt, Frankfurt am Main, Germany

Interacting with objects in our surroundings involves object perception and object location coding, the latter of which can be accomplished egocentrically (i.e., relative to the self) and/or allocentrically (i.e., relative to other objects). Allocentric coding for actions under more naturalistic scenarios can be influenced by multiple factors, (e.g., task relevance and prior knowledge). Within the hierarchy of scene grammar, the semantic relationship of local objects (small/moveable) can strengthen allocentric coding (i.e., stronger effects for local objects of the same vs. different object categories). One would assume that the next level of the scene grammar hierarchy, i.e., anchor objects (large/stationary), also modulates this process, since anchors tend to predict the identity and location of surrounding local objects that we interact with. Here, we investigated the effect of semantically congruent versus incongruent anchors on allocentric coding of local objects within two scene types (kitchen, bathroom). In a virtual environment, three local objects were presented on a shelf connecting two anchors (semantically congruent or incongruent with the local objects). After a brief mask and delay, the scene was presented again without the local objects and one of the anchor objects shifted (leftward or rightward) or not shifted. Then, one of the local objects appeared in front of the participant, who then had to grab the object with the controller and place it in its remembered location on the empty shelf. Our findings show systematic placement errors in the direction of the anchor shift, with no clear influence of semantic congruency. Eye movements confirm these findings, with gaze behaviour predominantly directed toward local objects over anchors (with no effect of semantics when gaze landed on these). The present results suggest that, even if they are task-irrelevant, anchors play an important role in allocentric coding of local objects in naturalistic, virtual environments for action.

JUSTUS Plus II program, Justus Liebig University, Giessen, Germany

36.328 FIND THE ORANGE: HOW RICH AND ACCURATE IS THE VISUAL PERCEPT THAT GUIDES ACTION?

Aryan Zoroufi[1] (aryanz@mit.edu), Nishad Gothoskar[1], Josh Tenenbaum[1], Nancy Kanwisher[1]; [1]Massachusetts Institute of Technology

Are the visual representations that guide our online interactions with the world sparse and impoverished, or richly detailed, including the 3D shape of objects and their spatial and physical relationship to each other? We address this question using a naturalistic virtual reality environment in which participants (N=10) are asked to find an occluded target object (orange) on a tabletop environment as quickly as possible, by pressing a button to indicate which occluder should be moved first, or by reaching directly for the occluder. The occluders differ in width, orientation, 3D shape, and the presence of holes (which enable the participant to see through parts of the occluder). As instructed, people launch their actions quickly, within 500 ms after stimulus onset. We find that the decisions about which of two occluder objects to move first are guided by fairly accurate estimates of the area behind occluders that take into account 1) the 3D structure of the scene (not just the 2D pixel area of the occluders) and 2) the relative size of the hidden object. We also find that the decisions are similarly fast and accurate whether participants explicitly report or move the objects. Overall, these results suggest that a fast and accurate 3D representation of both visible and occluded parts of a scene are rapidly available to guide rational search in naturalistic environments. Future work using this framework will investigate whether the information that is rapidly available during naturalistic viewing includes not only geometric but also physical properties of the scene.

Thanks to NSF NCS Project 6945933 for funding the study

36.329 VISUAL CUES IN NONVISUAL COOKING: ASSESSING THE ROLE OF TACTILE AND AI-ASSISTED TECHNOLOGIES

Lily M. Turkstra[1] (lturkstra@ucsb.edu), Alexa Van Os[1], Tanya Bhatia[1], Michael Beyeler[1]; [1]University of California, Santa Barbara

Understanding how individuals who are blind navigate instrumental activities of daily living can provide crucial insights into the indispensable role of visual cues in these tasks. Cooking, a complex and multi-step process, relies heavily on visual information, from the selection of ingredients to gauging the readiness of a dish. While alternative senses and assistive technologies offer some aid, the specific visual cues that guide the cooking process have not been extensively studied. To address this, we present an observational analysis of nonvisual cooking, highlighting the visual cues integral to the task and examining the interaction between these cues and assistive technologies, particularly smartphone-based applications. Eight either legally or totally blind participants (35-74 years of age) were trained on how to navigate a kitchen and its appliances using tactile tools (i.e., Wikki Stix, bump dots) and utilize an AI-based smartphone app (either Microsoft Seeing AI or Google Lookout). Participants were instructed to bake a pizza under two task conditions: either by relying solely on tactile tools or by combining tactile tools with a smartphone app. Their verbalized thoughts and requests for researcher assistance were recorded, with question frequency and topics used to gauge the importance of different visual cues. Participants exhibited high independence, rarely asking for researcher assistance and predominantly relying on tactile aids over smartphone apps, even when digital tools were designed specifically for the task. Apps were only used primarily when tactile tools were inadequate for acquiring crucial visual cues, such as selecting the correct pizza topping, or identifying similarly packaged ingredients. Only when tactile tools and apps failed in tasks like rolling out the dough did participants request researcher assistance. Our findings highlight a diverse range of user preferences and app usage patterns, providing valuable insights for the development of more effective assistive tools.

This work was partially supported by the University of California, Santa Barbara.

36.330 NO EFFECT OF REDUCING VISUAL REALISM ON MOTION SICKNESS IN VIRTUAL REALITY

Jeffrey Saunders[1] (jsaun@hku.hk), Calise Lau[1]; [1]University of Hong Kong

Simulated navigation in virtual reality often causes motion sickness. We tested whether motion sickness is modulated by the amount of visual realism of the simulated environment. Visual realism has been found to affect sense of presence, and some evidence suggests that motion sickness in virtual reality depends on presence. The only previous study that tested for a relationship between visual realism and motion sickness was small and did not control low-level motion information. We performed two experiments that compared motion sickness in conditions with different levels of visual realism but matched motion information. In the baseline realism condition, subjects navigated through a virtual town constructed of moderately realistic models with detailed texture maps. For the low realism conditions, the virtual town was modified to appear more cartoon-like without affecting motion information: complex 3D models were replaced with simple geometric shapes and detailed texture maps were replaced with simple patterns (Experiment 1) or solid shading (Experiment 2). Subjects performed both conditions in counterbalanced order. In each condition, subjects reported motion sickness (SSQ) and sense of presence (PQ) after navigation. We found no effects of visual realism. There was no difference between the motion sickness reported in the baseline and low realism conditions of either Experiment 1 (N=30) or Experiment 2 (N=24), and no differences in presence ratings. The lack of effect on presence might have been due to the limited realism of our baseline condition and the fact that subjects performed an active navigation task. Our results demonstrate that visual realism of a virtual environment can be varied over a wide range without affecting the tendency to produce motion sickness.

This research was supported by a grant from the Hong Kong Research Grants Council GRF 17407914.

36.331 FROM THE FLOW OF LIQUIDS TO THE FLOW OF TIME: GRANULARITY OF SPONTANEOUS LIQUID FLOW PREDICTIONS IN VISUAL PERCEPTION IMPACTS EXPERIENCED TIME

Yuting Zhang[1], Joan Danielle K. Ongchoco[1], Ilker Yildirim[1]; [1]Yale University

The flow of time ordinarily guides intuitive physics: in a visual scene, we spontaneously predict what will happen next, and can reason about what might have happened. Does intuitive physics in return impact how we experience time? Experienced duration of time is malleable and this plasticity has been suggested to reflect the intensity of cognitive processing, with more processing corresponding to subjective durations lasting longer than actual durations. Here, we suggest that spontaneous intuitive physical predictions during visual perception adaptively engender representations of various granularities, and such representational granularity impacts the experienced duration of time. A coarse-grained representation might suffice to predict whether a liquid will flow toward right or left, while a finer-grained simulation may be necessary to work out details of its trajectory. We predict subjective time will slow down in the latter case. Observers watched animations with differently-oriented “planks”, through which liquids would flow. To manipulate the granularity of simulations, we placed a row of vertical pegs at the bottom faced either Upwards (to encourage finer-grained processing — since the liquid could fall into only some of the narrow openings), or Downwards (to encourage coarse-grained processing — since it doesn't matter here where the liquid would fall). Without an overt intuitive physics task (i.e., no mention of prediction), observers simply reproduced the duration of animations by holding a key down. Across experiments, observers experienced animations with Upwards pegs as lasting longer than those with Downwards pegs. Critically, granularity independently modulated subjective time above and beyond mere predictability since these results held across animations that were highly predictable (when the liquid's viscosity made it more trackable) and ones that were more unpredictable (when using sparser Atari-like pixelated displays). Thus the granularity of our mental simulations may influence our experience of the flow of time itself.

36.332 LEARNING OR DOING? VISUAL RECOGNITION OF EPISTEMIC VS. PRAGMATIC INTENT

Sholei Croom[1] (scroom1@jhu.edu), Hanbei Zhou[1], Chaz Firestone[1]; [1]Johns Hopkins University

Whereas some actions are aimed at changing the world, others are aimed at learning about it. For example, someone might press on a door to open it, or to determine whether it's locked; someone might place their toe into a pool to enter it, or to gauge its temperature; someone might shake a container to shuffle its contents, or to figure out what's inside. The distinction between ‘pragmatic’ and ‘epistemic’ actions is recognized in other fields, but only recently entered vision science: In previous work (Croom et al., 2023), we found that, when watching videos of someone shaking a box, observers can infer what information they are trying to obtain (e.g., the number of objects inside vs. their shape). Here, we ask a broader question: Do epistemic actions share common visual features that distinguish them from pragmatic actions, even beyond particular action goals? We created a set of 216 videos, each showing a naive participant completing an epistemic action (determining the number, shape, or size of objects in a box) or a pragmatic action (shuffling the box's contents, making the objects collide, or causing them to jump into the air). Then, 100 observers viewed these videos and were given a different task: To distinguish pragmatic actions from epistemic actions—i.e., who was acting to do something vs. to learn something. While some observers were given details about the specific actions they would see, other observers were simply told that some videos showed ‘learning’ and others showed ‘doing’. Regardless of whether they were informed (Experiment 1) or uninformed (Experiment 2) of the candidate actions, observers correctly distinguished pragmatic from epistemic actions, based purely on the box-shaking dynamics. Thus, learning looks different from doing: Beyond recognizing the particular goals of an action, observers can visually recognize epistemic vs. pragmatic intent.

NSF BCS #2021053

36.333 SOCIAL AND PERCEPTUAL ATTRIBUTIONS DERIVED FROM MOVING SHAPES: A LANGUAGE MODEL ANALYSIS

Emily D Grossman[1] (grossman@uci.edu), Sajjad Torabian[1], Zhenze Zhang[1], John A Pyles[2], Hongjing Lu[3]; [1]University of California Irvine, [2]University of Washington, [3]University of California Los Angeles

Introduction. In 1944, Heider and Simmel showed that humans spontaneously generate social interpretations when viewing sparse animations with moving shapes. Further studies have followed, investigating how motion trajectories are recognized as human actions (Roemmele et al. 2016), how cuing can elicit social meaning in simple animations (Tavares et al. 2008), and how motion within a context can drive attributions of beliefs (Baker et al. 2017). Here, we combine and compare attributions of actions, intentions, emotions, and beliefs elicited by animations, as this intersectional approach has been understudied. Specifically, we transform humans' descriptions of Heider-Simmel like animations to a semantic space, where we can then examine representational structures that underlie perceptual and cognitive processes. Methods. Each participant viewed two subsets of 100 animations, while labeling the gist of each animation with single keywords in the first phase, and choosing from a predefined list of labels in the second phase. The list was created based on previous literature and was broadly categorized into action, intention, and belief. Human labels of each animation were then embedded into a semantic vector using Google's Universal Serial Encoder (USE) language model. We generated three models capturing emotion, interactivity, and mental-state attribution, and correlated each with the semantic similarity structure of the animations. Results. A network frequency analysis showed that participants most saliently identified emotional narratives from the sequences, with nodes of negative emotion and avoidant action appearing as hubs. The semantic structure of the animations as observed by the participants was strongly correlated with the emotion model, followed by the model of interactivity and more weakly by the mentalistic model. Conclusion. Humans are sensitive to perceiving emotional attributes from animations of moving shapes, as compared to action- and belief-based attributes.

NSF BCS-1658560 to EG and BCS-1658078 to JP

36.334 TESTING MENTAL COMPUTATIONS OF CENTER OF MASS USING REAL-WORLD STABILITY JUDGMENTS

Giuliana Bucci-Mansilla[1] (gbuccim1@jhu.edu), Jason Fischer[1]; [1]Johns Hopkins University

In many of our daily activities, we arrange objects to create stable configurations. For instance, when placing glasses on the tray, we tend to avoid clustering all glasses at one edge or grouping heavy and light glasses on opposite sides. These decisions are intuitive and effortless, although they might not be flawless. What are the mental computations underlying our everyday stability judgments? We hypothesized that people rely on a Newtonian mental model, computing the center of mass -CoM- to judge the stability of complex systems, rather than relying on heuristics guided by mainly perceptual cues (object size, position and shape). To test this idea, we designed a paradigm that allows us to assess when people employ CoM versus other heuristics. Our real-life experiment consisted of a tilted platform supporting 6 blocks -3 heavy and 3 light-, all of them at different positions, resembling a tray full of glasses. Participants were tasked with making the platform as flat as possible by removing only one block. We manipulated the variables "weight" and "position", in such a way that the best block to remove varied in position and weight trial by trial. With this design, no single visual heuristic would yield successful performance, allowing us to test whether participants nonetheless employed heuristic-based strategies. Our results show that participants predominantly employed a strategy utilizing CoM computations, consistently choosing blocks that moved the CoM of the platform to a balanced state while exhibiting a slight bias toward choosing heavier blocks. People may not execute a flawless computation of the center of mass, but their approach is remarkably close. Our findings shed light on the cognitive processes involved in stability judgments in everyday tasks, emphasizing the nuanced interplay between physics-based computations and perceptual heuristics.

36.335 VELOCITY– NOT PERCEIVED AS SUCH: THE ROLE OF PERCEIVED MASS ON MOTION ESTIMATION

Abdul-Rahim Deeb[1] (abdul-rahim_deeb@brown.edu), Jason Fischer[1]; [1]Johns Hopkins University

In physics, velocity is conventionally defined as the first derivative of an object's position with respect to time. However, a longstanding assumption within vision science has equated the visual hierarchy of object motion with the principles of Newtonian mechanics– treating velocity as a unitary and fundamental variable, akin to its role in physics. This assumption has persisted despite compelling evidence suggesting that the human visual system is biased in perceiving velocity, influenced by factors such as viewing distance and angle. Contrary to the presumed universality of velocity as a visual primitive, numerous studies have revealed pronounced biases in human observers' ability to discern object motion, particularly influenced by the size of moving objects. Notably, larger objects are consistently perceived as moving more slowly. Building upon these findings, our study investigates the impact of perceived mass on the perceived speed of moving objects. Through a 2-AFC paradigm involving objects of varying sizes and materials, our findings unveil a discernible trend: as objects increase in perceived mass due to size or apparent density, their perceived speed decreases. In light of these observations, we challenge the prevalent assumption that velocity operates as a visual primitive. Instead, we contend that momentum, rather than velocity, may offer a more suitable conceptual framework within object motion perception, acknowledging the inherent integration of information about mass in the visual processing of motion.

36.336 PERCEIVING ANIMACY THROUGH SCHEMATIC INTUITIVE PHYSICS: SHARED CONCEPTUAL STRUCTURE OF ANIMACY BETWEEN VISION AND LANGUAGE

Ning Tang[1] (ningtangcog@gmail.com), Enjie Xu[1], Siyi Gong[2], Jifan Zhou[1], Mowei Shen[1], Tao Gao[2]; [1]Zhejiang Univeristy, [2]UCLA

In the real world, agents are not omnipotent. Their motions often deviate from intended goals due to physical constraints, as exemplified by a leashed dog moving erratically while chasing a cat. Humans typically perceive this animacy, yet empirical psychophysical studies contradict this observation, indicating that spatial deviations and line connections between objects can impair perceived animacy. In this study, we propose that these limitations are special cases within a more unified theory, where animacy is inferred by intelligently interpreting various forces imposed on the agent through the lens of intuitive physics. Previous failures are caused by physically arbitrary deviations and connections. We aim to investigate whether perceived animacy improves when deviations from goal-directed motion are explained away by a constraint imposed by a physical line. Moreover, we examined two intuitive physics models in perceived animacy: first, a realistic Newtonian Physics model implemented by a modern physics engine; second, a schematic force dynamics model from cognitive linguistics, where an agonist's motion is determined by its intrinsic force and the force imposed by an antagonist. While the former excels in explaining perceptual physical judgments, the latter offers a linguistic framework of the conceptual structure underlying verb usage in language. Our findings reveal that perceived animacy, assessed through free reports and visual searches, significantly increases when deviations and connections are physically explainable. Importantly, the force dynamics model yielded higher perceived animacy than Newtonian Physics. Extensive testing of physical parameters confirmed that the weaker results of Newtonian physics were a generic phenomenon. Collectively, these results indicate that vision effectively perceives animacy by explaining spatial deviations and line connections through intuitive physics. The intuitive physics in perceived animacy is better modeled by force dynamics from linguistics, making it an interesting case study showing that there is a common conceptual structure underlying both vision and language.

SUNDAY, MAY 19, 2:45 – 6:45 PM, BANYAN BREEZEWAY

Spatial Vision: Models

36.337 EFFECTS OF NORDIC DENOISING ON POPULATION RECEPTIVE FIELD MAPS

Christian Windischberger[1] (christian.windischberger@meduniwien.ac.at), David Linhardt[1], Michael Woletz[1], Garikoitz Lerma-Usabiaga[2,3]; [1]Medical University of Vienna, Austria, [2]BCBL. Basque Center on Cognition, Brain and Language, San Sebastián, Spain, [3]IKERBASQUE. Basque Foundation for Science, Bilbao, Spain

While achieving higher spatial resolution is an important goal in fMRI studies, it also results in reduced signal-to-noise ratio (SNR). NORDIC is a noise reduction method based on patchwise principal component analysis (PCA) and increases SNR by removing thermal noise components. Critically, the NORDIC approach should reduce spatial smoothing effects that were frequently described with other noise-reduction methods. Here we investigate NORDIC's effects on population receptive field (pRF) mapping, particularly on pRF size estimation. We acquired fMRI data from three healthy participants using a SIEMENS PrismaFit 3T scanner. A bar aperture moving in eight directions was presented, revealing reversing checkerboards over a 9° radius field of view. NORDIC denoising was applied after standard scanner reconstruction. Both original and denoised data underwent minimal preprocessing using fMRIPrep, followed by pRF mapping analysis using containerized solutions prfprepare and prfanalyze-vista. Our findings show that NORDIC increases variance-explained values without inflicting differences in pRF position (eccentricity and polar angle), while pRF size estimations increase considerably (median 22% increase in pRF size with NORDIC). Cohen’s d effect sizes show a small effect on pRF size and a large effect on variance explained. It might be concluded that this increase in pRF sizes is caused by an increase in SNR due to NORDIC. Our simulations, however, clearly show that increased SNR yields reduced pRF sizes. We also studied the effects of spatial smoothing in the pRF sizes estimated and found that spatial smoothing leads to increases in pRF sizes. Taken together, these increases in pRF size seem not to be directly linked to spatial image smoothness, but may arise from the retinotopic organization of neighboring voxels. Although NORDIC only marginally increases image smoothness, its impact on pRF size estimations necessitates careful interpretation. Our results underscore the importance of considering NORDIC's influence on pRF size in fMRI preprocessing.

Austrian Science Fund P33180, P35583

36.338 COMPARING PRF MAPPING ESTIMATES FOR WORDS AND CHECKER PATTERNS

David Linhardt[1] (david.linhardt@meduniwien.ac.at), Yongning Lei[2], Pedro M. Paz-Alonso[2,3], Brian Wandell[4], Christian Windischberger[1], Garikoitz Lerma-Usabiaga[2,3]; [1]Medical University of Vienna, Austria, [2]BCBL. Basque Center on Cognition, Brain and Language, San Sebastián, Spain, [3]IKERBASQUE. Basque Foundation for Science, Bilbao, Spain, [4]Stanford University, Department of Psychology

Functional MRI responses from voxels in the visual cortex are driven by stimuli within restricted regions of the visual field, their so-called population receptive fields (pRF). The central position and size of every voxel’s receptive field can be quantified using pRF mapping. In a previous report, we measured the pRF centers of individual voxels using words and contrast-reversing checker patterns shown within the stimulation area. The pRF centers measured with words differed from those measured with checkers. Voxels with a pRF center in the near periphery (5-10 degrees visual angle from the fixation) are more eccentric (1-3 degrees) when measured with checkers compared to words. To gain a deeper understanding of these effects, we acquired new datasets that differed significantly from the previous data. These variations included using different MRI scanners, fields of view, and acquisition sequences with either high-temporal or high-spatial resolution. Additionally, we adapted the used stimuli (such as variations in bar width and speed, flickering frequency) and participant populations, including individuals with both uncorrected and corrected visual acuity. Data were analyzed using a highly reproducible containerized public analysis platform (prfprepare and prfanalyze-vista). Results confirm the main effect in higher visual areas (hV4, VO1, IPS0-1). Further, initial measurements suggest specific stimulus manipulations (including defocus) impact the size of the change in eccentricity. Moreover, these manipulations may have different impacts on the eccentricity shift measured in different visual field maps. These findings offer a compelling starting point to further investigate stimuli induced pRF differences.

This work was supported by the Austrian Science Fund (FWF; P35583), the Spanish Ministerio de Ciencia e Innovación (IJC2020-042887-I; PID2021- 123577NA-I00, RYC2022-035502-I; CEX2020-001010/AEI/10.13039/501100011033) and Basque Government (BERC 2022-2025; PIBA_2022_1_0014; KK-2023/00090).

36.339 THE INFLUENCE OF ATTENTIONAL LOAD ON POPULATION RECEPTIVE FIELD PROPERTIES

Sumiya Sheikh Abdirashid[1,2,3] (s.abdirashid@spinozacentre.nl), Tomas Knapen[1,2,3], Serge O. Dumoulin[1,2,3,4]; [1]Spinoza Center for Neuroimaging, KNAW Netherlands, [2]Computational Cognitive Neuroscience and Neuroimaging, Netherlands Institute for Neuroscience, KNAW Netherlands, [3]Experimental and Applied Psychology, Vrije University Amsterdam, Netherlands, [4]Experimental Psychology, Helmholtz Institute, Utrecht University, Netherlands

Spatial attention enhances perception by attracting population receptive fields (pRFs) towards the attended locus. Our recent work has shown the degree of this attraction can be altered by precision of attention. In addition to spatial features of attention, understanding how qualitative aspects of attention, such as attentional load, influence pRF properties requires further investigation. The Gaussian attention field model is often used to summarize the locus and precision of attention by a Gaussian. In this context, attentional load is the amplitude of the attention field. Notably, this model predicts no effect of altering the amplitude of the attention field. Based on previous results, we added an offset of 1 to the attention field (AF+1 model), this better captured the effects of precision of attention while maintaining the core prediction of attraction towards the attended locus. In contrast, the AF+1 model predicts the amplitude of the attention field does influence resulting pRFs. Here, we compared these model predictions and investigated the effect of attentional load on pRF properties. We used 7T MRI to map pRFs while participants carried out an RSVP detection task at fixation. A stream of symbols which varied in color and orientation was presented at a rate of 5.3 Hz. Two attention conditions with identical stimuli were compared: low load where participants detected only a target symbol regardless of color and orientation and high load where participants detected two targets symbols each of a specific color and orientation combination. Additionally, separate pRF-mapping sessions were carried out to obtain pRF parameters independent of the two conditions. Behavioral results showed that the difficulty of the two conditions varied, with higher D' values for the low-load condition compared to the high-load condition. Preliminary data shows some pRF property differences between conditions, indicating that attentional load alters visual representations of space.

36.340 A NOVEL APPROACH FOR POPULATION-RECEPTIVE FIELD MAPPING USING HIGH-PERFORMANCE COMPUTING

Siddharth Mittal[1] (siddharth.mittal@meduniwien.ac.at), Michael Woletz[1], David Linhardt[1], Christian Windischberger[1]; [1]Medical University of Vienna

Population-receptive field (pRF) mapping is a tool for mapping visual information encoding in the brain using fMRI. The pRFs are typically modelled as 2D Gaussian, where the mean (μ_x, μ_y) signifies location, and the variance (σ) indicates the receptive field size. Traditional pRF mapping tools like mrVista and SamSrf yield accurate results but require long computation times. Conversely, newer methods like fast-pRF from CNI_toolbox favour speed over accuracy. To bridge this gap, we propose the novel implementation GEM-pRF (GPU-Empowered Mapping of pRF), which combines high accuracy with greatly shortened computation time. This method involves two steps: (1) initial grid search for pRF positions and size estimates (μ_x, μ_y, σ) executed through efficient matrix operations on high-performance GPU cores and (2) approximating the residual sum of squares (RSS) error function as quadratic, using partial derivatives in the neighbourhood to refine estimates. We evaluated the accuracy of our results using simulated and real fMRI data. For simulated data, we employed a validation framework by Lerma-Usabiaga et al. (2020), to create a noisy simulated dataset with known pRF parameters. Our implementation (GEM-pRF) was compared with mrVista and SamSrf, showing similarly high accuracy in pRF parameter estimation. For analysis on real data, we scanned a healthy male on a 3T Siemens PrismaFit scanner (CMRR EPI, TR/TE=1000/38ms, 1.5mm iso, MB=3). Using this real data, we compared our method's pRF parameter estimations with mrVista (a commonly used tool for pRF mapping). The results revealed a high correspondence between the two. Notably, our method estimates pRF parameters for a 10,000-voxel fMRI dataset in just 30-40 seconds, a significant improvement from the approx. 10 minutes taken by mrVista. These findings underscore the remarkable speed and maintained accuracy of our GPU-based implementation, enabling comprehensive analysis of large datasets and unlocking new possibilities for exploration with complex pRF models.

This work was supported by the Austrian Science Fund (FWF; grant numbers: P35583)

36.341 POPULATION RECEPTIVE FIELD MODELS CAPTURE EVENT-RELATED MEG RESPONSES

Kathi Eickhoff[1,2,3], Arjan Hillebrand[4], Maartje C. de Jong[1,2,5], Serge O. Dumoulin[1,2,3,6]; [1]Spinoza Centre for Neuroimaging, Amsterdam, the Netherlands, [2]Netherlands Institute for Neuroscience, Amsterdam, the Netherlands, [3]Vrije Universiteit, Amsterdam, the Netherlands, [4]Department of Clinical Neurophysiology and Magnetoencephalography Centre, Amsterdam UMC location Vrije Universiteit Amsterdam, the Netherlands, [5]University of Amsterdam, the Netherlands , [6]Utrecht University, the Netherlands

The visual system is organized retinotopically. In humans, this organization can be studied non-invasively by estimating the receptive fields of populations of neurons (population receptive fields; pRFs) with functional magnetic resonance imaging (fMRI). However, fMRI is too slow to capture the temporal dynamics of visual processing that operate on the scale of milliseconds. Other non-invasive techniques such as magnetoencephalography (MEG) provide this temporal resolution while lacking the spatial resolution to disentangle the precise locations of pRFs in the cortex. Here, we introduce a forward modeling approach that combines fMRI's spatial- and MEG's temporal resolution, enabling us to estimate pRFs on the neuronal timescale. Using fMRI, we estimated the participants pRFs using conventional pRF-modeling. With MEG, we measured event-related field (ERF) responses while the participants (N=5) viewed briefly presented (100ms) contrast-defined bar and circle shapes. Next, we combined the pRF models with a forward model that maps the pRFs' electrical brain activity to the sensor level, to predict MEG's sensor-level responses to the stimuli. We computed the goodness of fit between the predicted and measured ERF responses at each time point using cross-validated variance explained. To evaluate whether the recorded ERFs were optimally fitted by the pRFs, we moved the pRF positions away from the fMRI-estimated positions and refitted the new predictions to the experimental ERF data. We found that the fMRI-estimated pRFs explained up to 90% of the variance in individual MEG sensor's ERF responses. The variance explained varied over time, but the pRF model accurately captured the ERF responses between 80ms to 250ms after stimulus presentation. Perturbing the pRF positions decreased the explained variance, suggesting that the ERF responses were driven by the pRFs. In conclusion, pRF models accurately capture event-related MEG responses, enabling routine investigation of the spatiotemporal dynamics of human pRFs with millisecond resolution.

36.342 SPATIAL FREQUENCY TUNING IN EARLY VISUAL CORTEX IS NOT SCALE INVARIANT

Michaela Klimova[1,2], Emily Wiecek[1,3,4], Luis Ramirez[1], Sam Ling[1]; [1]Boston University, [2]Northeastern University, [3]Boston Children's Hospital, [4]Harvard Medical School

The visual system can distinguish patterns over a wide variety of spatial scales. Based on physiological evidence, this robustness is thought to reflect the retinotopic scale-invariant organization of primary visual cortex, whereby the spatial area over which a cell responds is proportional to the spatial information that cell is tuned to convey. Specifically, larger receptive fields (RFs) when moving away from the foveal representation reflect the integration of information from increasingly larger regions of the visual field and a reduction in spatial sampling density. The assumption made by scale invariance is that the information sampled by a RF is constant across the visual field. Model-based fMRI methods, such as population receptive field (pRF) and population spatial frequency tuning (pSFT) analyses, have enabled the characterization of spatial tuning (eccentricity and size) and SF preferences (peak preference and tuning bandwidth) of cortical sub-populations sampled in a voxel. Using these tools, we evaluated whether scale invariance, defined as a ratio between peak pSFT and pRF size (cycles/RF), is constant across the visual field in early visual areas V1–V3 (N=8). We found a marked departure from scale invariance: there was an expansive increase in cycles/RF for voxels near the central visual field, and an eventual decrease for voxels that prefer higher eccentricities. This departure from scale invariance at innermost eccentricities was most pronounced in V1. In extrastriate cortex (V2–V3), the relationship of cycles/RF and eccentricity appeared more constant across the visual field — closer to the predictions of scale invariance. Taken together, these results reveal that SF preference does not scale linearly with RF size, a central assumption commonly used to adjust for cortical magnification factors.

This work was funded by the National Institutes of Health (NIH) Grant R01EY028163 to S. Ling, NIH Grant K23EY034212 to E. Wiecek, NIH Grant F99NS124144 to L. Ramirez, and involved use of equipment funded by the National Science Foundation (NSF) Major Research Instrumentation Grant 1625552.

36.343 HIERARCHICAL GAUSSIAN PROCESS MODEL FOR HUMAN RETINOTOPIC MAPPING

Sebastian Waz[1] (scw8734@nyu.edu), Yalin Wang[2], Zhong-Lin Lu[13]; [1]New York University, [2]Arizona State University, [3]NYU Shanghai

In human subjects, visual neuroscience often depends, critically, on the estimation of retinotopic maps from BOLD fMRI. Intepreting brain activity as it relates to basic visual perception requires knowledge of the boundaries of (e.g.) areas V1, V2, and V3 on the cortex and the mapping of retinal eccentricity and polar angle onto these areas. Any approach to drawing the areas' boundaries and the mappings therein is limited in accuracy by the resolution of the fMRI scanner and is susceptible to noise. These issues can be addressed using hierarchical Bayesian modeling, wherein information is pooled across subjects, with potential to enhance estimates in both precision and unbiasedness. We introduce a novel hierarchical Bayesian Gaussian process model (HGP) to estimate retinotopic maps for 162 subjects of the Human Connectome Project (Van Essen et al., 2013; Uğurbil et al., 2013), taking population receptive field model estimates (Dumoulin and Wandell, 2008) on these data as input to the HGP. Subjects were functionally aligned using optimally chosen wedges from their cortical ROIs. We found that the standard deviation across subjects of the estimated V1 ventral-dorsal boundary was 0.042 wedge-arcs on average, ranging from 0.036 to 0.049 wedge-arcs along the length of this boundary. For the contour line at 4° eccentricity, the standard deviation was 0.056 wedge-radii on average, ranging from 0.044 to 0.070 wedge-radii along this contour. At the population level, a systematic distortion of the eccentricity map was observed along the ventral V1/V2 boundary. Future work will be done to assess how the hierarchical framework in HGP mitigates noise-induced bias when generalized to a new data set. Since the HGP was fitted to relatively high-resolution 7T scanner data, it may prove especially valuable as a prior for 3T scanner data which are commonly used but require additional constraints to accurately estimate retinotopic maps.

Supported by National Eye Institute Grant R01 EY032125

36.344 DIFFEOMORPHIC REGISTRATION ENHANCES RETINOTOPIC MAPPING IN 3T

Negar Jalili Mallak[1] (njalilim@asu.edu), Zhong-Lin Lu[2,3,4], Yalin Wang[1]; [1]School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA., [2]Division of Arts and Sciences, NYU Shanghai, Shanghai, China., [3]Center for Neural Science and Department of Psychology, New York University, New York, USA., [4]NYU-ECNU Institute of Brain and Cognitive Science, NYU Shanghai, Shanghai, China.

Retinotopic mapping, a fundamental component of visual cognitive neuroscience, helps us understand how the brain processes visual stimuli. Despite its significance, the process relies on BOLD functional magnetic resonance imaging (fMRI), which has a low signal–noise ratio (SNR) and low spatial resolution. These limitations impede the creation of accurate and precise retinotopic maps. This study introduces a novel application of Diffeomorphic Registration for Retinotopic Maps (DRRM; Tu, et al, 2022) to enhance the alignment of retinotopic maps using the 3T NYU Retinotopy Dataset, encompassing analyze-PRF and mrVista results. Diffeomorphic Registration for Retinotopic Maps (DRRM) quantifies the diffeomorphic condition, ensuring accurate alignment of retinotopic maps under topological conditions by leveraging the Beltrami coefficient. We evaluated the quality of the registered retinotopic maps utilizing visual coordinate change, flipped triangles, and goodness of fit to BOLD time series. Minimized visual coordinate changes and the absence of flipped triangles validate the diffeomorphic nature of DRRM registration. The results indicate that the DRRM-registered retinotopic maps provide a superior fit to the fMRI time series which is evident in the reduced Root Mean Square Error (RMSE) from the DRRM fits (mean RMSE=1.06) compared to the structurally registered retinotopic maps and Benson's inference map (mean RMSE= 1.3). This improvement suggests that visual coordinates from the DRRM provided a better account of the fMRI time series than the original population receptive field solutions. In conclusion, our work demonstrates that DRRM is a valuable tool that can significantly improve the quality of retinotopic maps in the realm of 3T fMRI data. This successful adaptation positions DRRM as a promising method for advancing retinotopic map research and applications, addressing the challenges posed by the limited SNR and spatial resolution of BOLD fMRI and enhancing the accuracy and interpretability of retinotopic maps.

The work was partially supported by NSF (DMS-1413417 and DMS-1412722) and NIH (R01EY032125 and R01DE030286)

36.345 BEHAVIORAL AND NEURAL SIGNATURES OF EFFICIENT SENSORY ENCODING IN THE TILT ILLUSION

Ling-Qi Zhang[1,2] (zlqzcc@gmail.com), Geoffrey K. Aguirre[1], Alan A. Stocker[1]; [1]University of Pennsylvania, [2]Janelia Research Campus, Howard Hughes Medical Institute

Human perception of orientation is shaped by spatial and temporal context. For example, in the tilt illusion, the surround orientation induces a distinctive bias pattern in the perceived orientation of the center (Gibson, 1937). At the neuronal level, it has been shown that stimulus context alters the response properties of visual neurons (e.g., Benucci, Saleem & Carandini, 2013). Connecting from neural coding to behavior, however, is difficult as it requires specific assumptions about how orientation is represented (encoding) and interpreted (decoding). Here, we conduct a study that characterizes the surround modulation of orientation encoding simultaneously at both levels. Across 1,200 trials during fMRI scanning, each of 10 subjects estimated the orientation of a briefly displayed grating (1.5s, 1 Hz contrast modulation) by rotating a line probe after a short delay (3.5s - 4.5s). The stimuli were presented within an annular surround of either non-oriented noise, or gratings with one of two fixed orientations (±30 degrees off vertical). We extracted encoding accuracy, expressed as Fisher information (FI), based on a lawful relationship between FI and the bias and variance in behavioral responses (Noel et al., 2021). We also obtained the neural population FI for each of the retinotopically organized visual areas by fitting voxel-wise probabilistic encoding models to the fMRI data (van Bergen & Jehee, 2018). At both behavioral and neural levels, sensory encoding for the stimulus condition with an unoriented surround accurately reflects the natural scene statistics of orientation. In the presence of the oriented surround, encoding accuracy is significantly increased at the corresponding surround orientation, which also matches the conditional orientation statistics in natural scenes. Further, the effect of surround modulation upon FI increased steadily across the ventral visual area hierarchy. Our results are consistent with the notion that contextual modulation represents a form of efficient coding.

This work is supported by a MindCORE Collaborative Research Grant from the University of Pennsylvania

36.346 A MODULAR IMAGE-COMPUTABLE PSYCHOPHYSICAL SPATIAL VISION MODEL

Jannik Reichert[1] (jannik.reichert@cis.mpg.de), Felix A. Wichmann[2]; [1]Max Planck Institute for Software Systems, [2]University of Tübingen

To explain the initial encoding of pattern information in the human visual system, the standard psychophysical spatial vision model is based on channels specific to spatial frequency and orientation, followed by divisive normalization (contrast gain-control). Schütt and Wichmann (2017, Journal of Vision) developed an image-computable implementation of the standard model and showed it to be able to explain data for contrast detection, contrast discrimination, and oblique and natural-image masking. Furthermore, the model induces a sparse encoding of luminance information. Whilst the model's MATLAB code is publicly available, it is non-trivial to extend, or integrate into larger pipelines because it does not provide a modular, pluggable programming framework. Based on the previous MATLAB implementation we developed a modular image-computable implementation of this spatial vision model as a PyTorch framework. Furthermore, we added a number of refinements, like a jointly spatially and spatial frequency dependent contrast gain-control. With luminance images as input, it is easy to employ the model on real-world images. Using the same psychophysical data, we compare our model's predictions of contrast detection, contrast discrimination, and oblique and natural-image masking with the previous implementation. The major advantage of our framework, however, derives from its modularity and the automatic differentiation offered by PyTorch as these facilitate the implementation and evaluation of new components for the spatial vision model. Furthermore, our framework allows the integration of this psychophysically validated spatial vision model into larger image-processing pipelines. This could be used to take inputs from retina models instead of from pre-computed luminance images or to further process the model's outputs with higher-level vision models. Given its flexibility, the model could also be used as a plug-in for or replacement of parts of artificial neural networks, which would enable comparison of aspects of human and machine vision.

JR was supported by CS@max planck – The Max Planck Graduate Center for Computer and Information Science. FAW is a member of the Machine Learning Cluster of Excellence, funded by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy – EXC number 2064/1 – Project number 390727645.

36.347 ANALYTIC MODEL OF RESPONSE STATISTICS IN NOISY NEURAL POPULATIONS WITH DIVISIVE NORMALIZATION

Daniel Herrera-Esposito[1] (dherrera1911@gmail.com), Johannes Burge[1]; [1]University of Pennsylvania

Divisive normalization is an essential neural computation that is commonly used in models of visual processing. However, normalization models rarely incorporate neural noise. This is not realistic. Here, we expand the classic model of normalization to include the effects of noise. The classic model is characterized by a noiseless unnormalized drive (e.g., linear-receptive-field outputs) that is divided by a scalar normalization signal. The normalization signal is, itself, a function of the unnormalized drives (e.g. L2-norm). The result is a normalized population response. We expand this classic model by incorporating noise into the unnormalized drive. We model the drive as multivariate Gaussian, entailing that noise affects both the numerator (drive) and the denominator (normalization signal) of the normalization equation. We derive analytic formulas for the mean and covariance of the normalized responses, given the unnormalized-drive statistics. The formulas can incorporate arbitrary noise correlations and different types of biologically-plausible normalization (e.g. broadband, feature-specific). The formulas are also differentiable, making them suitable for optimization routines used in computational models, and in fitting neural data. The behavior of the model is consistent with classic noiseless models (e.g. responses saturate at high contrasts), but non-obvious behaviors are also predicted. First, normalization transforms independent drive-noise into noise correlations in the normalized responses. Second, these normalization-induced response noise correlations are stimulus-dependent. Third, increasing noise in fixed-mean unnormalized drives reduces mean normalized responses. Our model includes elements

(e.g. large neural populations, dependence of normalization on noisy population drive) that are missing from other attempts to analytically model the interaction of normalization and noise. The formulas derived here describe in a principled way the consequences of normalization on noisy visual processing, and constitute a tool for quantitatively modeling the behaviors of real neural systems that have not previously been analytically linked to normalization.

This work was supported by the National Eye Institute and the Office of Behavioral and Social Sciences Research, National Institutes of Health Grant R01-EY028571 to J.B.

36.348 A HIGHLY REPLICABLE MODEL OF ACHROMATIC CONTRAST SENSITIVITY BASED ON INDIVIDUAL DIFFERENCES IN OPTICS AND SPATIAL CHANNELS: ROBUST CONSISTENCY IN FACTOR STRUCTURE ACROSS >6 VERY DIFFERENT DATASETS

David Henry Peterzell[1] (davidpeterzell@mac.com); [1]Fielding Graduate University, [2]National University (JFK School of Psychology)

For 40 years, individual differences and factor analysis have been used to investigate the underlying variable structure of achromatic contrast sensitivity in hopes of understanding underlying spatial mechanisms. Some perceive these factors to be volatile and difficult to replicate. However, the robust stability and consistency of underlying factors are shown here. Six+ studies were considered and reviewed, each using some methodologies that differed from the other studies (e.g,, adult vs. infant participants, psychophysical vs EEG/sweep-VEP thresholds; photopic vs. scotopic illumination; data collected by author vs archival data of others). The datasets provide a consistent and congruent set of results: 1) Below 2.25 c/deg, two factors correspond to Wilson's channels A and B. (Wilson's A and B channels were estimated using modeling masking data decades ago, with A and B of the 6-channel model peaking in at approximately .7 and 1.3 c/deg in adults, and at predictably lower spatial frequencies in infants). 2) Above 2.25 c/deg, a third factor corresponds to optical variation's influence on contrast sensitivity. That is, individual differences in contrast sensitivity are primarily determined by optics. This third factor was recently validated by an analysis of CSFs collected using adaptive optics. 3). For scotopic vision, there is an additional fourth factor tuned to very low spatial frequencies. This third factor is consistent with findings from adaptation and masking studies. The existence of a highly replicated factor structure confirms that contrast sensitivity variability below 2.25 c/deg is determined by two "Wilson" channels (adults with photopic viewing) or more channels (infants, adults with photopic viewing). Above 2.25 c/deg, optics determine variability in CSFS. Unless adaptive optics methods are used, CSF variability should not be used to estimate channel properties above 2.25 c/deg.

36.349 A ROBUST CO-VARIATION OF THE STIMULUS-SPECIFIC BIAS AND VARIABILITY ACROSS DIFFERENT VIEWING CONDITIONS AND OBSERVERS, AND ITS IMPLICATION ON THE BAYESIAN ACCOUNT OF ORIENTATION ESTIMATION

SANG HUN LEE[1] (visionsl@snu.ac.kr), Dong-gyu Yoo; [1]Seoul National University

The prevalence of cardinal orientations in our visual surroundings appears to be reflected in our perceptual system, possibly influencing our estimation performance in two error metrics. When assessed for bias, orientation estimates are typically displaced away from their nearest cardinal orientations, often referred to as the "repulsive bias." When assessed for variability, orientation estimates are typically more variable across trials for oblique orientations than cardinal orientations, often referred to as the "oblique effect." While these two phenomena have been widely reported, they have been studied in quite an isolated manner. Consequently, little effort has been made to relate the orientation-specific bias to the corresponding variability. Given the intimate interdependence between these two error metrics in the estimation theory, such as the Cramer-Rao bound, clarifying and accounting for the relationship between the orientation-specific bias and variability will refine the current understanding of how the visual system operates to estimate the state of its surroundings. In this perspective, we gathered the orientation estimation data from four independent studies with different viewing conditions and examined whether any significant relationship exists between the orientation-specific bias and variability. Here, we report one such relationship: as the repulsive effect grows in the bias, the oblique effect retreats while the anti-oblique effect (i.e., cardinal effect) begins to dominate the variability. When we quantified the strength of the repulsive bias and the dominance of the oblique effect in the variability, their relationship was well captured by a linear trend across the 10 different datasets ($r = -0.948$), and across individual observers within the datasets ($r = [-0.618, -0.703]$). By leveraging the "efficient Bayesian observer model" proposed by Wei and Stocker (2015), we attempted to offer a principled account for the observed bias-variability covariance, which points to the variation in "loss function" as a computational-level source of the covariance.

This work was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (No. NRF-2015M3C7A1031969, NRF-2017M3C7A1047860, NRF-2021R1F1A1052020), and SNU R&DB Foundation (339-20220013).

36.350 LIMITS ON HUMAN CONTRAST SENSITIVITY IMPOSED BY THE INITIAL VISUAL ENCODING

Fangfang Hong[1] (fh862@sas.upenn.edu), Nicolas P. Cottaris[1], David H. Brainard[1]; [1]University of Pennsylvania

Contrast sensitivity (CS) is a fundamental measurement that reveals the capabilities and limitations of human vision. Understanding the stimulus dependence of CS provides links to the underlying neural mechanisms and supports applications ranging from display optimization to diagnosis of retinal disease. Our broad objective is to establish principles and computations that enable prediction of CS from models of the initial visual encoding. Here we consider how encoding by the midget retinal ganglion cells (mRGCs) shapes CS for achromatic (L+M+S) and chromatic (L-M) patterns, presented at multiple eccentricities. We modeled cosine-windowed gratings (2 deg field) presented at eccentricities 0, 2.5, and 7 deg, with spatial frequencies from 0.25 to 64 cpd. The visual scenes were converted into retinal images using eccentricity- and wavelength-dependent optical point spread functions, derived from published wavefront-aberration measurements. Then, we computed the cone excitations of a simulated retinal mosaic and passed these through a model of mRGC receptive fields (RFs). The spatial linear model RFs are constrained by published anatomical and physiological data, and are non-selective with respect to input from the L- and M-cones. Finally, we derived CS by training a classifier with stimulus-labeled mRGC responses and computing the accuracy of the trained classifier using a simulated psychophysical task. The derived CSs recapitulate core features of human performance: i) CS as a function of spatial frequency is bandpass for achromatic gratings but lowpass for chromatic gratings; ii) peak CS shifts to lower spatial frequencies with increasing eccentricity; iii) at low spatial-frequencies, chromatic CS declines more rapidly with eccentricity than achromatic CS. Our results confirm the value of investigating limits of visual performance imposed by known features of the initial visual encoding, and provide a framework for integrating diverse experimental data into a generalizable model of CS.

36.351 AN INTERNAL REPRESENTATION OF CONTRAST BASED ON MAGNITUDE ESTIMATION COMPATIBLE WITH DISCRIMINATION

Cristina Rodríguez Arribas[1] (cristinarodriguezarribas@gmail.com), Daniel Linares[1], Joan López-Moliner[1]; [1]Vision and Control of Action Group, Department of Cognition, Development, and Psychology of Education, Institute of Neurosciences, Universitat de Barcelona, Barcelona, Catalonia, Spain

The relationship between physical and perceived magnitude is a longstanding question that remains unsolved. This relationship could be characterized by an internal representation that could be modeled by a function that maps physical intensity into perceived intensity (transducer) and some perceptual variability (noise). This internal representation has been constrained by studies using discrimination tasks, but the challenge is that for a given pattern of discriminability, there are infinite combinations of transducers with noise that are compatible. Recently, Zhou, Duong and Simoncelli (2022) have proposed that magnitude estimation tasks could, not only constraint, but fully specify the internal representation being the transducer the mean estimation response and the noise the standard deviation. Under this proposal, delta (δ), the slope of the transducer at each intensity divided by the noise, should be consistent with the discriminability pattern obtained from discrimination tasks. To test this framework, five participants conducted a discrimination and a magnitude estimation task for contrast. We used contrast because discrimination experiments have shown a nonlinearity at low intensities—the pedestal effect—which should be revealed by δ obtained from magnitude estimation. Our results for the discrimination task replicate the pedestal effect. For magnitude estimation, we found a linear or slightly compressive behavior for the transducer and an expansive behavior for the noise. Notably, δ estimated from magnitude estimation showed a pedestal effect, consistent with the discrimination task. Overall, our findings suggest that magnitude estimation is a valid procedure to characterize the internal responses.

This work was funded by grants PID2020-114713GB-I00 funded by MCIN/AEI/10.13039/501100011033 to JLM and PID2020-119843RB-I00 by MCIN/AEI/ 10.13039/501100011033 to DL. CRA was supported by grant FPI PRE2021-099277 from the Spanish Ministry of Science and Innovation.

36.352 SPATIALLY-SPECIFIC FEATURE TUNING DRIVES RESPONSE PROPERTIES OF MACAQUE IT CORTEX

Akshay V Jagadeesh[1], Sohrab Najafian[1], Kendrick N Kay[2], Michael J Arcaro[3], Margaret Livingstone[1]; [1]Harvard Medical School, [2]University of Minnesota, [3]University of Pennsylvania

Neurons in primate visual cortex respond to stimuli within local regions of visual space, i.e. receptive fields. Traditionally, receptive fields are mapped by characterizing neuronal or population activity in response to low-level, high-contrast stimuli, such as bars or gratings, presented at various locations. This approach has proven effective for mapping receptive fields in early visual cortex; however, it is notably more difficult to map receptive fields in high-level regions of visual cortex, such as inferior temporal (IT) cortex, due to the complex stimulus selectivity exhibited by neurons in these regions. Here, to determine whether there is an interaction between the spatial and featural selectivity of IT neurons, we measured multiunit electrophysiological responses in fMRI-defined face-selective patches of IT cortex of rhesus macaques. Monkeys fixated while viewing a diverse set of naturalistic images, including faces, face parts, hands, objects, and scenes, in a grid of locations, spaced apart by 1-2 visual degrees, spanning the central 17 degrees around fixation. We modeled IT receptive field structure as a two-dimensional Gaussian with compressive spatial summation (Dumoulin & Wandell, 2008; Kay et al., 2013) and were able to estimate reliable receptive field positions for face-selective units. We observed apparent shifts in receptive field position depending on stimulus type: specifically, a vertical shift for inverted compared to upright faces, consistent with prior human fMRI results (Poltoratski et al., 2021). However, this apparent shift can be explained by the spatially non-homogeneous contribution of face parts/features to the model responses, which we determined by modeling receptive field position as a function of eye or mouth position. These preliminary findings, that receptive fields of face cells in IT exhibit spatially specific feature tuning for face parts, indicate the need to jointly consider both spatial and featural selectivity when mapping receptive fields in high-level visual cortex.

SUNDAY, MAY 19, 2:45 – 6:45 PM, BANYAN BREEZEWAY

Visual Search: Cueing, context, scene complexity, semantics

36.353 EXPLORING THE LIMITS OF RELATIONAL GUIDANCE USING CATEGORICAL AND NON-CATEGORICAL TEXT CUES

Steven Ford[1] (steven.ford@ucf.edu), Younha Collins[1], Daniel Go[1], Joseph Schmidt[1]; [1]University of Central Florida

Objects in the environment do not exist in isolation; they exist relative to other objects (your wallet may be to the right of your keys). Recent work suggests that following a pictorial target preview, spatial relationships between objects do guide search, as measured by the proportion of trials in which the target pair is fixated first (Ford et al., 2021; Ford et al., In revision). To parameterize this finding, we conducted three experiments to assess the oculomotor guidance of attention generated by spatial relationships in response to text cues.

In all three experiments, participants searched for arbitrary object pairs in particular spatial arrangements (e.g., "fish above car"), amongst other pairs of random objects and we assessed performance between matched (target pairs matched the cued spatial relationship) and swapped (target pairs relationship was reversed) search displays. Experiment one investigated relational guidance using categorical text cues, with one or both objects cued. The second also used categorical text cues, but two objects were always cued, and the search array contained both, one, or neither of the cued objects in matched or swapped arrangements. Relational guidance did not emerge in either experiment, suggesting that relational guidance might rely on highly specific visual features. To test this, we conducted a final experiment in which participants memorized a limited set of targets so that they could verbally describe each object's specific visual features. They were then given text cues pertaining to the specific targets they memorized. In this case, relational information impacted oculomotor search guidance. The findings suggest that relational guidance can be extended beyond pictorial previews, but depends on well-learned visual features that can be precisely coded. Variance in visual features that result from a category of objects may eliminate relational guidance.

36.354 EXOGENOUS CUES MAKE SEARCH LESS EFFORTFUL

Sangji Lee[1] (lee73769598@tamu.edu), Andrew Clement[2], Brian Anderson[1]; [1]Texas A&M University, [2]Millsaps College

In visual search, a cue predicting the location of an upcoming target facilitates the speed of target report, with the benefit of the cue being proportional to its predictive validity. Although behavioral data supports the facilitation of search performance by a predictive cue, it is unclear whether validly cuing the target location can reduce the mental effort involved in searching for the target. In the current experiment, participants searched for a target among distractors and identified whether the target had a gap on the left or the right. Each trial was preceded by an exogenous cue that could appear at the location of one of the items. Participants completed three types of blocks where the cue validity was initially set to 25%, 45%, or 65%. After the initial validity was indicated, participants were given the option to use a hand dynamometer to exert physical effort to increase the cue validity up to 85%. The duration of each trial was fixed at three seconds using a dynamic inter-trial-interval, and participants were explicitly informed that there was nothing they could do to shorten the duration of the experiment. We observed a robust validity effect for the exogenous cue in both response time and accuracy, indicating that participants used the cue to help localize the target. Participants were also strongly motivated to increase cue validity, exerting near-ceiling levels of force regardless of block type. That is, regardless of how high the validity of the cue was initially set to, participants were willing to exert their maximal physical effort to bring validity up to 85%. Our findings suggest that validly cuing the target location reduces the mental effort of searching, to the degree that participants are willing to trade off some measure of physical effort in order to achieve this reduction in mental effort.

36.355 CONTEXTUAL CUEING IS NOT RESTRICTED TO A LOCAL CONTEXT, WHEN THE LOCAL CONTEXT CANNOT BE EASILY SEGREGATED

Aner Zheng[1] (clorisyeon@gmail.com), Sang Chul Chong[1]; [1]Yonsei University

Contextual cueing is a phenomenon where visual search becomes faster in repeated contexts. Previous studies have suggested that contextual cueing is based on only stimuli near the target (Brady & Chun, 2007; Olson & Chun, 2002). These experiments had a spatial constraint, where only three stimuli were presented within a quadrant. We hypothesized that this spatial constraint restricted the scope of context learning. We compared contextual cueing effects when the spatial constraint was either present or absent. The task was to find a T-shaped target among 11 L-shaped distractors. We manipulated the structure of a display (constrained vs. non-constrained), scope (global vs. local), and repetition (repeated vs. non-repeated). The constrained condition was a between-subject, while the other conditions were within-subject variables. In the constrained condition, three stimuli were presented in each quadrant, whereas all 12 stimuli were presented randomly in the non-constrained condition. In the global-repeated condition, all stimuli were presented at the same location in every block. In the local-repeated condition, only the two distractors and the target were repeated. These three locally repeated stimuli were confined to one quadrant in the constrained condition, whereas they were the target and its two closest distractors in the non-constrained condition. In the non-repeated condition, all stimuli were randomly generated and the potential locations of a target were the same as in the repeated condition. In the constrained condition, we found the contextual cueing effects in both the global and local conditions. However, in the non-constrained condition, we did not find the effect in the local condition. Therefore, the perception of a spatial structure, such as a quadrant, may serve as a cue to divide a display into four distinct areas. This division could result in the limitation of learning to only the distractors presented in the same quadrant as the target.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A2C3004133).

36.356 CONTEXTUAL CUEING IN HIGHLY COMPLEX REAL-WORLD STIMULI

Tom Tomshe[1] (dtomsh1@lsu.edu), Melissa R. Beck[1], Brandon Eich[1]; [1]Louisiana State University

The contextual cueing effect describes the benefit of repeated contexts on visual search performance relative to non-repeated contexts. Larger contextual cueing effects, defined by a larger benefit of repeated contexts, have been found in real-world stimuli relative to simplistic stimulus arrays. However, it is unclear how much of this difference is due to increased complexity of the display versus other factors such as semantic information. The current work examined the impact of stimuli rich in complexity but with little semantic information on facilitating contextual cueing. We assessed participants' visual search performance using highly complex aeronautical charts across five blocks in which half of the charts in each block were shown either once in the experiment (Non-Repeated) or repeated once each block (Repeated). Participants also completed a recognition post-test consisting of repeated charts from the search task, a subset of non-repeated charts from the search task, and novel charts which were not present in the search task to assess their memory for the charts. Non-linear modelling revealed no significant differences in search response time across blocks of trials between repeated and non-repeated charts. However, post-test results demonstrated strong recognition performance for repeated, followed by non-repeated charts compared to novel charts. These findings suggest that high complexity and search difficulty do not necessarily lead to a reliable contextual cueing effect. This lack of a contextual cueing effect was not due to a lack of memory for the arrays: Strong explicit memory for the charts was confirmed. Thus, the larger contextual cueing effects typically found in real-world stimuli may be due to semantic information more so than stimuli complexity. Furthermore, there may be an optimal level of complexity for contextual cueing that may have been surpassed in the current experiment.

36.357 EXAMINING DIFFERENTIAL EFFECTS OF TARGET AND CONTEXT REPETITION IN VISUAL SEARCH: INSIGHTS FROM A BIG DATA APPROACH

Emma M. Siritzky[1], Audrey Siqi-Liu[1], Dwight J. Kravitz[1], Timothy J. Vickery[2], Stephen R. Mitroff[1]; [1]The George Washington University, [2]University of Delaware

Much remains unknown about how visual statistical learning—incidental learning of stimulus and display regularities—influences future performance. Visual search paradigms may be useful for investigating visual statistical learning given that recent stimulus repetitions facilitate search performance and the effects decay over time. Specifically, past studies on the “Previous Trial Effect"—same/different content in preceding trials influences performance—have found priming boosts consecutive trial performance and decays over ~5 trials (Maljkovic & Nakayama, 1994). Additionally, contextual cueing—implicit learning of repeated contextual information—shows performance improves following the repetition of context, regardless of its task relevance (Chun & Jiang, 1998). As the extent and potential overlap of such priming effects remain unclear, the current study aimed to determine how target repetition was affected, and possibly strengthened, by general search context (e.g., background characteristics). With a massive dataset (~3.8B trials, ~15.5M participants) of visual search data collected from the Airport Scanner mobile application (Kedlin Co.), it is possible to look at subtle and large changes in target and context repetitions, and at varying delay intervals. For example, performance was evaluated for a specific trial as a factor of whether a preceding trial contained (1) the same target and context, (2) the same target but different context, (3) a different target but same context, or (4) a different target and context. Moreover, these conditions were assessed for multiple preceding trials to measure decay and other time-focus effects. Search was faster in all match conditions compared to baseline performance, with larger benefits for the matching target than context alone. However, response times were fastest following recent trials where the combination of matching target and context occurred, suggesting a possible benefit for stimulus and display associations. These and other results provide insights from visual search performance on how statistical regularities influence cognitive performance.

This research was funded by US Army Research Laboratory Cooperative Agreements #W911NF-21-2-0179, #W911NF-23-2-0210, and #W911NF-23-2-0097.

36.358 TASK COMPLEXITY AND ONSET OF VISUAL INFORMATION INFLUENCE ACTION PLANNING IN A NATURAL FORAGING TASK

Danilo A. Kuhn[1] (danilo.kuhn@uni-marburg.de), Jan Tünnermann[1], Anna Schubö[1]; [1]Philipps-University Marburg

Picking fruits or playing with LEGO® bricks are everyday examples for tasks that not only involve visual search for multiple objects, but also action planning to grasp the visually selected objects and place them somewhere. How a selection movement is planned should also depend on the complexity of a task. We first asked if differences in task complexity would elicit differences in movement onset, and second, if the availability of visual information during task instructions influences action planning. Participants engaged in a non-exhaustive natural foraging task, where they had to pick and place LEGO® bricks of a particular color in predefined areas. Placing instructions (collect, sort, pile) differed between trials to implement different precision requirements. More complex tasks should require more time for planning the grasping movement and accurate placing, which should be reflected in differences in movement onset. Further, participants wore shutter glasses to manipulate when visual information became available. On one day, participants listened to the task instructions while the shutter glasses were open. On the other day, the shutter glasses were closed while participants listened to the task instructions. Results revealed a later movement onset in the piling condition, irrespective of vision availability during instructions. Moreover, the trajectory of the first reach in the late vision condition showed a larger curvature, indicating that vision during task instructions aids action preparation for target selection. Our findings show that more complex tasks can delay action initiation even if vision is available during instruction in a natural foraging task. However, the availability of vision still influences the curvature of the planned movement, potentially reflecting uncertainty in target choice when vision was not available during task instruction.

This research was supported by “The Adaptive Mind”, funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art.

36.359 THE INTERACTION OF CLUTTER AND SCENE SIZE ON VISUAL SEARCH IN NATURAL SCENES

Wentao Si[1] (wsi@bates.edu), Lindsay Houck[1]; [1]Bates College

It is well understood that content in natural scenes impacts attention. Clutter, the level of disorder within a scene, also impacts visual search, where search time increases with scene clutter. Clutter is operationalized as feature congestion, edge density, and subband entropy, which all correlate with target response time (RT). Clutter also correlates with object quantity, though this depends on scene size - larger scenes have higher object capacities. It is less known how scene size interacts with visual clutter to impact search time. Here, we used stimuli from Park, Konkle, and Oliva (2015) with visual clutter and scene size ratings to test this interaction. Scenes were categorized by dimensions of clutter (low, medium, high) and size (small, medium, large) from levels 1, 3, and 5 from the stimuli. 49 online participants

completed 216 visual search trials, half target-present and half target-absent, with an 18-item T and L array overlaid on the scene. A linear mixed-model analysis revealed a main effect of scene clutter on RT ($F(2, 9366.1) = 116.55$, $p < 0.001$), with more clutter increasing RT, but no main effect of scene size ($F(2, 9366.1) = 2.33$, $p = 0.097$). However, the interaction between clutter and size was significant ($F(4, 9366.1) = 33.32$, $p < 0.001$): for low and medium clutter, RT increased with scene size, but for high clutter, RT decreased with scene size. A post-hoc analysis revealed no impact of the background contrast surrounding the target on RT, ruling out low-level feature contrast in more cluttered scenes as a confounding factor. This reveals a previously unknown interaction between scene size and clutter beyond the operationalized measures of clutter as visual disorder within a scene. Future work will untangle this interaction with eye-tracking and other measures of visual clutter, such as feature congestion.

Thanks to the Bates Neuroscience Program for funding data collection for this study

36.360 ROBUST TARGET-RELATED CLUTTER METRICS FOR NATURAL IMAGES

Elizabeth Zhou[1], Yelda Semizer[2], Melchi Michel[1]; [1]Rutgers University, [2]New Jersey Institute of Technology

Vision researchers have long studied the effect of clutter on visual search performance. Traditionally, researchers have characterized clutter using “target-agnostic” measures that rely only on characteristics of the image itself, irrespective of the observer’s task (e.g., the identity of the search target). Here, we are interested in clutter in natural images, and in how the effect of clutter might vary with the target of the search. Along these lines, Semizer and Michel (2022) instructed observers to search natural scenes for objects belonging to a particular category (e.g., cellphones). They then proposed two novel metrics to characterize image clutter in a target-relevant way, based on the similarity between the background scene features and those of the target. The “exemplar-level” metric only considers the features of the target exemplar present in a particular search image. This metric is only available when a target is present in the image. In contrast, the “category-level” metric considers the feature distribution across all exemplars of the target category within the image set. This latter metric is available regardless of whether a target is present in the image. Semizer and Michel (2022) defined both metrics using Steerable Pyramid features (Simoncelli & Freeman, 1995). Here, we redefined these metrics using a more flexible feature description, Histograms of Oriented Gradients (Dalal & Triggs, 2005), that is robust to perturbations in scale and position. We then fitted generalized linear models to the exemplar-level and category-level metrics, along with a traditional (target-agnostic) clutter metric to see how well they predict search time. Our results show that, for images with a target present, the target-agnostic and exemplar-level metrics predict search time better than does the category-level metric. Moreover, when the search target is absent, the category-level metric contributes significantly (i.e., beyond the predictions of the target-agnostic metric alone) to explaining search time.

36.361 ASSESSING THE EFFECT OF STIMULI COMPLEXITY IN WEB-BASED VISUAL FORAGING

Enilda Velazquez[1], Nelson Roque[2]; [1]University of Central Florida, [2]The Pennsylvania State University

The visual foraging task investigates observer visual search behavior in respect of multiple targets. Across previous studies, abstract stimuli are commonly used to explore foraging behaviors. To fully explore visual search as it occurs in real-world environments, it was of interest to substitute abstract stimuli for ecologically valid stimuli in a foraging task. The present study investigated the effects of stimuli complexity on observer foraging performance using ecologically valid foraging stimuli. Participants ($N = 221$) completed a 14-block web-based foraging task on their own devices. The stimuli were categorized into control sets (simple shapes with 1 feature layer, and conjunctive shapes with 2 feature layers) and experimental sets (stop signs and yield signs with 1 to 3 layers of features). Each block contained pages with 6 targets and 6 distractors. Participants viewed a target example commencing an exhaustive search. They were instructed to complete as may pages as possible per block (45 seconds). Foraging performance was assessed using mean inter-target time (ITT), where higher ITT suggested less efficient foraging performance. Linear multi-level models showed increased ITT with stimuli complexity, except in the three-layer stop sign condition. Observers using smaller screens foraged faster than those on larger screens. The variation of ITT between blocks for each stimuli type was minimal, suggesting that individuals maintained consistent foraging speeds within each stimulus category. However, greater variation in ITT within blocks was observed for the experimental stimuli, suggesting more variability in foraging performance in the experimental conditions. These findings are consistent with the theory that complex stimuli prolong search times and contribute to research on the "pop-out" search effect. Further, this research sheds light on the impacts of device differences on task performance. Overall, this study emphasizes the importance of considering the ecological context in understanding observer foraging performance.

36.362 VOWELWORLD 2.0: USING ARTIFICIAL SCENES TO STUDY SEMANTIC AND SYNTACTIC SCENE GUIDANCE

Yuri Markov[1] (yuamarkov@gmail.com), Melissa Le-Hoa Vo[1], Jeremy M Wolfe[2]; [1]Goethe University Frankfurt, Scene Grammar Lab, Germany, [2]Brigham and Women’s Hospital, Harvard Medical School

Scene guidance is difficult to investigate in realistic scenes because it is hard to systematically control complex, realistic images. Parameters like set size are often ambiguous in real or even VR scenes. We created a new version of VowelWorld 2.0 (Vo & Wolfe, 2013), where we control various parameters of a highly artificial “scene”. Scenes are 20x20 grids of colored cells with 120 cells containing letters. Participants search for a vowel, present on 67% of trials. Each scene contained three big disks (2x2 cells) with consonants on them. These served as “anchor objects” which are known to predict target locations in real-world searches (Vo, 2021). An additional 96 cells featured rings which were grouped into larger analogs of surfaces. A vowel’s placement could follow three rules. Color rule (semantic): certain targets were associated with one background color “gist” (e.g., A’s appear in red scenes). Structure rule (syntactic): vowels were placed near or inside the small rings. Anchor rule (syntactic): vowels were close to a big circle containing a neighboring consonant (e.g., “B” implies “A”). Two vowels followed all three rules, two vowels followed color and surface rules, and one vowel was placed randomly. On half of the trials, participants were precued with a specific vowel. Otherwise, participants searched for any vowel. For the first three blocks, participants attempted to learn the rules from experience.

Then, we explained the rules. Participants failed to fully learn rules but did benefit from the learned anchor rule (shorter RTs). Knowing rules markedly speeded performance for vowels that followed only color and surface rules. Anchor rule vowels showed less improvement over initial learning. Knowing rules had a major impact on ending absent trials. Future work will systematically vary the predictability of different rules to test under which circumstances rule learning becomes more or less optimal.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 222641018 SFB/TRR 135 TP C7 granted to MLHV and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project 'The Adaptive Mind').

Sunday, May 19, 2:45 – 6:45 pm, Banyan Breezeway

Visual Search: Eye movements, suppression

36.363 CAN MULTIPLE EQUALLY SALIENT DISTRACTORS BE SUPPRESSED SIMULTANEOUSLY?

Brandi Lee Drisdelle[1,2] (brandi.drisdelle@bristol.ac.uk), Martin Eimer[2]; [1]University of Bristol, [2]Birkbeck, University of London

While unique salient items tend to "pop out" from a visual scene, capturing our attention, there is a growing body of research suggesting that task-irrelevant salient items can instead be suppressed. Previous research has primarily focused on the suppression of one salient item. Using behavioural and electrophysiological (PD component) measures, we recently showed that, for displays with two salient items, suppression was stronger for the more salient item (Drisdelle & Eimer, 2023). Here, we compared displays with one salient item and two salient items using behavioural (Experiment 1) and electrophysiological (Experiment 2) measures. In both tasks, participants searched for a shape-defined target among target-coloured distractors and zero (Experiment 1 only), one, or two different-coloured (salient) distractors. Salient distractors were always the same colour. Experiment 1 employed a capture-probe paradigm, where participants reported probe letters superimposed on all shapes. Critically, suppression effects (an impaired ability to report letters at a salient distractor location) were equally strong regardless of whether another salient distractor was present, indicating efficient multiple-item suppression. Experiment 2 provided converging electrophysiological evidence, based on PD components (ERP markers of suppression) elicited by displays with one or two lateral salient distractors. Lateralised distractor related activity was isolated from target related activity by only examining displays where the target was absent or on the midline. Critically, the PD elicited by displays with two lateral salient distractors was significantly larger than the PD elicited by displays with a single lateral salient distractor. This indicated that PD components can reflect the combined contribution of suppression that is applied simultaneously to multiple salient distractors. Taken together, our results suggest that it is possible to suppress multiple distractor items that are equally salient, and that this suppression does not result in any behavioural costs relative to single-item suppression.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Fonds de recherche du Québec – Nature et technologies (FRQNT) postdoctoral fellowship programs, and grant ES/V002708/1 from the Economic and Social Research Council (ESRC), UK.

36.364 CAN PEOPLE SUPPRESS SALIENT VISUAL DISTRACTORS WITHOUT FOREKNOWLEDGE OF THEIR COLORS?

John McDonald[1] (jmcd@sfu.ca), Daniel Tay[1], Jessica Green[2], Ali Jannati[3]; [1]Simon Fraser University, Burnaby, Canada, [2]University of South Carolina, [3]Linus Health, Boston, USA

Several lines of evidence suggest that observers can suppress salient visual distractors to prevent them from capturing attention. Currently, it is unclear whether such proactive suppression is possible when the defining feature of the distractor varies unpredictably across trials. Using probe-recall rates and oculomotor data, Gaspelin and Luck (2018, JEPHPP) showed that suppression is not possible when target and distractor features swap unpredictably across trials. These results indicate that suppression may be tied to the visual feature that defines the distractor (feature suppression). However, we previously reported that salient distractors elicit an event-related potential (ERP) component associated with suppression (the distractor positivity, PD) even when a salient distractor is randomly intermixed with a less-salient distractor that requires no suppression (Gaspar et al., 2016, PNAS). Here, we tested the feature-suppression hypothesis more directly by varying the color of the distractor while maintaining its high salience. In one experiment, participants viewed displays containing eight or nine green circles, a green diamond (shape-singleton target), and, on distractor-present trials (50%), a nongreen circle (color-singleton distractor). Critically, the distractor color was varied randomly across trials (magenta, red, orange, blue, and cyan) to prevent feature-based suppression. Both target and distractor elicited contralateral positivities over the posterior scalp in the time range of the P1/N1 components (Ppc; 100—200 ms). The target elicited a subsequent N2pc (index of attentional selection), while the distractor elicited a subsequent PD. These findings indicate that salience-based suppression can occur without foreknowledge of the distractor's color and are thus inconsistent with the feature-suppression hypothesis. The Ppc results are also inconsistent with suppressive interpretations of the pre-N2pc positivity (i.e., it does not appear to be an early PD).

36.365 DISTRACTOR SUPPRESSION IN PRIMARY VISUAL CORTEX

David Richter[1,2] (david.richter.work@gmail.com), Dirk van Moorselaar[1,2], Jan Theeuwes[1,2,3]; [1]VU Amsterdam, [2]Institute Brain and Behavior Amsterdam (iBBA), [3]William James Center for Research, ISPA-Instituto Universitario

Selective attention is fundamental for effective interaction with our surroundings. A primary mechanism behind this capability is the automatic suppression of the numerous distracting stimuli that compete for our attention. This suppression allows us to focus on essential tasks, like driving a car, without being overwhelmed by salient but irrelevant inputs. While the behavioral benefits of distractor suppression are well-established, its neural underpinnings are not yet fully understood. In an fMRI study, we examined where and how sensory responses in the visual brain display signs of distractor suppression after incidental learning of spatial statistical regularities. Participants were exposed to an additional singleton task where, unbeknownst to them, one location more often contained a highly salient distractor. We then analyzed whether visual responses in terms of fMRI BOLD were modulated by this predictability. Our findings indicate that such implicit spatial priors shape sensory processing even at the earliest stages of cortical visual processing in V1, evident as a suppression of stimuli at locations which frequently contained the distracting information. Notably, this suppression occurred when distractor or target stimuli appeared at the high probability distractor location, suggesting that suppression arises before stimulus identification completed. These results highlight a proactive strategy employed by early visual cortex, where potential distractions are suppressed preemptively, possibly even before the onset of the stimulus display. In sum, our study underscores how the brain leverages prior knowledge, for example from statistical learning, to optimize sensory processing and attention allocation.

36.366 REVISITING THE TIMING OF SALIENT-SIGNAL SUPPRESSION

Daniel Tay[1] (dta22@sfu.ca), John McDonald[1]; [1]Simon Fraser University, Burnaby, Canada

According to the signal suppression hypothesis, salient stimuli automatically trigger “attend-to-me” signals that cause these stimuli to capture attention unless they are suppressed. In support of this hypothesis, Sawaki and Luck (2010, Atten Percept Psychophys) showed that a colour-singleton distractor elicited an event-related potential (ERP) marker of suppression called the distractor positivity (PD) rather than a marker of attentional selection (N2pc) in a letter-detection task. The PD was reported to begin 100 ms after stimulus onset, which was hypothesized to be early enough to prevent the distractor from eliciting the N2pc in its conventional time range (180–300 ms). Here, we tested whether this “early PD” reflected suppression or the attend-to-me signal itself. The study was motivated by two considerations. First, it may take more than 100 ms to establish a salience-based attend-to-me signal from cortex. Even if such a fast signal were possible, no ERP activity resembling such signal was observed prior to onset of the PD. Second, salient targets sometimes elicit a positivity akin to an early PD prior to the onset of N2pc. This positivity cannot be ascribed to suppression because the subsequent N2pc demonstrates that the item was attended. We replicated Sawaki and Luck’s task and then reversed the target and distractor stimuli such that participants detected the colour singleton rather than a nonsingleton letter. The singleton target elicited a small early positivity that rapidly turned into an early N2pc due to the salience of the target. The early positivity was larger for upper-visual-field targets because the temporally overlapping N2pc is smaller for upper-field stimuli. The results indicate that the early posterior contralateral positivity (Ppc) is associated with the attend-to-me signal and the later PD is associated with suppression of that signal.

36.367 EFFECTS OF SEARCH PRIORITY ON WORKING MEMORY-GUIDED SEARCH FOR REAL OBJECTS: EVIDENCE FROM EYE-MOVEMENTS

Hanane Ramzaoui[1] (hramzaoui@lsu.edu), Melissa R. Beck[2]; [1]Louisiana State University, [2]Louisiana State University

Working memory (WM) serves current goals while enabling future ones to be planned and maintained. This latter role of WM is crucial, as everyday activities require finding task-relevant objects in a timely manner (visual search) and therefore prioritizing WM representations with respect to moment-by-moment task-relevance of items held in WM. How prioritization in WM affects the different processes during visual search remains unanswered. Here, we compare current and prospective search templates in a sequential search task. The current search template is the first target to be searched for, and the prospective search template is the target to be searched for in a subsequent display. On each trial, the participants perform two consecutive searches. Before the search displays, both targets are presented, and the one to be searched for in subsequent displays is indicated by the nature of its outline. Each array includes six colored objects, each unique in color and semantic category, arranged in a circle. Using an eye-tracker, response times were segmented into distinct processes during visual search: initiation time (first saccade latency), scanning time (elapsed time between the first saccade and the first fixation on the target) and verification time (elapsed time between fixation on the target and manual response). Search priority affected all three processes, resulting in shorter times for the first search than for the second search. This finding suggests that search priority affects both the process of setting up the template and the process of matching a fixated object to an internal representation of the target. Specifically, the information currently relevant provides a quickly accessible and precise template, enabling search to begin as soon as the array appears, as well as faster comparison of the fixated object to the template.

36.368 EVALUATING THE CONTRIBUTIONS OF TOP-DOWN AND BOTTOM-UP PROCESSING ON EYE MOVEMENTS DURING PARALLEL VISUAL SEARCH

Howard Jia He Tan[1] (jiaheht2@illinois.edu), Zoe (Jing) Xu[1], Alejandro Lleras[1], Simona Buetti[1]; [1]University of Illinois Urbana-Champaign

Most models of visual attention refer to the interplay between top-down and bottom-up processing, emphasizing factors like attentional templates and salience. Prior modeling work in the field has often used complex, real-world scenes, that do not generalize well to the more simplified and controlled stimuli used in most lab studies on visual search. In the current study, we used an efficient visual search paradigm in a pseudo-realistic environment, with well-controlled search stimuli that allows a simultaneous evaluation of the impact of top-down and bottom-up factors on eye movement patterns. Our stimuli varied along the color dimension to manipulate target-distractor similarity. In addition, our displays contained a salient stimulus of higher salience than target and other distractor stimuli. Eye gaze was tracked using an Eyelink 1000 Plus and fixations were categorized by the item they landed closest to. We manipulated task instructions, introducing a free-viewing instruction condition to serve as a baseline for how bottom-up contrast guided eye movements in one group of participants, and a top-down search instruction in a second group, where subjects were asked to find the red target in the scene. Experiment 1 assessed the impact of set size of less-salient distractors across both instructions. Experiment 2 examined target-distractor similarity effects for the less-salient distractors. We computed the ratio of fixations selective towards the target (top-down) versus the high-salience singleton (bottom-up) across all fixations made and how selectivity evolves throughout a trial across all conditions tested.

Results across free-viewing conditions showed selectivity for the high-salience item during the first fixation was 14% higher than for the stimulus used as target stimulus in the search task, yet it was 30% lower in the goal-driven group. Selectivity towards bottom-up factors decreased throughout a trial in free-viewing conditions across both experiments and decreased as target-distractor similarity increased across both instructions.

This material is based upon work supported by the National Science Foundation under Grant No BCS1921735 to SB.

36.369 VIEWPOINT SELECTION IN ACTIVE VISUAL SEARCH

Tiffany Wu[1] (tiffwu@yorku.ca), John K. Tsotsos[1]; [1]York University

Most visual search studies use a 2D, passive observation task, where subjects search through artificial stimuli on a screen. In contrast, real world search involves physical 3D scenes and searchers who choose relevant scene viewpoints as the search proceeds. Searchers employ eye, head, and body movements to investigate the scene; they are active observers. To investigate viewpoint selection in active observation during 3D search, an active search task was conducted in a controlled real-world environment, a 3x4m space furnished with tables and wire cage shelving acting as surfaces to place the stimuli. Stimuli were miniature everyday objects, scattered in various orientations on the tables and cages. Targets were placed in upright, sideways, face-up, or diagonally tilted positions, but the target image probe was always presented in an upright orientation. Observers moved freely, untethered, and their eye and head movements, reaction time, and accuracy, were synchronized and measured over 12 trials each. Results indicate that similar to 2D search tasks, target-absent trials take longer than present trials and require more fixations and head travel. Interestingly, efficiency in these metrics was found to increase over time only in target present trials, not in absent trials. Collected eye and head movement data further revealed head tilts for subjects to match canonical orientations of non-upright objects. Indeed, targets placed in non-canonical orientations required more fixations before subjects would confirm them as present. Subjects were also found to crouch in order to fixate on objects placed at lower levels (such as the table surface, which was approximately 70cm high). Our results provide novel analyses on eye and head movement metrics during search in an active observation environment, demonstrating the important nuances of movements that can be induced by requiring viewpoint selection to complete a task.

This research was supported by the Canada Research Chairs Program (950-231659), and the Natural Sciences and Engineering Research Council of Canada (RGPIN-2022-04606).

SUNDAY AFTERNOON POSTERS IN PAVILION

SUNDAY, MAY 19, 2:45 – 6:45 PM, PAVILION

Face and Body Perception: Neural mechanisms 1

36.401 MULTIVARIATE ANALYSIS OF STRUCTURE-FUNCTION-BEHAVIOR RELATIONS SUPPORTING FACE RECOGNITION BEHAVIOR IN AUTISTIC AND NON-AUTISTIC ADOLESCENTS

Yiming Qian[1] (yimingqian2015@gmail.com), K. Suzanne Scherf; [1]Department of Psychology, the Pennsylvania State University

Most neuroimaging studies assessing brain-behavior relations tend to evaluate pairwise associations between a single neural metric (e.g., fMRI activation or microstructural properties) and a behavioral metric (e.g., accuracy). However, this approach does not characterize how brain structure and function interact holistically to subserve behavior. To address this gap, we integrated fMRI, sMRI, DTI, and functional connectivity measures to investigate how relational patterns among these metrics predict face recognition (FR) behavior over time. The data presented here represent the baseline comparison between autistic (N = 29) and non-autistic (N = 22) adolescents (11-17 years). We chose to study the FR system due to the prominent difficulty in face recognition behavior that exists in autism and the whole-brain nature of the FR neural system. For the fMRI and sMRI metrics, we individually defined six ROIs—the bilateral fusiform face area (FFA) and amygdala as the core regions in the FR system, and bilateral early visual cortex (EVC). Structural connectivity was defined as the radial diffusivity of the white matter tracts seeding from each ROI. Functional connectivity was defined across all ROIs using GIMME and quantified using overall node strength. Partial Least Squares, a multivariate method, was then utilized to understand how the relation among these metrics predicted FR behavior. The results showed robust brain-behavior associations. Specifically, for both groups, higher functional activation within regions (right FFA and left amygdala), functional connectivity (left EVC), and radial diffusivity in most ROIs within the network together positively predicted higher FR scores (i.e., CFMT). The relation between the brain metrics and FR behavior was stronger in the autistic than the non-autistic group. This work may inspire new ways of thinking about how neural networks dynamically organize through development to support behavioral change and, more specifically, why face recognition is a challenging social behavior for autistic individuals.

The research was supported by Pennsylvania Department of Health SAP grant 4100047862 (M.B., K.S.S., N.M.), NICHD/NIDCD P01/U19 HD35469-07 (M.B., PI Nancy Minshew).

36.402 N170 AND N250 SENSITIVITY TO DIAGNOSTIC FACIAL INFORMATION DURING WHOLE-FACE RECOGNITION

Pierre-Louis Audette[1] (audp02@uqo.ca), Justin Duncan[1], Caroline Blais[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais

In our presentation at VSS2023, we explored the correlation between N170 face sensitivity and diagnostic information processing during a 10-identities recognition task (Audette et al., 2023). We utilized sparse facial stimuli generated with Bubbles, and previously published classification images (Royer et al., 2018) were employed to quantify the available diagnostic information on a stimulus basis (0-100%). Remarkably, we observed a linear increase in N170 amplitude as diagnostic information increased (r=-0.98), indicating thorough processing of diagnostic information during face identification. However, the revealed amount of diagnostic information strongly correlated with the amount of facial surface (r=0.99), suggesting that this effect may be influenced not only by diagnostic information but also by the overall amount of facial information. Consequently, the second phase systematically examined whether N170 and N250 reflect the quantity of diagnostic information processed by the brain while keeping the amount of facial information constant. Sparse facial stimuli were generated using a method similar to the first phase. To standardize the facial surface across stimuli, we applied inverse bubbles to an average face of the 10 identities, replacing face regions hidden by bubbles with non-diagnostic facial information. The stimuli were categorized into 12 bins ranging from 0.001% to 100%, with an additional 0% condition (average face). EEG data were collected from 10 participants during 1,300 trials of a 10-identities recognition task. We analyzed N170 and N250 peak amplitudes at PO8, along with behavioral responses. Results indicated that only N250 peak amplitude varied: as diagnostic information increased, N250 amplitude linearly increased (r=-0.90). Moreover, N250 amplitudes across bins strongly correlated with recognition accuracies (r=-0.86). Thus, N250 seems to reflect in-depth processing of diagnostic information during face identification, while N170 responds to the ease of categorizing the stimulus as a face.

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2022-04350).

36.403 NEURAL CORRELATES OF FAMILIAR FACE RECOGNITION DO NOT BENEFIT FROM COLOUR INFORMATION

Holger Wiese[1] (holger.wiese@durham.ac.uk), Emma Smedley[1], Linda H. Lidborg[1]; [1]Durham University

Humans are highly efficient at recognizing familiar faces. This remarkable ability is based on image-invariant long-term representations of known faces which can be activated from highly variable instances, including pictures that have never been seen before. While the existence of such representations is well-established, it is less clear what specific information is stored in them. Here, we examined whether colour information is beneficial for activating familiar face representations, which would suggest that they contain such information. We used event-related brain potentials (ERPs) in an immediate repetition priming paradigm. Images of well-known celebrities were combined into prime/target pairs, consisting of either different images of the same facial identity (repetition condition) or two different identities (non-repetition condition). Critically, while the target face was always presented in colour, the prime could be either in colour or greyscale. We observed clear ERP priming effects, with more negative amplitudes at occipito-temporal channels in the repetition relative to the non-repetition conditions starting approximately 220ms after target onset. This N250r effect was highly similar for colour and greyscale primes. These findings show that face representations are efficiently activated by both colour and greyscale images, and accordingly that colour information is not beneficial for facial identity processing. This is in line with the suggestion that familiar face representations do not contain colour information.

LHL and HW are funded by the Economic and Social Research Council (ESRC, UK)

36.404 RACE SHAPES RAPID NEURAL FACE CATEGORIZATION

Pauline Schaller[1] (pauline.schaller@unifr.ch), Lisa Stacchi[1], Peter de Lissa[1], Anne-Raphaëlle Richoz[1], Roberto Caldara[1]; [1]University of Fribourg

Face processing in the human brain is relatively automatic, rapid, effortless, and very effective. However, this critical biological skill is markedly impaired when observers are confronted with less familiar other-race or inverted faces. Interestingly, on the one hand, when more sensitive neural adaptation paradigms are used, same-race upright faces elicit early larger N170 amplitudes compared to other-race faces. On the other hand, inverted same-race faces lead to a greater recognition impairment and elicit larger N170 amplitudes compared to inverted other-race faces. Yet, whether similar neurofunctional electrophysiological signatures for those impairments could also be found with fast periodic visual stimulation (FPVS) remains unknown. The FPVS is a fast, reliable, and highly sensitive method to isolate neural brain categorization responses. We thus recorded the electrophysiological signals of Western observers while they were viewing upright and inverted Western and Eastern faces by using an FPVS paradigm. The first experiment consisted in a stream of 6 Hz composed of either same- or other-race faces among which the faces of the alternate race were periodically embedded (oddball). In the second experiment, same- or other-race oddball faces were respectively embedded in object streams either in upright or inverted condition. Surprisingly, the race of the faces did not modulate the FPVS response in the upright conditions, regardless of whether faces were embedded in a stream of faces or objects. Critically, however, in the inverted condition, the neural response was significantly lower for same- than other-race faces, with the response to other-race faces remaining comparable to the upright condition. Our data show that face inversion affects the FPVS neural responses of the visual category human observers are more expert with. These findings contribute to a deeper understanding of the intricate mechanisms underlying neural processing in face perception and its interaction with race.

This work was supported with funding from the Swiss National Science Foundation awarded to Roberto Caldara (100019_189018).

36.405 SIMULTANEOUS ACTIVATION OF MULTIPLE FACE REPRESENTATIONS

Linda H. Lidborg[1] (linda.h.lidborg@durham.ac.uk), A. Mike Burton[2], Holger Wiese[1]; [1]Durham University, [2]University of York

Humans are highly skilled at familiar face recognition, as we are able to recognize and distinguish between thousands of individual facial identities. Given this remarkable number of known faces, it seems surprising that we confuse familiar faces only very rarely. One potential explanation is that the activation of a memory representation of a particular known face results in the simultaneous inhibition of all other face representations. If correct, this would allow the activation of only a single facial representation at any given time, so preventing the brain from mixing up different facial identities. Here, we used event-related potentials to examine the neurophysiological basis of this hypothesized inhibitory process, and tested whether it is possible to have multiple face representations activated at the same time. Using individualized stimuli sets of highly familiar celebrity faces in an immediate repetition priming paradigm, we presented participants (N = 30) with prime displays consisting of (i) two images of the subsequent target facial identity (Double Repetition), (ii) two images of different facial identities, one of which showed the target identity (Single Repetition), or (iii) two images of two different facial identities (Non-Repetition), neither of which showed the following target identity. We observed clear N250r effects, consisting of more negative amplitudes at occipito-temporal channels for both the Double and Single Repetition conditions relative to the Non-Repetition condition. Moreover, Double and Single Repetition conditions were highly similar and did not significantly differ from each other. At variance with the inhibition hypothesis, this finding suggests that the brain can keep at least two different facial identities activated simultaneously.

This work was funded by the Economic and Social Research Council (ESRC, UK).

36.406 THE NEURAL BASIS OF HUMAN UNFAMILIAR FACE IDENTITY RECOGNITION WITH FMRI FREQUENCY-TAGGING

Marie-Alphée Laurent[1] (marie-alphee.laurent@univ-lorraine.fr), Pauline Audurier[2], Jean-Baptiste Durand[2], Jacques Jonas[1,3], Bruno Rossion[1], Benoit R. Cottereau[2]; [1]University of Lorraine, CNRS, Nancy, France, [2]CerCo, University of Toulouse 3 Paul Sabatier, CNRS, France, [3]CHRU-Nancy, Neurology Department, France

The face has great significance for social interactions and is the most diagnostic information for identifying individuals in humans. Understanding the neural mechanisms involved in face identity recognition (FIR) is critical, particularly for the individuation of unfamiliar faces, which cannot be based on encoded multimodal semantic associations, but on available visual cues only. To investigate the neural basis of FIR, we capitalize on an original fMRI approach based on fast periodic visual stimulation, providing objective, sensitive, and reliable measures of human face recognition. FMRI recordings were performed in ten healthy human subjects. Natural images of a single unfamiliar identity were presented within a rapid 6Hz stream in two conditions: (1) with the same face image across low-level changes (size, luminance, contrast) only, or (2) with different images, introducing higher-level changes (background, head orientation, expression). Every 9s during a 243s run, 7 images of different unfamiliar identities were introduced in bursts. For each participant and each condition, we recorded 3 runs with upright faces and 3 with inverted faces. Analyses were performed within face-selective regions (defined from a frequency-tagging localizer) and in the Fourier domain where individual face discrimination responses were objectively identified and quantified, at the peak of the identity change frequency (0.111Hz). Robust image-based individual face discrimination responses were found across both conditions in core face-selective ventral regions (FFA, OFA) and exhibited inversion effects, invariant to high-level stimulus changes. In contrast, responses in low-level visual regions and in the pSTS were negligible in our second condition, which involved generalization across changes of views. Interestingly, we also found specific responses to FIR in the IFG which were significantly reduced for inverted faces. Overall, our results highlight the cortical network involved in human FIR and suggest that fMRI frequency-tagging provides a valid approach to characterize the cortical network underlying this function.

36.407 UNRAVELING THE NEURAL CODE FOR REAL LIFE FACIAL EXPRESSION PERCEPTION

Arish Alreja[1], Michael Ward[2], Taylor Abel[3], Mark Richardson[4], Louis-Philippe Morency[1], Avniel Ghuman[3]; [1]Carnegie Mellon University, [2]University of California, Los Angeles, [3]University of Pittsburgh, [4]Massachusetts General Hospital and Harvard University

We study face perception to understand how our brains process the identity, expressions, and facial movements of friends, family, coworkers, and others in real life. Controlled experiments have revealed many aspects of how the brain codes for faces, but little is known about how the brain codes for the natural intensity and expressions during real life interactions. We collected intracranial recordings from epilepsy patient-participants who wore eye-tracking glasses to capture everything they saw on a moment-to-moment basis during hours of natural unscripted interactions with friends, family, and experimenters. Face pose, identity, expressions, and motion were parameterized using computer vision, deep learning, face AI and state space models. Fixation locked facial features and brain activity were related using a bidirectional model which maximized the correlation between them in a jointly learned latent neuro-perceptual space. The model predicted brain and face dynamics from each other accurately (d' of approximately 1.8, 2.47, 1.02 for overall, between and within identity comparisons). Reconstructed brain activity revealed an important role for the recently proposed putative social vision pathway alongside traditional face areas in ventral temporal cortex. Probing the representational space for facial expression and motion revealed a person's resting facial expression as an important anchor point and that neural populations were more sharply tuned to changes in expression than their intensity. Lastly, the brain exhibited greater sensitivity to small changes from a person's resting face, such as a coy smile, compared to similar differences between a big and a slightly bigger smile, a potential analog of the Weber-Fechner law for facial expressions. Together, these results demonstrate that during real world interactions, instances of individual fixations on a person's face are coded with "oval" shaped tuning spaces wherein the oval pointed to the resting expression (norm) and became bigger further from that expression.

NSF (1734907) and NIH (R01MH132225, R01MH107797)

36.408 WAVELET-BASED IMAGE DECOMPOSITION AFFECTS SSVEP SIGNAL AMPLITUDE FOR FACE IDENTIFICATION

Jérémy Lamontagne[1] (lamj54@uqo.ca), Laurianne Côté[1], Justin Duncan[1], Caroline Blais[1], Émilie Deslauniers[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais

Previous studies have predominantly examined N170 sensitivity in a binary manner, focusing on the presence or absence of distinct facial features, either independently or within a facial context (e.g., Parkington & Itier, 2018). However, recent work has suggested that the N170 operates more like a continuum, with amplitude increasing as diagnostic information accumulates (Audette et al., 2023). In parallel to the study of ERPs, the method of steady state visual evoked potentials (SSVEP) has been instrumental in exploring neuronal responses to oscillating visual stimuli, shedding light on the brain's capacity to synchronize with and process visual information across various frequencies. Seeking to replicate the amplitude continuum observed in ERPs, we utilized SSVEP, incorporating wavelets into our stimuli to enhance decomposition while preserving low-level information. Presenting modified faces at five decomposition levels (0 to 20%) and three flickering frequencies (4, 5, or 6 Hz) to 11 observers, we implemented an oddball paradigm featuring identity changes (AAAAAB). Participants completed 45 trials of 53 stimulation cycles, encompassing three trials for each of the 15 conditions. Our results suggested no effect of stimulus presentation frequency ($F(10) = 0.45$, $p = 0.630$) but high responsiveness to the level of decomposition in presented faces ($F(10) = 8.65$, $p < .001$). In essence, as the diagnostic information in faces decreased, neural activity synchronization to identity diminished. In other words, the less diagnostic information was available in faces, the less the participant's neural activity synchronized to the change in identity. These findings not only advance our comprehension of cognitive processes in face recognition but also hold promise for optimizing facial feature extraction in real-world applications.

36.409 WHERE IN THE BRAIN ARE FACE FORM AND MOTION ENCODED INDEPENDENTLY?

Fabian A. Soto[1] (fasoto@fiu.edu), S. Sanaz Hosseini[1]; [1]Florida International University

A human face is a multi-dimensional structure conveying several important social signals that are critical in day-to-day interactions. Recently, it has been proposed that two parallel pathways process form (ventral pathway) and motion (dorsal pathway) in faces. However, it is still unclear in what way the information carried by these two pathways may be independent. In this study, we investigated whether the ventral and dorsal pathways represent facial form and motion independently, using formal definitions of independent neural representation that can be tested through fMRI decoding analyses. We created highly controlled computer-generated stimuli, consisting of 3 identities showing 3 different facial motions (9 total dynamic stimuli). Two motions involved the exact same face expression, while the third motion involved a different expression. Twelve participants completed 3 hours of fMRI scanning during which they completed a one-back task. Employing MVPA, we assessed whether we could decode facial form and motion from face-selective areas in the two visual pathways. Our results suggest that most areas in the two pathways encode information about both face form and motion. In trials in which both motion and expression varied across stimuli, most areas showed evidence for invariant encoding, with the exception of aIT. In trials in which only motion varied across stimuli (i.e., no expression information was available), evidence of invariance was still observed in FFA (both motion and form) and IFG (only for identity), whereas results for other areas were inconclusive. In defiance of general opinion regarding aIT forming invariant representation of facial identity, representation of facial form untwined from motion was observed in IFG but not aIT. Individual-level analyses showed considerable differences across participants and in many cases qualified conclusions from group-level analyses, but they confirmed a key role of IFG in the invariant representation of face form.

This work was supported by the National Science Foundation under grant No. 2020982 to Fabian A. Soto. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

36.410 JUDGING THE PERCEPTUAL SIMILARITY OF OWN- AND OTHER-RACE FACES

Megan Lall[1] (lallm@mcmaster.ca), Eric Mah[1], Brett Roads[2], Jim Tanaka[1]; [1]University of Victoria, [2]University College London

Most people are experts at recognizing faces, however, research has shown that we are less expert at recognizing faces from other races - the Other-Race Effect (ORE). An influential framework to account for the ORE is the face space model where faces are represented as points in a multi-dimensional similarity space with each dimension signifying a specific facial feature or attribute. According to this approach, own-race faces are perceived as more distinct and therefore, their representational points are more spatially separated in face space. In contrast, other-race faces are perceived as more similar and therefore, their points are more densely clustered in face space. To examine the face space representations of participants for own- and other-race faces, we employed PsiZ, a novel method for obtaining psychological embeddings–rich multi-dimensional representations of psychological similarity spaces that are inferred from behavioural similarity judgements. We predicted that the psychological embeddings will be more differentiated for own-race faces and more densely packed arrangements of other-race faces. For this study, we recruited 60 African, 60 Caucasian and 60 Chinese online participants who made similarity judgments to blocks of 20 African, 20 Caucsian and 20 Chinese faces. Our main results indicated that there was limited evidence to support the face space account of the ORE. Inspection of the psychological embeddings by race of the participant and face showed that own-race faces were not more differentiated in face space than other-race faces. However, upon examining faces by different racial groups, distinct patterns emerged. African participants exhibited an ORE for Caucasian faces, while Caucasian participants demonstrated an ORE for Chinese faces. Notably, Chinese participants did not display a discernible ORE, indicating variability in cross-race recognition effects among the studied groups. Participant similarity judgements were moderately correlated with simulated judgements based on VGG-16 perceptual features, with some differences by race.

Face and Body Perception: Neural mechanisms 2

36.411 AFTEREFFECTS FOLLOWING ADAPTATION TO FACE MENTAL IMAGES

Mathias Salvas-Hébert[1], Vincent Taschereau-Dumouchel[2,3], Ian Charest[1], Frédéric Gosselin[1]; [1]Cerebrum, Département de psychologie, Université de Montréal, [2]Département de psychiatrie et d'addictologie, Université de Montréal, [3]Centre de recherche de l'Institut Universitaire en Santé Mentale de Montréal

Recent neuroimaging studies using fMRI and EEG have consistently revealed overlapping brain activation during both mental imagery and visual perception. Yet, the extent to which these processes share underlying mechanisms remains elusive. Our prior work uncovered a weak correlation between perceptual thresholds and mental imagery (i.e., vividness judgments) for identical natural scenes (Charest et al., 2023), prompting a new investigation using adaptation as a psychophysiological tool. Adaptation is an invaluable tool for non-invasive exploration of low- to high-level visual processing, including face (e.g. Leopold et al., 2001), object (e.g. Feng & He, 2005), and scene (e.g., Greene & Oliva, 2005) processing. Importantly, adaptation has been previously used to induce aftereffects of imagined motion (Winawer et al., 2010). However, it is still unknown if adaptation can elicit aftereffects following high-level adaptation. In each trial, participants viewed the same two full-frontal, color faces simultaneously for 1 second, one on each side of a fixation cross. Subsequently, they were instructed to imagine the face previously shown either on the left or the right for 6 seconds. Participants then assessed whether a morph resembled the face initially shown on the left or the right (40 repetitions × 7 morph levels × 2 imagined faces, totaling 560 trials). Preliminary results from five participants revealed significantly different means for the cumulative Gaussian distributions fitted to the proportions of responses in favor of face B as a function of morph levels when face A or face B was imagined ($t(4) = -4.69$, $p < 0.01$; adapted to face A: M = 0.55; and to face B: M = 0.61). These initial results offer a promising avenue for finely comparing high-level visual processing and mental imagery across individuals with diverse cognitive proficiencies, paving the way for a deeper understanding of the interconnected nature of these cognitive domains.

36.412 AN EFFICIENT MULTIMODAL FMRI LOCALIZER FOR HIGH-LEVEL VISUAL, AUDITORY, AND COGNITIVE REGIONS IN HUMANS

Samuel Hutchinson[1] (samhutch@mit.edu), Ammar Marvi[1], Freddy Kamps[1], Emily M. Chen[1], Rebecca Saxe[1], Ev Fedorenko[1], Nancy Kanwisher[1]; [1]Massachusetts Institute of Technology

Although localizers for functional identification of category-selective regions in individual participants are widely used in fMRI research, most have not been optimized for the reliability and number of functionally-distinctive regions they can identify, or for the amount of scan time needed to identify these regions. Further, functional localizers for regions in high-level visual cortex do not enable localization of cortical regions specialized for other domains of cognition. Here we attempt to solve these problems by developing a single localizer that enables reliable localization in just 23 minutes of fMRI scan time of cortical regions selectively engaged in processing faces, places, bodies, words, and objects, as well as cortical regions selectively engaged in processing speech sounds, language, and theory of mind. To this end, we use a blocked design in which participants watch videos from five different visual categories (of scenes, faces, objects, words, and bodies), while simultaneously listening to and performing tasks on five different kinds of audio stimuli (false belief sentences, false photo sentences, arithmetic problems, nonword strings, and texturized speech). We counterbalance these conditions across five runs of ten blocks each, with each block consisting of one 21-second auditory stimulus and seven three-second videos from one visual category. Each visual stimulus occurs equally often with each audio stimulus, so that contrasts in each modality are unconfounded from conditions in the other. Data from ten participants show that this Efficient Multimodal Localizer robustly identifies, within individual participants, cortical regions selectively engaged in processing faces, places, bodies, words, and objects, as well as speech sounds, language, and theory of mind, as tested against established standard localizers for these functions. The stimuli and presentation code for this new localizer will be made publicly available online, enabling future studies to identify functional regions of interest with the same procedure across multiple labs.

This work was supported by NIH grant 1R01HD103847-01A1 (awarded to RS) and NIH grant 5UM1MH130981.

36.413 BEHAVIOURAL AND ERP CORRELATES OF EYE-MOVEMENT PATTERNS IN FACE PERCEPTION

Nianzeng Zhong[1,2], Janet Hsiao[2], William Hayward[1]; [1]Lingnan University, Hong Kong, [2]The University of Hong Kong

Faces all show the same basic facial features in the same general arrangement, but observers do not all adopt the same eye-movement patterns when looking at them. We conducted two experiments to investigate how face-viewing preferences affect face-specific event-related potentials (ERPs). In Experiment 1, participants were asked to complete a free eye-movement face identification block and two fixation number-controlled blocks in which the initial fixation location was restricted to the left eye or the nose. During this task, we used EEG to record participants' scalp electrical potentials. Participants were separated into two groups according to their eye movements in the free eye-movement block, with an upper-focused group who favored the eyes of faces, and a lower-focused group who favored the nose and mouth. In fixation number-controlled blocks, the upper-focused group performed better, fixated longer, and elicited a larger N170 amplitude for the eye-fixation condition than the nose-fixation condition. In contrast, the lower-focused group performed and fixated evenly and elicited comparable N170 amplitudes between fixation conditions. In addition, the P1 component was larger for the nose-fixation condition than the eye-fixation condition, and this difference was bigger for the lower-focused group than the upper-focused group. On the other hand, the P1 appeared to have no relation to behavioral performance. In Experiment 2, new subjects completed two fixation duration-controlled blocks in which subjects were asked to keep fixating on the left eye or the nose of faces for 250ms or 500ms, and the behavioral and ERP results in Experiment 1 were replicated. These findings suggest that face identification is associated with an

individual's favored looking pattern tuned by fixation location on a face, and the N170 provides an index of this identification performance.

36.414 CONCURRENT FACE-SELECTIVE NEURAL ACTIVITY ACROSS THE VENTRAL OCCIPITOTEMPORAL CORTEX REVEALED WITH HUMAN INTRACEREBRAL RECORDINGS

Corentin Jacques[1] (corentin.g.jacques@uclouvain.be), Jacques Jonas[1,2], Louis Maillard[1,2], Sophie Colnat-Coulbois[1,3], Bruno Rossion[1,2]; [1]Université de Lorraine, CNRS, Nancy, France, [2]Université de Lorraine, CHRU-Nancy, Service de Neurologie, Nancy, France, [3]Université de Lorraine, CHRU-Nancy, Service de Neurochirurgie, Nancy, France

Recognition of object categories from the visual environment is a fundamental brain function, supported in humans by a bilateral network of regions extending ventrally from the occipital pole to the anterior temporal cortex. While dominant models of visual recognition describe this network as being organized hierarchically, with distinct processing stages increasing in representational complexity, abstraction and timing from posterior to anterior regions, direct evidence for this hierarchical organization in the human brain is lacking. Here we take advantage of the high temporal and spatial resolution provided by intracerebral recordings in a large sample of participants (N=140) to provide an extensive characterization of the time-course of category-selective responses to natural images of faces – arguably the most familiar and ecologically valid stimulus in the human environment – across the whole ventral occipitotemporal cortex (VOTC). Objectively tagged face-selective responses in the high-frequency broadband range (30-160 Hz) were found in 10% of recording sites distributed over the whole VOTC, with regional peaks of activity around and along the length of the fusiform gyrus, leading to three key observations. First, face-selective neural activity is characterized both by signal increase (57% of sites; 65% of total response amplitude) and decrease (43% of sites; 35% of amplitude) relative to other object categories, with a respective lateral-medial distinction along the VOTC. Second, an increase in degree of selectivity is observed along the postero-anterior axis, with about 50% of exclusive response to faces in the anterior regions of the ventral temporal lobe. Third and most importantly, face-selective neural activity occurs largely concurrently (80-100ms onset – 350-450ms offset) across the VOTC, with highly overlapping time-courses between regions. Overall, these observations directly challenge the standard view of a hierarchical neural organization of visual object recognition in the human brain and call upon alternative neurofunctional models of this key brain function.

36.415 CROSS-VALIDATING THE NEUROFUNCTIONAL ELECTROPHYSIOLOGICAL MARKERS OF EARLY FACE CATEGORIZATION

Lisa Stacchi[1], Fazilet Zeynep Yildirim-Keles[1], Roberto Caldara[1]; [1]University of Fribourg, Switzerland

The N170 event-related potential (ERP) is the most widely investigated neurofunctional marker of early face categorization. In the past ten years, there has been a surge in research using the fast periodic visual stimulation (FPVS) methodology to delve into face categorization. FPVS studies have consistently reported robust bilateral face categorization responses over the occipitotemporal cortex with a right hemispheric dominance, closely mirroring the N170 topography. Yet, the precise neurofunctional correspondence between these two electrophysiological markers of face categorization remains elusive. To address this issue, we recorded the electrophysiological signals of human observers who viewed natural images of faces and objects using both ERP and FPVS paradigms. In the FPVS paradigm, participants were exposed to rapid presentations of objects with periodically intervening face stimuli, while in the ERP paradigm faces and objects were presented in isolation. We quantified the FPVS response in the frequency domain and extracted ERP visual components, including the P1, N170 and P2 in response to face stimuli, from both the ERP paradigm as well as the time domain of the FPVS response. Only the N170-P2 peak-to-peak amplitude recorded in the ERP paradigm significantly explained the amplitude of the FPVS frequency response. The N170-P2 peak-to-peak amplitude was also significantly associated with the respective component in the FPVS time domain, while the N170 was not. Altogether, our data show that the FPVS frequency response is not singularly contingent on any isolated ERP component, such as the N170, but rather reflects a later complex neural integration. These findings raise crucial methodological and theoretical considerations on the relationship between FPVS and ERP responses, urging caution when interpreting the neurofunctional role of both electrophysiological signals. Both markers relate to distinct neural computations and dynamics of how the human brain processes faces.

This work was supported with funding from the Swiss National Science Foundation awarded to RC (10001C_201145).

36.416 DYNAMICS OF FACE PERCEPTION: UNRAVELING THE ROLE OF EYES AND MOUTH IN NEURAL PROCESSING

Yasemin Gunindi[1] (yasemingunindi@sabanciuniv.edu), Çiçek Güney[1], Huseyin Ozkan[1], Nihan Alp[1]; [1]Sabanci University

Faces, as crucial conveyors of social information, are often studied using static images or dynamic videos to analyze emotional aspects, overlooking the significance of part-based dynamic information in neutral faces. Hence, the extent to which part-based dynamic information, primarily derived from the eyes and mouth, contributes to dynamic face perception remains elusive. In this study, using neutral dynamic face stimuli, we investigate how the brain processes part-based information during dynamic face perception with specific emphasis on its ability to discriminate between forward and backward face videos over time. Participants fixated on a central cross while watching 3-second grayscale muted videos featuring individuals speaking in neutral state. We manipulate the face orientation (right-side-up, upside-down) and the presence of eye blink (with/without blink), and ask participants to indicate the temporal order of the dynamic face videos as forward or backward. The eyes and mouth were contrast-modulated at 6 and 7.5 Hz, respectively. Steady-state visual evoked potentials were recorded from 64 EEG channels. Behavioral results (d' primes > 0) indicate that participants performed the task well. EEG results reveal an orientation effect consistent with the literature. The topographic map of neural responses indicates a central-occipital focus for the eye and a lateral-occipital focus for the mouth. The overall neural response shows that there is a bias towards the mouth when face orientation is right-side-up. We observed that neural responses to eyes tend to be differentially more elevated

compared to mouth when there is a blink. In intact faces (forward, right-side-up), the blink information closes the mouth bias. Whereas the most-distorted case (backward, up-side-down) requires a separate search for a cue (blink) in the eyes, suppressing the mouth part. Overall, this study significantly contributes to our understanding of dynamic face perception, emphasizing the role of dynamic part-based information, particularly eyes and mouth movements.

This work was funded by TUBITAK 1001 (122K922).

36.417 EFFECTIVE CONNECTIVITY OF THE HUMAN CORTICAL FACE NETWORK THROUGH CONCURRENT INTRACEREBRAL ELECTRICAL STIMULATION AND FREQUENCY-TAGGED VISUAL PRESENTATION

Luna Angelini[1] (luna.angelini@univ-lorraine.fr), Bruno Rossion[1,2], Corentin Jacques[1], Marie-Alphée Laurent[1], Louis Maillard[1,2], Sophie Colnat-Coulbois[1,2], Jacques Jonas[1,2]; [1]University of Lorraine, CNRS, Nancy, France, [2]CHRU-Nancy (Nancy Regional University Hospital), Neurology department, France

The neural basis of human face recognition has been extensively studied for decades. The contribution of several face-selective brain regions in the ventral occipito-temporal cortex (VOTC; e.g., in the fusiform gyrus, FG; inferior occipital gyrus, IOG) has been established, but the functional organization of this cortical network remains largely unknown. In rare epileptic patients implanted with intracerebral electrodes in several face-selective VOTC regions, an original combination of frequency-tagging and direct electrical stimulation (DES) can provide unique information to address this issue. These depth electrodes allow us to stimulate a local node of the network while recording the functional activity of other implanted regions, with high spatial and temporal resolution. Here we report the case of a 34-year-old man, presented with 70-second sequences of natural images (objects with one face inserted every five images, at a 6Hz rate), while focal stimulation (1,0mA at 55Hz) was separately applied for 10s to four independently defined face-selective areas in the right and left VOTC (right and left lateral FG, right IOG, right anterior FG). Upon stimulation, face-selective neural activity was reduced or abolished both locally and at remote bilateral VOTC recording sites. These remote effects of DES were found in both postero-anterior (i.e., DES to lateral occipital cortex affecting fusiform face-selective activity) and antero-posterior (DES to fusiform gyrus affecting lateral occipital cortex activity) directions as well as across face-selective sites of homologous regions of the two hemispheres. Most importantly, this reduction was extremely specific to the face-selective response, as it did not extend to the general visual response (i.e., 6Hz and harmonics). Overall, these results shed original light on the functional connectivity of the human cortical face recognition network and pave the way for a more widespread and systematic development of this approach to reveal the functional and effective connectivity of human brain networks.

36.418 EXPLORING VISUAL STRATEGIES AND THEIR ELECTROPHYSIOLOGICAL CORRELATES IN SAME AND OTHER-RACE FACE PROCESSING

Isabelle Charbonneau[1], Vicki Ledrou-Paquet[1], Anthony Proulx[1], Arianne Richer[1], Laurianne Côté[1], Caroline Blais[1], Justin Duncan[1], Daniel Fiset[1]; [1]Universite du Quebec en Outaouais

In the realm of face perception, it has been suggested that faces belonging to one's own race are processed differently than those of other races, leading to superior recognition of same-race faces (Meissner & Brigham, 2001; Malpass & Kravitz, 1969). This phenomenon, known as the Other-Race Effect (ORE), has been extensively examined, notably through eye-tracking studies that have shown that White individuals allocate less attention to the eyes of Black faces compared to White faces (e.g. Kawakami et al., 2014). To better understand this bias, we first asked 15 White participants to complete a face memory task, following an old/new paradigm with both Black and White faces. Replicating the ORE (i.e. better accuracy (d') in memorizing white (M= 1.59, SD = .70) than black faces (M= .75, SD = .33): $t(14) = 7.02$, $p < .001$, Cohen's d = 1.8, 95% CI [0.59, 1.1]), participants then completed two other tasks (gender and smile/neutrality discrimination) while their EEG signals were recorded (for a total of 6000 trials/participant). In each trial, distinct parts of Black and White faces were revealed using the bubbles method (Gosselin & Schyns, 2001). Multiple linear regression analyses using a Pixel Test (Stat4Ci Toolbox; Chauvin et al., 2005) on EEG amplitudes at specific electrodes of interest (e.g., PO8, PO7) revealed strong associations with the eye region within the N170 time window, regardless of the task or the race of the faces. These findings suggest that same and other-race faces undergo similar processing during the early stages of face perception, with differences likely emerging later in the face identification stream.

Social Sciences and Humanities Research Council (SSHRC)

36.419 FACE-RELATED ACTIVITY IN SUPERIOR COLLICULUS AND TEMPORAL CORTEX OF PRIMATES

Gongchen Yu[1] (yugongchen1990@gmail.com), Leor Katz[1], Christian Quaia[1], Richard Krauzlis[1]; [1]Laboratory of Sensorimotor Research, National Eye Institute, NIH

Face processing has been extensively studied in higher-order visual cortex (e.g., temporal cortex 'face patches'). Recently, we reported that neurons in the macaque superior colliculus (SC) display a preference for face stimuli at extremely short latencies. Here, we directly compared face-related visual responses in the SC and temporal visual cortex. Using the same subject, stimuli, and paradigm, we recorded neurons in the middle face patch of the temporal cortex, as well as the SC, under both foveal and peripheral (6 degree contralateral) presentation conditions. In temporal cortex, with foveal presentation, 53% of visually responsive neurons exhibited a significant face preference at 100ms after stimulus onset, but only 8% displayed a face preference at 50ms. With peripheral presentation, only a minority of the same neurons exhibited a significant face preference either early (12% at 50ms) or late (14% at 100ms). In contrast, a majority of SC neurons displayed a significant face preference both early and late, for both foveal and peripheral stimulus presentations (foveal: 69% at 50ms, 43% at 100ms; peripheral: 49% at 50ms, 45% at 100ms). Overall, for temporal cortex neurons, the median latency for the face preference was 79ms and 93ms, for foveal and peripheral conditions respectively. For SC neurons, the median latencies were markedly shorter – 40ms and 42ms. Thus, a face preference emerges earlier for neurons in the SC than in the temporal cortex, and the preference is equally strong for foveal and peripheral locations for SC neurons, whereas the preference in the temporal cortex shows a strong bias towards central vision. Our results reveal distinct properties of face processing in the SC and temporal cortex that may support complementary functions in the primate visual system – rapid face detection in the visual periphery by the SC and more sophisticated face processing by temporal cortex following foveation.

This work was supported by the National Eye Institute Intramural Research Program at the National Institutes of Health ZIA EY000511

36.420 LANGUAGE MODEL PREDICTION OF VISUAL CORTEX RESPONSES TO DYNAMIC SOCIAL SCENES

Emalie McMahon[1], Colin Conwell[1], Kathy Garcia[1], Michael F. Bonner[1], Leyla Isik[1]; [1]Johns Hopkins University

Recent work has shown that language models based on sentence captions of images are good models of high-level ventral visual cortex, on par with vision models. Text manipulation experiments reveal that this match to the ventral stream is strongly dependent on the nouns in the image captions, suggesting that language models perform well because they represent the things (i.e., agents and objects) in an image. However, the visual world is much richer than static things. We see people dynamically interacting with objects and other people. These dynamic scenes have been shown to more strongly activate visual cortex, and high-level lateral regions, in particular, uniquely respond to dynamic social content. Can vision and language models predict responses to dynamic social scenes in ventral and lateral visual cortices? To investigate this question, we used a large-scale dataset of three-second clips of social actions and collected sentence captions of each clip. Comparing the prediction of vision and language models, we first find that language and visual models similarly predict responses in ventral visual cortex, extending prior work with static images to dynamic scenes. In contrast, we find that language models outperform vision models in predicting lateral visual cortex. Next we performed sentence manipulation experiments in which we selectively remove parts of speech from the sentence captions. First, we replicate prior work that nouns, but not verbs, alone yield high prediction of ventral visual cortex. In contrast, in lateral visual cortex, verbs and nouns are similarly highly predictive, and removing only the verbs impairs performance more than only removing the nouns. Taken together, these results suggest that language models' match to lateral visual cortex relies on action information and that good models of these regions must contain representations of not just agents and objects but also their actions and interactions.

Funded by NIH R01MH132826 awarded to LI.

SUNDAY, MAY 19, 2:45 – 6:45 PM, PAVILION

Face and Body Perception: Development, experience

36.421 EXAMINING SPONTANEOUS NEURAL ACTIVITY PATTERNS IN DEVELOPMENTAL PROSOPAGNOSIA USING RESTING STATE EEG

Maruti Mishra[1,2,3] (maruti.mishra07@gmail.com), Erin Hugee[4], Patrick Sutphin[3], Kevin Spencer[2], Joseph DeGutis[2], Cindy Bukach[3]; [1]California State University Bakersfield, CA, USA, [2]VA Boston Healthcare/Harvard Medical School, Boston, MA, [3]University of Richmond, Virginia, USA, [4]Northwestern University, IL, USA

Developmental prosopagnosia (DP) is characterized by a lifelong difficulty in learning and recognizing faces. While task-based studies have identified brain regions associated with face processing deficits in DPs, less is known about the spontaneous, intrinsic brain activity patterns in the absence of explicit tasks. To address this, resting-state electroencephalography (EEG) data were obtained from 32 DPs and 25 controls during five-minute open-eye recordings. The data were preprocessed at 0.5 to 100 Hz, ICA artifact corrected and average referenced to extract absolute spectral power values for delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (14-30 Hz), and gamma (30-80 Hz) bands for each electrode and then pooled into regional electrode clusters (frontal, central, parietal, temporal and occipital). A 3 way mixed ANOVA for the two groups, hemispheres, and regions revealed significant interactions for alpha and gamma bands, highlighting hemispheric differences. Individuals with DP displayed significantly lower alpha power in the left hemisphere compared to the right. This was driven by central and temporal electrode regions. In contrast, both groups exhibited significantly greater gamma power in the left hemisphere, with controls showing a larger hemispheric asymmetry. This asymmetry was driven by electrodes in the parietal, temporal, and occipital regions in controls, leading to significant DP vs control group differences. Alterations in alpha power have been associated with neuronal excitability and gamma in higher-level feature binding, both critical for face processing. Furthermore, both alpha and gamma have also been implicated in N170 ERP for face specific processing. Our findings provide evidence for the first time that there are cross-hemisphere differences in resting state alpha and gamma band power between DPs and controls.

NSF and James S. McDonnell Foundation Scholar Award to Cindy Bukach; R01EY032510-02 to Joe DeGutis

36.422 LIKE ADULTS, CHILDREN CANNOT DISTINGUISH BETWEEN AI-SYNTHESIZED FACES FROM REAL ONES BUT EXHIBIT A RESPONSE BIAS TOWARD SYNTHESIZED FACES

Sarina Hui-Lin Chien[1,2] (sarinachien@mail.cmu.edu.tw), Pei-Tzu Liu[1], Yu-Ting Ma[1], Chien-Kai Chang[1]; [1]Insititue of BIomedical Sciences, China Medical University, Taichung, Taiwan, [2]Center for Neuroscience and Brain Diseases, China Medical University, Taichung, Taiwan

Introduction. Generative adversarial networks have become popular platforms for synthesizing realistic visual content. In a large-sample survey, Nightingale & Farid (2022) demonstrated that AI-synthesized (GANsyn2) faces are indistinguishable from real faces and more trustworthy to adults. Here we extended this study to test children and adults in the East Asian context and further explored whether such ability is correlated with proficiency in featural/configural processing. Methods. Thirty-four 5- to 12-year-old children (18 boys) and 34 adults (17 males) participated in the study. All participants received three tasks: a computerized face judgment task whereby participants judged whether the image was a real person or synthesized by AI, a trustworthiness rating task whereby participants rated the trustworthiness of each face on a 1 (very untrustworthy) to 5 (very trustworthy) point-scale, and a paper-and-pencil face discrimination test containing two target faces and 12 comparison faces with modifications either on the eyes, nose, or mouth that participants had to mark the locus of alterations. Results. Neither adults (M = 0.476 ±0.015, d'=0) nor children (M =0.460 ±0.015, d'=0) were able to discriminate between real and synthesized faces better than chance. However, adults performed relatively better with identifying real faces (HIT/CR<1), whereas children were better at identifying synthesized faces (HIT/CR>1). Synthesized faces (M =3.116±0.042) were rated more trustworthy than real ones (M =2.696 ± 0.049) in all participants. Lastly, adults (M = 7.853±0.368) performed better than children (M = 4.971±0.361) in the face discrimination test, but the individual's test score did not correlate with the accuracy of the face judgment task. Conclusion. Like adults, children cannot distinguish between AI-synthesized faces from real ones but exhibit a response bias toward synthesized faces. Proficiency in featural/configural processing does not contribute to the face authenticity judgment.

We are thankful to Dr. Sophie Nightingale for providing valuable suggestions and stimuli. This project is funded by Taiwan National Science and Technology Council grant NSTC 112-2423-H-002-002 to Dr. SHL Chien and Dr. CC Chen, and by MOST 111-2410-H-039-007 grant to Dr. SHL Chien.

36.423 THE RELATIVE CONTRIBUTION OF FACE AND BODY FOR PERSON RECOGNITION IN INFANCY

Megumi Kobayashi[1] (mkobayashi@human.niigata-u.ac.jp), So Kanazawa[2], Masami K. Yamaguchi[3], Alice J. O'Toole[4]; [1]Niigata University, [2]Japan Women's University, [3]Chuo University, [4]The University of Texas at Dallas

In everyday life, people can identify others from a distance. Viewing distance plays a key role in determining the relative importance of face versus body information for person recognition (e.g., O'Toole et al., 2011). When a person is nearby, adults rely more on the face than body. However, at a distance, adults rely on both the face and body (Hahn et al., 2016). Despite evidence indicating that infants observe the bodies of other people from a young age (Jayaraman et al., 2017), no studies have explored infants' recognition of people based on both face and body information. This study fills this gap by examining 5- to 7-month-old infants' ability to recognize approaching people. Here, we used a familiarization/novelty-preference procedure. In Experiment 1, infants (N = 120) were familiarized with videos of a person approaching from a distance and tested on their recognition from distant and close views. We found that 5- to 7-month-olds were able to recognize the approaching person from both distances (all ps < .01). In Experiment 2, infants (N = 217) were familiarized with an approaching person, as in Experiment 1, but were tested with a video of approaching people with blurred faces (allowing access to body information) or blurred bodies (allowing access to face information) from distant and close views. At a distance, 5- and 7-month-olds recognized the approaching person based on body information (all ps < .01). When nearby, 5- and 6-month-olds recognized the approaching person based solely on face information (face: p < .01, body: p < 1), whereas 7-month-olds recognized the person based on both face and body information (face: p < .01, body: p = 0.02). We conclude that the relative contribution of face and body for person recognition would change with increasing age.

This research was supported by JSPS KAKENHI grant for MK (23H01055) and the Ministry of Education, Culture, Sports, Sciences and Technology (MEXT) KAKANHI grant for MKY (17H06454)

36.424 DOES CONFIDENCE PREDICT FACE-IDENTIFICATION ACCURACY FOR SAME-RACE AND OTHER-RACE FACES?

Geraldine Jeckeln[1] (geraldine.jeckeln@utdallas.edu), Alice J. O'Toole[1]; [1]The University of Texas at Dallas

Confidence in face-identification decisions can influence legal outcomes. Additionally, people are less accurate at recognizing other-race than same-race faces (other-race effect [ORE]). Although confidence predicts face-identification accuracy for same-race faces (Hahn et al., 2021; Jeckeln, et al., 2022), it is unclear whether this finding holds for other-race faces. Here, we used a comparative-confidence task (Mamassian, 2016) to examine observers' ability to evaluate the correctness of their face-matching decisions for same-race and other-race faces. Participants (27 White/Caucasian [W/CA], 27 Black/African American [B/AA]) completed a comparative-confidence task embedded in a face-identity matching test: On each face-matching trial (24 B/AA trials, 24 W/CA trials), participants viewed three face images (two same-identity images and one different-identity image), and selected the odd-one-out. After completing two face-matching trials (2 B/AA trials or 2 W/CA trials), participants selected the trial on which they felt more confident (confidence-selected trial). There was an ORE for identification accuracy ($F(1, 54) = 6.25$, $p = .015$, $\eta2= .10$). However, the ORE did not affect the confidence-accuracy relationship: Accuracy was greater for confidence-selected trials than confidence-rejected trials for both same-race and other-race faces ($p < .01$). Additionally, trial difficulty (DifficultyTrial1-DifficultyTrial2) predicted B/AA participants' confidence choice for same-race ($R^2 = .51$, $F(1,12) = 10.59$, $p < .01$) and other-race faces ($R^2 = .69$, $F(1, 10) = 21.87$, $p < .001$), and W/CA participants' confidence choice for same-race faces ($R^2 = .47$, $F(1, 10) = 8.712$, $p = .0145$), but not other-race faces ($p = .2519$). These results demonstrate that confidence serves as a predictor of accuracy in the context of same-race and other-race face-identity matching and that confidence is driven by trial difficulty.

Research funded by The National Institute of Standards and Technology, Grant to A.OT.

36.425 EXAMINING THE ROLE OF FACIAL EXPRESSIONS IN UNFAMILIAR FACE RECOGNITION

Shruti Vyas[1] (shruti.vyas@torontomu.ca), Margaret Moulson[2]; [1]Toronto Metropolitan University

Facial expressions are both consistent and unique visual components of a face for each individual person, and this information may be useful when learning and recognizing faces. By using matching paradigms, some studies have shown that facial expressions (e.g., smiles) facilitate face recognition, but only when the faces are shown simultaneously and with limited exposure to faces with natural variability. It remains unclear whether facial expressions help or hinder face recognition when naturally varying face stimuli are shown sequentially. To address this gap in the literature, we are conducting a study to examine whether facial expressions play a role in face recognition for naturally varying unfamiliar faces. This study aims to investigate differences in matching accuracy after viewing an array of either expressive faces (smiling) or non-expressive (neutral) faces, followed by a test face with a matched or mismatched identity. Participants are presented with 40 smiling and 40 neutral trials in random order, with congruent expressions between learning and test faces in each trial. Currently, 24 participants have been recruited (expected n=50) from the Toronto Metropolitan University undergraduate research pool. We hypothesized higher matching accuracy (% correct responses) in conditions with smiling faces at test vs. neutral faces, suggesting better identity recognition for faces displaying an expression. A preliminary 2 (Identity: matched vs. mismatched) x 2 (Expression: Neutral vs. Smiling) within-subjects ANOVA showed a statistically significant main effect of identity $F(1, 23.00) = 57.03$, $p < 0.001$), and expression $F(1, 23.00) = 9.63$, $p = 0.005$). Specifically, participants showed greater matching accuracy for match trials vs. mismatch trials regardless of expression, and greater accuracy for smiling match trials ($M = 18.54$, $SD = 1.92$) vs. neutral match trials ($M = 17.96$, $SD = 2.07$). These findings will be further interpreted when data collection is completed (expected completion by March 2024).

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A2C3004133).

36.426 FAST READOUT OF IDENTITY-RELATED INFORMATION IN MULTIPLE IMAGES AFTER INCIDENTAL LEARNING

Jieun Cho[1] (jieuncho@yonsei.ac.kr), Sang Chul Chong[1,2]; [1]Graduate Program in Cognitive Science, Yonsei University, [2]Department of Psychology, Yonsei University

Unfamiliar face images often lead to the erroneous categorization of variations as distinct individuals (Jenkins et al., 2011). This tendency is presumed to arise from the multifaceted features inherent in faces, which naturally exhibit variations. Hence, encountering natural variations in a person, such as viewing the person from different angles, under various lighting conditions, with different facial expressions, or makeups, can handle this difficulty through which we identify stable, unchanging features across multiple instances. Different studies on face learning have emphasized the necessity of conceptual learning which involves direct supervision or label information over mere perceptual learning, especially for better generalization on recognizing new instances of learned identities. In the current study, we investigated how exposure to a same person's variable images via incidental learning, without explicit supervision, influences the perception of newly encountered variations of a learned face—particularly when integrating dissimilar images into a unified identity. In the learning phase, target identities were intermittently presented amid a lot of other distractor faces, with variations within the same identity. Before and after the learning phase, we asked participants to approximate the number of distinct identities among 24 face images from two different people as soon as possible within 8 seconds. Participants either performed the task observing the encountered exemplars during the learning phase or novel exemplars of the learned identities. When comparing the estimated number of identities before and after learning, a prominent decrease in the number was observed after learning. This suggests that the incidental, perceptual learning enables rapid readout of abstract information about specific identities from multiple images. Moreover, the number was smaller in the novel-exemplar condition than the old-exemplar condition. These results indicate that people can equally generalize to new faces through perceptual learning, as long as a variety of instances is presented during learning.

36.427 FASTER CATEGORIZATION OF THE MAJORITY-RACE FACES IN A MULTIRACIAL SOCIETY

Charles C.-F. Or[1] (charlesor@ntu.edu.sg), Bo Zhen Lu[1], Nadia Mazlan[1], Yi Jie Toh[1], Alan L.F. Lee[2]; [1]Nanyang Technological University, [2]Lingnan University

The effect of race on face perception is less understood in multiracial societies, especially when faces are categorized against other objects (face categorization). As facial appearances differ across racial groups, categorizing faces of different races into the same generic "face" category is not a trivial process. In addition, the racial background of the observer could modulate such face-categorization performance. Thus, we seek to understand whether the race of the face and the observer's race influence the speed of face categorization among racially diverse Singaporeans, of which Chinese are the majority (>70%) while Indians and Malays are significant minorities (>20% together). Such effects were examined by a verified, sensitive rapid serial visual presentation (RSVP) paradigm (Or, Goh, & Lee, 2021, Vision Research). Singaporean observers (N=107; 32 Chinese, 39 Indians, 36 Malays) were presented with a rapid sequence of 4650 color natural images each containing either a face (Chinese, Indian, or Malay; 150 in total) or a non-face object randomly (SOA=83.3ms). Observers were instructed to press a key as soon as they saw a face each time in the RSVP sequence. Generally, observers responded faster to Chinese faces than to Indian and Malay faces, regardless of the race of the observer. Importantly, even minority observers showed speed advantages in categorizing the majority-race Chinese faces. This was surprising as we did not find the same-race preferences commonly reported in face-identification studies conducted in monoracial communities. Interestingly, such majority-race advantages were eliminated for these minority observers who performed a second sequence in grayscale, indicating that color interacts with face race during face categorization. This novel preference to the majority-race faces suggests that face-categorization performance depends on the racial diversity of the society which influences people's racial experience.

Supported by Singapore MOE AcRF Tier 1 Grants (RG51/19 to C.O.; RG59/19 to C.O. & A.L.), SGUnited Traineeships Programme to N.M., NTU CoHASS Incentive Schemes to C.O., and the NTU-URECA programme.

36.428 HOW THE LEARNING OF UNFAMILIAR FACES IS AFFECTED BY THEIR SIMILARITY TO ALREADY-KNOWN FACES

Maddie Atkinson[1] (m.atkinson2@pgr.reading.ac.uk), Kay Ritchie[2], Peter Hancock[3], Katie L.H Gray[1]; [1]University of Reading, [2]University of Lincoln, [3]University of Stirling

What happens as an unfamiliar face becomes familiar is a process that is poorly understood. Recent evidence suggests that unfamiliar faces are processed in relation to pre-existing face representations. It is proposed that, when a novel face resembles someone whom we already know, the familiar representation is somewhat activated at encoding. Although familiarity has been shown to have a facilitatory effect on novel face recognition, it is unknown how the learning of unfamiliar and 'similar-to-familiar' faces differ. To examine this, 91 participants completed a training-test procedure. During training, participants were exposed to multiple arrays containing ambient images of novel 'similar-to-familiar' (i.e., UK or American celebrities' less well-known siblings) and unfamiliar (i.e., Spanish celebrities' siblings) identities. Exposure duration was also manipulated; training arrays were presented for either 60s or 120s. Following each training array, participants provided an estimate of the total number of identities present. At test, participants were required to make binary 'familiar'/'novel' judgements to previously unseen target and distractor images; d' scores were calculated for each participant. Results revealed that identity estimates followed a decreasing linear trend across training trials; although participants become significantly more accurate over time, this trend did not vary as a function of familiarity or exposure duration. At test, sensitivity (d') was significantly higher for 'similar-to-familiar' faces relative to unfamiliar faces. As expected, there was also a significant effect of exposure duration, where longer encoding durations resulted in better learning. In addition, these main effects were qualified by an interaction; sensitivity was higher for 'similar-to-familiar' compared to unfamiliar faces in the long exposure condition but were not significantly different in the short exposure condition. The present findings suggest that 'similar-to-familiar' faces are learned more readily than unfamiliar faces, demonstrating that already stored face representations impact how we learn new faces.

36.429 PHYSICALLY BLURRED FACES ARE MORE RECOGNIZABLE AT A DISTANCE

Lei Yuan[1] (cyuan26@student.ubc.ca), Claudia Wu[1], Ipek Oruc[1]; [1]University of British Columbia

Previous work has shown that faces with a high degree of blur are more recognizable at smaller sizes. In addition, blurry faces generate identity adaptation aftereffects at small, but not large, sizes. This small-size advantage for blurry faces has been observed using digitally scaled and blurred face images. Here, we examine whether the small-size advantage persists in a paradigm with greater ecological validity, where size variation is achieved through viewing distance, and blurring is achieved via physical blurring filters. We tested 12 participants (5 females, 7 males, ages: 20-44 years, M=26.67, SD=7.16) in a face recognition protocol at two viewing distances (1m and 0.35m) corresponding to small (1.66 degree) and large (4.75 degree) face sizes. At each trial, a celebrity face was displayed on a computer screen until response. Trials were blocked by viewing distance (close vs. far), where a randomly selected half of 100 celebrity images were viewed at each distance. The order of blocks was counterbalanced across participants. Faces were viewed behind a 1-degree Luminit holographic Light Shaping Diffusers (LSD®), such that blur level was fixed throughout the experiment. Participants were asked to identify the celebrity at each trial. If the participant was unable to correctly identify the face, they were immediately shown intact images of the celebrity to determine whether they were unfamiliar with the celebrity, in which case the trial was discarded. Face recognition accuracy was 66.22% at the small size, significantly higher than the 60.19% at the large size (p=0.014, d=0.39). Ten out of 12 participants showed the small-size advantage at the individual level. These results replicate earlier findings of better recognition for blurry faces at smaller sizes with digital manipulation of blur and size, and extend them by demonstrating that this small-size advantage generalizes to physically blurred faces viewed at varying distances.

This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) USRA award (LY), an NSERC Discovery Grant RGPIN-2019-05554 (IO) and an Accelerator Supplement RGPAS-2019-00026 (IO).

36.430 THE IMPORTANCE OF CONCEPTUAL KNOWLEDGE WHEN LEARNING NEW FACES DURING NATURALISTIC VIEWING

Kira Noad[1] (kira.noad@york.ac.uk), David M. Watson[1], Timothy J. Andrews[1]; [1]University of York

Although the ability to recognise familiar faces is a critical part of everyday life, the process by which a face becomes familiar in the real world is not fully understood. Previous studies have focussed on the importance of perceptual experience with faces. However, in natural viewing, perceptual experience with faces is accompanied by increased knowledge about the person and the context in which they are encountered. Here, we used a natural viewing paradigm to investigate how the context in which events are presented affects subsequent face recognition. Participants viewed clips from the TV show Life on Mars, where context was manipulated by presenting events either in 1) the original sequence, or 2) a scrambled sequence. While the manipulation significantly affected the understanding of the events, it had no effect on the perceptual exposure to the faces. Nevertheless, we found that recognition of faces was greater in participants from the original group compared to the scrambled group. Moreover, individual differences in conceptual understanding correlated with face recognition. Next, we investigated how our understanding of conceptual information is reflected in the brain. After having previously seen either the original or scrambled version of the movie, participants now viewed a new movie from Life on Mars while neural activity was recorded using fMRI. Inter-subject correlation (ISC) was used to measure the similarity of the response across participants in the original or scrambled groups. We found higher ISCs between participants in the original group within a network of regions that are typically associated with the processing of semantic, episodic and affective information. Together, these findings suggest that conceptual information is important for learning new faces and that the neural correlates are evident in a distributed response across a network of regions beyond the visual brain.

SUNDAY, MAY 19, 2:45 – 6:45 PM, PAVILION

Visual Memory: Working memory and neural mechanisms, models, decision making

36.431 "REPULSIVE-FOLLOWED-BY-ATTRACTIVE" PAST-PRESENT NEURAL INTERACTIONS UNDERLIE SERIAL DEPENDENCE

Huihui Zhang[1,2,3,4] (zhanghuihui.cn@gmail.com), Minghao Luo[1,2,3], Huan Luo[1,2,3]; [1]School of Psychological and Cognitive Sciences, Peking University, [2]Beijing Key Laboratory of Behavior and Mental Health, Peking University, [3]PKU-IDG/McGovern Institute for Brain Research, Peking University, [4]Peking-Tsinghua Center for Life Sciences, Peking University

Current perception tends to be spontaneously influenced by the previous trial, namely serial dependence. This has been posited to arise from a perceptual temporal continuity field or a two-stage process. Hence, examining how past reactivation interacts with current information throughout various stages within each trial, i.e., encoding and decision-making, is essential to tackling the question. Here we performed two spatial perception tasks with electroencephalography (EEG) and magnetoencephalography (MEG) recordings. In Experiment 1, participants memorized the location of a dot within a 2-D continuous space (encoding stage) and reproduced it later (decision-making stage). In Experiment 2 with attentional modulation added, participants memorized locations of two dots and recalled the cued dot location later. Behaviorally, both experiments showed attractive serial bias, i.e., spatial perception is biased toward the previously reported location. Importantly, past-trial reactivation co-occurs with current-trial information during both the encoding and decision-making stages, signifying past-present interactions. Most interestingly, the past-present neural interactions exhibit a two-stage dynamic profile: repulsive interactions during encoding and attractive interactions during decision-making, arising in the visual cortex and orbitofrontal cortex (OFC), respectively. Finally, only the late attractive interaction is modulated by attention and correlates with serial bias behavior, while the early repulsive interaction is task-irrelevant. Overall, our study provides novel neural evidence supporting that serial dependence involves a repulsive-followed-by-attractive two-stage process, wherein past information first repulses present processing during sensory encoding in a task-irrelevant way and is then integrated with it in OFC based on task modulation during decision-making.

36.432 SERIAL DEPENDENCE IN VISUAL WORKING MEMORY COULD IMPROVE REPRESENTATIONAL PRECISION WITHOUT OUR AWARENESS

Jenny W.S. Chiu[1] (jennychiuws@gmail.com), Alan L.F. Lee[1]; [1]Lingnan University, Hong Kong

Our perception of a visual stimulus tends to depend on previously presented ones, a phenomenon known as serial dependence. Previous studies have shown that confidence judgments on the preceding stimulus can influence the magnitude of serial-dependent bias (e.g., Samaha et al., 2019), but it remains unclear whether we are metacognitively sensitive to perceptual effects arising from serial dependence. In the present study, we investigated the metacognitive sensitivity of serial dependence effect in a visual working memory task. In each trial, one or six Gabor patches (as a between-subjects factor) were briefly presented at locations around 5-6 degrees away from fixation. After a delay with a randomly-sampled duration (normally-distributed with mean = 2,000ms, SD = 500ms, limited within 250 - 3,750ms), a cue would appear at one of the Gabor's locations. Participants first reproduced the orientation of the Gabor patch presented at the cued location by adjusting the orientation of an on-screen line, and then rated their confidence on the orientation response using a slider bar. We manipulated the relative difference in orientation between every two consecutive trials such that there were 11 possible relative differences (inclusively from -75 to +75 degrees, 15 degrees apart). When there was only one Gabor presented (i.e., a set size of 1), we found that precision in the reported orientations (measured as the inverse of across-trials response variability) was significantly higher when the target orientation was the same as the one in the preceding trial. Interestingly, confidence ratings remained constant across values of relative differences. In a separate experiment with a set size of 6, we did not find any serial dependence effects on either first-order or second-order responses. Our findings suggest that, while serial stability could improve the precision of stimulus representation, we are metacognitively insensitive to such improvement.

This work was supported by the Research Matching Grant Scheme (LU Fund Code: 185218) from the University Grants Committee of the Hong Kong SAR

36.433 VISUAL WORKING MEMORY LOAD PERSISTS DURING THE COMPARISON PHASE

Chong Zhao[1] (chongzhao@uchicago.edu), Temilade Adekoya[1], Sintra Horwitz[1], Edward Awh[1], Edward Vogel[1]; [1]University of Chicago

Working memory is often measured by presenting arrays of visual items to be remembered over a short delay that must be compared with a test display. Extensive work has elucidated the neural mechanisms that support the encoding and maintenance periods of these tasks. Though, little is still understood about the comparison process itself. For example, the contralateral delay activity (CDA) is a sustained EEG component that provides a sensitive measure of the current working memory load during the retention period. However, it is not clear whether this activity continues to track the full working memory load during the comparison phase of the task or if it is reduced to just the one item from the array that is being tested. In Experiment 1, we used a change localization task with 2-item and 4-item arrays of colors. At test, subjects had to report which item changed. We observed that the response-phase CDA for set size 4 was significantly larger than the CDA for set size 2. In Experiment 2 we used a single-probe change detection design, in which only a single item from the original array was shown at test. Despite only a single item shown on the screen, we again observed that the response-phase CDA for set size 4 was still significantly larger than the CDA for set size 2. Our results suggest that the working memory load during test reflects the load from all of the items from the array that were stored.

We acknowledge the funding from National Institute of Mental Health (grant ROIMH087214); Office of Naval Research (grant N00014-12-1-0972).

36.434 FLEXIBLE ALLOCATION OF VISUAL SELECTION AND ACTION PLANNING DURING VISUAL WORKING MEMORY

Rose Nasrawi[1] (r.nasrawi@vu.nl), Freek van Ede[1]; [1]Institute for Brain and Behavior Amsterdam, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam

Visual information maintained in working memory can be used to plan for and guide upcoming behavior. Yet, in daily life, the order in which visual information is encoded into working memory does not always correspond with the order in which this information will become relevant for behavior: visual information that is encoded early might only become relevant for behavior later, and vice versa. We asked how the dynamics of visual selection and planning for prospective memory-guided behavior depend on the order in which visual information is anticipated to become relevant for behavior. To address this, we developed a visual-motor working-memory task in which participants were asked to sequentially encode and memorize two colored tilted bars, that were reproduced consecutively following two separate response cues. The order in which the items needed to be reproduced could always be anticipated by the participant, and varied between trials: the cueing-order could either be the same as the encoding-order (first item cued for report first), or be reversed (first item cued for report second). We used EEG time-frequency analyses to track the selection of visual information (item location) and action planning (prospective response hand, linked to item tilt) during working memory. We show that visual selection of both items is largely unaffected by the order in which they are expected to become relevant for behavior. In contrast, signatures of action planning exhibit a clear prioritization of the action that will become relevant for behavior first. Furthermore, they show how planning to act on one memory item can take place alongside visual encoding of another item. These results reveal that visual encoding and planning for prospective manual actions can be decoupled, whereby the brain can flexibly encode new visual information into working memory, while concurrently planning to guide behavior using previously encoded visual information.

This research was supported by an ERC Starting Grant from the European Research Council (MEMTICIPATION, 850636) and an NWO Vidi grant by the Dutch Research Council (grant number 14721)

36.435 THE DYNAMICS OF VISUAL WORKING-MEMORY PARTIAL-UPDATE DEPENDS ON SET-SIZE AND SPATIAL LAYOUT

Shaoying Wang[1] (shaoying.wang@du.edu), Srimant Tripathy[2], Haluk Ogmen[1]; [1]University of Denver, [2]University of Bradford

Working memory (WM) has limited capacity, which necessitates an efficient update process. Previous studies showed that updating WM involves an active removal process whereby information that is no longer relevant is discarded to make room for new information (Ecker et al, 2014a, 2014b). We focused on determining the dynamics of this active removal and updating process, particularly in cases of partial modifications within WM. We used a modified version of Ecker's (2014b) experimental design. In Experiment 1, the stimuli consisted of three random letters placed in boxes with black boundaries. Before new letters appeared, some of these box boundaries turned red to indicate which letters were to be updated. The duration between the onset of this "removal cue" and the appearance of the new letter(s) varied from 50 to 1500 ms. After this delay, only the new, updated letters were displayed to ensure a partial update. Subjects pressed the space-bar to indicate when they finished their update. Results showed that, as the duration of the removal cue increased, the reaction time for updating decreased. This suggests that subjects were actively removing information from their WM. The time required to remove one letter was similar to that required for two letters. To investigate the relationship between WM capacity and the removal/update dynamics, Experiment 2 increased set-size to four, while keeping all other parameters unchanged. Our findings revealed that set-size and the spatial layout of the stimuli influenced the removal/update dynamics. We found that the time required for WM content removal and update is influenced by two factors: 1) the presence of a spatial gap, i.e. an unchanged letter, which may prevent the grouping/chunking of to be removed/updated items leads to longer removal/update times; and 2) under similar gap conditions, updating the leftmost letter takes more time compared to the rightmost letter.

36.436 THE OBLIQUE EFFECT IN VISUAL WORKING MEMORY IS ENHANCED BY DISTRACTION, REGARDLESS OF TDCS MANIPULATIONS

Harun Yörük[1] (harunyoruk42@gmail.com), Benjamin J. Tamber-Rosenau[1]; [1]University of Houston

The sensory recruitment model of visual working memory (VWM) suggests that VWM representations depend on the early visual cortex (EVC). However, the model has been challenged by the logic that ongoing perception should interfere with EVC VWM representations. Moreover, previous research showed that the contents of VWM can be decoded from the parietal cortex regardless of distractor presence but EVC decoding was reduced or absent if visual distractors were presented during the memory delay (Bettencourt & Xu, 2016, doi:10.1038/nn.4174; Rademaker et al., 2019, doi:10.1038/s41593-019-0428-x). Here, we adopted two manipulations to evaluate VWM reliance on EVC vs. parietal cortex. First, we applied anodal (or sham) transcranial direct current stimulation (tDCS) to either occipital cortex (Oz) or right posterior parietal cortex (P4) prior to a Gabor patch orientation VWM task. We predicted distinct effects of EVC and parietal tDCS on overall VWM performance. Second, we displayed distractors during the delay period in half of the testing blocks. We reasoned that the oblique effect, which is thought to rely in part on EVC (Berkley et al., 1975, doi:10.1016/0042-6989(75)90213-8), would be reduced by distractor presence and the accompanying shift to parietal mechanisms. Contrary to our predictions, neither tDCS manipulation was effective. The oblique effect was present under all conditions, and surprisingly, it was stronger under visual distraction. To the extent that the oblique effect indeed has its basis in EVC, these results suggest paradoxically increased dependence on EVC when memories are maintained through incoming visual stimuli. Alternatively, such an effect would be consistent with the loss of VWM detail during distraction, or these results could suggest that the oblique effect is less dependent on EVC than previously believed.

This material was supported by the United States National Science Foundation under grant number 2127822 and Sigma Xi Grants in Aid of Research number 4213

36.437 HIGH RESOLUTION IMAGING OF THE HUMAN PREFRONTAL CORTEX DURING WORKING MEMORY

Zhengang Lu[1] (zl451@nyu.edu), Logan Dowdle[2], Kendrick Kay[2], Clayton Curtis[1]; [1]New York University, [2]University of Minnesota

Working memory (WM) increases the duration with which stimulus representations are available for further processing. The canonical theory of WM posits that persistent activity in neurons in lateral prefrontal cortex (PFC) stores WM representations. While electrophysiological data from nonhuman primate studies supports this hypothesis, data from fMRI studies have largely failed to find persistent activity in human lateral PFC. One possible explanation for this puzzling discrepancy stems from recent analyses of the anatomical distribution of neurons tuned to stimulus features across macaque PFC (Miller et al., 2022; Xiang et al., 2023). Namely, the resolution of previous fMRI measurements may be too coarse relative to the fine-grained spatial scales of tuned neurons. To address this potential limitation, we used high-resolution (900 micron isotropic voxels) fMRI at 7T to measure delay period activity during a memory-guided saccade task within a partial slab covering lateral PFC (N=5;

TR 2100 ms, 60 slices, partial Fourier 6/8, in-plane acceleration 2, multiband acceleration 2). Custom pre-processing methods were implemented to achieve high spatial accuracy, including EPI undistortion, head motion correction, and run-wise nonlinear anatomical coregistration. We also performed population receptive field mapping to identify topographically organized visual field maps and spatially selective voxels. In our preliminary analyses, we observed persistent activity in several areas of the PFC. Moreover, in some voxels this activity contained information about the location of the memoranda. These results suggest that high-field fMRI with sub-millimeter spatial resolution can bridge the gap between nonhuman primate neurophysiology and human neuroimaging research.

This work was funded by the US National Institutes of Health (R01 EY034118).

36.438 INTERACTIONS BETWEEN VISUAL WORKING MEMORY AND FACE PERCEPTION

Sanika Paranjape[1] (sanikap@gmail.com), Sarah Shomstein[1], Dwight Kravitz[1]; [1]The George Washington University

Visual working memory (VWM) is a process by which visual information is temporarily maintained and manipulated. The sensorimotor recruitment model posits a shared recruitment of the sensorimotor cortices by both perception and VWM. Previous studies show that this bidirectional relationship exists for simple and low-level features, like orientation, color, and motion. However, few studies investigated whether this bidirectional relationship is unique to low-level features or extends to high-level and complex stimuli, like faces. We hypothesized that faces held in VWM influence perception of faces we see later, consistent with the common recruitment of face-selective regions for VWM and perception. In the experiment, two discrimination conditions are shown to the participants: middle and side. In the middle condition, two discrimination stimuli are equidistant from the maintained face image, whereas in the side condition, they are both off to one side. We hypothesized and found that in the middle condition, both discrimination faces would be pulled inwards towards the maintained face, reducing the perceived difference in the two faces, resulting in a higher discrimination threshold. In the side condition, the closer discrimination face would be more biased towards the maintained face, resulting in a greater perceived difference and lower thresholds. Further, based on the continuous report, the discrimination faces in the side condition introduced a systematic bias towards themselves, demonstrating directional interference between perception and the face held in VWM. We also established the specificity of these effects to actively maintained information, by asking participants to maintain the orientation of a gabor patch in the center and then make a same/different judgment on the faces, which eliminated any difference in the discrimination thresholds between the middle and side conditions. Overall, this bidirectional interference suggests that face-selective areas are involved in VWM maintenance, further eroding the distinction between supposed perceptual and cognitive processes.

NSF BCS-2022572 to SSS and DJK

36.439 THE ART OF REMEMBERING: APPLYING VISUAL DESIGN PRINCIPLES TO MEMORY

Mohsen Rafiei[1] (moo.rafiei@gmail.com), Bahareh Jozranjbar[1]; [1]Providence College, Psychology Department, [2]Providence College, Psychology Department

This study delves into the rule of thirds, a pivotal principle in visual arts and design, assessing its application in enhancing memory recall and accuracy. We investigated if visual stimuli alignment with rule of thirds (ROT) points within a visual field could optimize memory performance, a question of significant interest in the field of visual perception. Participants completed an odd-one-out visual search task featuring a 5x5 matrix of lines, each with a distinctively oriented target line. The matrix incorporated 4 ROT lines at key intersection points, 4 Non-ROT lines centrally placed at similar distances as ROT lines, 16 Far lines at the margins, and 1 central line, with slight jitter added. The focus was on comparing the accuracy of memory recall for target orientations across these spatial positions. The results demonstrated that positioning lines at ROT points significantly enhanced the accuracy of reported orientations compared to other screen locations, specifically non-ROT, indicating an advantage in spatial memory. Reaction times (RTs) were similar between ROT and Non-ROT points, suggesting the improvement in memory accuracy was not due to faster processing. However, both ROT and Non-ROT locations yielded faster RTs than Far locations, indicating central positioning's role in expedited response. Our research underscores the vital impact of spatial positioning on visual memory, particularly emphasizing the rule of thirds as an influential factor in visual perception and human-computer interaction. By strategically aligning items with the rule of thirds, we observed a notable enhancement in memory accuracy. This finding has considerable implications for the design of user interfaces and educational materials, moving beyond mere aesthetics to practical applications in spatial memory. Such insights pave the way for further HCI research, aimed at refining visual layouts to boost user engagement and learning efficiency. Keywords: Visual attention, working memory, Rule-of-thirds, human computer interaction

36.440 ATTRACTOR DYNAMICS IN MORPH-SHAPE WORKING MEMORY IN MACAQUE PREFRONTAL CORTEX

Wenhao Dang[1] (wenhao.dang@vanderbilt.edu), Rye Jaffe[2], Christos Constantinidis[1]; [1]Vanderbilt University, [2]Wake Forest University

The neural basis of working memory has been an important topic in current research. The 'bump attractor' model describes a network where structured recurrent excitation and broader feedback inhibition maintain information via persistent neural population codes. This model successfully explains behavioral variations and many neural activity features. However, while the neural basis of working memory in nonhuman primates has been studied extensively with spatial working memory tasks, the neural mechanisms of object working memory remain neglected. We thus trained two macaque monkeys to perform a delayed match-to-sample task using morphed object silhouettes, with five levels (0%-40% morphing with 10% steps between neighboring difficulty levels) of similarity between the target and the distractor. In 22 recording sessions, monkeys were able to perform the task well above chance level for all morphing axes we used (mean performance was 86%, 83%, 77%, 66%, and 54% for 0%, 10%, 20%, 30%, and 40% morphing respectively). We used multi-

contact linear probes to record single-neuron activity from the prefrontal cortex as monkeys performed the task. Out of 423 cells recorded, we found 42 that were selective to shape during the delay period. In these cells, we analyzed ones (n=28) with enough error trials when presented with their selective shape pair. We found that decreased persistent activity (comparing average firing rates in a 700ms window), was present in correct compared to error trials following the presentation of a preferred stimulus, paired t-test,t(27)=2.38, p=0.025). Conversely, increased persistent activity following the presentation of a non-preferred stimulus was more likely to result in errors (paired t-test,t(27)=2.08, p=0.047). These results are consistent with the idea that object working memory is being maintained in the prefrontal cortex and is predictive of performance in shape working memory tasks, just as prior research has already demonstrated in spatial tasks.

Research reported in this abstract was supported by the National Institutes of Health National Eye Institute under award numbers R01 EY017077

36.441 EEG-BASED DECODING OF SHAPES AND THEIR CATEGORIES IN VISUAL WORKING MEMORY

Frida Printzlau[1,2] (fridaprint@gmail.com), Olya Bulatova[2], Michael Mack[1], Keisuke Fukuda[1,2]; [1]University of Toronto, [2]University of Toronto Mississauga

Visual working memory (VWM) allows us to store information in a highly accessible format for an upcoming task. Traditionally, VWM studies require participants to keep a precise copy of a stimulus in mind. But in the real world, we might need to store the same information for different types of tasks, such as recognition or categorisation judgements. For example, when deciding if a bike is the exact model you want, or the same brand. In this study, we asked how categorisation modulates VWM representations. Participants first learned to group unfamiliar shapes from the Validated Circular Shape (VCS) space (Li et al., 2020) into two categories based on their visual features. They then completed a shape VWM task that either required delayed match-to-sample or delayed match-to-category judgements on different blocks while we collected electroencephalography (EEG) data. We tracked the emergence of stimulus-, category- and task level information with high temporal resolution using multivariate pattern analyses of EEG. The neural activity pattern over posterior electrodes contained information about the memorised shape for about one second following VWM encoding. Initially, the stimulus code overlapped across the two tasks, but quickly separated according to task. Later in the delay, stimulus coding persisted only for the match-to-category task and was accompanied by a neural category signal, indicating that categorisation may require an active stimulus representation. To our knowledge, this is the first illustration that the VCS space is decodable from EEG, preserving the circular similarity structure. This provides a fruitful avenue for researchers looking to characterise neural representations of unfamiliar visual stimuli with high temporal resolution. The results of this study will help elucidate the neural mechanisms supporting VWM under different task demands.

This research was funded by the University of Toronto Faculty of Arts and Sciences Postdoctoral Fellowship Award to FP; the Natural Sciences and Engineering Research Council (KF and MM); the Connaught New Researcher Award to KL; Brain Canada Future Leaders in Canadian Brain Research Grant to MM.

SUNDAY, MAY 19, 2:45 – 6:45 PM, PAVILION

Visual Memory: Imagery

36.442 THE PUPIL AND THE MIND'S EYE: PORTABLE PUPILLOMETRY CAPTURES ROBUST RESPONSES TO IMAGINARY LIGHT

Michael McPhee[1] (mcphee.mi@northeastern.edu), Morgan McCarty[1], Michael Young[2,3], Jorge Morales[1]; [1]Northeastern University, [2]Massachusetts General Hospital, [3]Harvard Medical School

The pupillary response is driven by a range of factors, most notably by retinal illumination and arousal. Strikingly, the pupil's diameter also adjusts in reaction to endogenous perceptual features, such as the brightness of imagined objects, the brightness or darkness evoked by certain words, and scene interpretation. Pupil size is, moreover, affected by all manner of cognitive processes, including pattern detection, mental effort, attention, and deception. Reliable measurement of pupil changes, however, usually requires tightly controlled conditions – a dark room lacking exogenous variable light sources is the gold standard. Such requirements effectively limit pupillometry to the laboratory, precluding its use with certain populations (e.g., hospitalized patients) as well as in public spaces like schools or parks. Here, using a commercially available head-mounted display (HMD) and an internally mounted eye tracker, we replicate results from two different groups demonstrating the pupil's sensitivity to imaginary light (Laeng & Sulutvedt, Psychol Sci, 2014; Kay et al., eLife, 2022). In each trial, subjects were shown triangles with four possible luminance levels and four possible orientations against a mid-gray background. This perceptual phase was immediately followed by a long dark-adaptation period. Then, while staring at a gray screen, subjects were asked to vividly imagine the exact shape they had previously seen (with the same brightness and orientation). In line with previous findings, we show that subjects' pupillary response is proportional not only to the luminance of the perceived shapes, but also to the brightness of their imagined counterparts. Our use of a portable HMD to limit the impact of ambient lighting conditions demonstrates the feasibility of conducting sophisticated pupillometry outside of the laboratory. If the subject cannot come to the darkness, the darkness must go to the subject.

This material is based on work supported by Templeton World Charity Foundation, Inc. under Prime Award no. TWCF0495 and The Regents of the University of California.

36.443 PUPILLARY RESPONSE IN VISUAL IMAGERY

Sin Wan Sharon Hon[1] (swhon@connect.hku.hk), Sing Hang Cheung[2]; [1]The University of Hong Kong

The extent and nature of the overlap between visual imagery and visual perception have been debated over the past century. Can visual imagery result in presumably automatic physiological response such as the pupillary light reflex (PLR)? Laeng & Sulutvedt (2014) reported

pupillary responses to dark and bright imagined scenarios. Based on such findings, Kay, Keogh, and Pearson (2022) proposed using the magnitude of imagery-induced PLR as a measure of the ability to generated vivid imagery. We aimed to replicate Kay et al.'s (2022) findings on the PLR response in visual imagery. Ninety-five normally sighted participants were asked to view 16 stimuli in four luminance levels, and then imagine the previously seen stimulus. Pupillary responses were measured during both the perception and imagery periods. PLR response was examined by comparing the pupil diameter in the two darker luminance conditions against that in the two brighter conditions. PLR response was statistically significant in both the perception and the imagery periods (perception: $F(1,94) = 598$, $p < .001$; imagery: $F(1,94) = 14.7$, $p < .001$). Statistically significant bivariate correlations were consistently observed among the self-report questionnaires (VVIQ, OSIVQ, and SUIS) and the trial-by-trial vividness ratings, suggesting a shared mechanism underlying the subjective evaluation of imagery vividness. However, we could not replicate Kay et al.'s (2022) findings on the association between trial-by-trial vividness ratings and the magnitude of PLR response during the imagery periods. Subjective reports of imagery vividness could reflect both the ability to generate vivid imagery and the metacognitive evaluation of such ability.

36.444 ASSESSING VISUAL MENTAL IMAGERY IN CEREBRAL VISUAL IMPAIRMENT

Claire E. Manley[1] (cemanley@meei.harvard.edu), Madeleine Heynen[1], John Ravenscroft[2], Lotfi B. Merabet[1]; [1]Massachusetts Eye and Ear, Harvard Medical School, [2]Moray House School of Education and Sport, The University of Edinburgh

Cerebral (cortical) visual impairment (CVI) is an umbrella term for visual disorders associated with maldevelopment of or damage to retrochiasmal visual processing areas of the brain, often (but not necessarily) in the absence of major ocular disease. Individuals with CVI may show lower-level visual function deficits, such as reduced acuity, contrast sensitivity, visual field impairment, and ocular-motor disorders. However, higher-order perceptual deficits, such as impaired visual search and image recognition, are also common. One aspect that has not been investigated is whether visual mental imagery is also impaired in these individuals. We used the Vividness of Visual Imagery Questionnaire (VVIQ; Marks 1973) to investigate differences in visual imagery abilities in individuals with CVI compared to age-matched controls. VVIQ and verbal IQ scores (WAIS-IV) were collected from 12 individuals with CVI (3 males, mean=22.58 years ± 5.32 SD) and 12 controls with neurotypical development (4 males, mean=22.42 years ± 3.35 SD). There was no significant difference in age [$t(18.952)=0.091$, $p=0.929$, $d=0.037$] or verbal IQ [meanCVI=108.50 ± 29.266 SD, meanControls=119.73 ± 13.425 SD, $t(12.365)=-1.111$, $p=0.288$, $d=-0.502$] between the two groups. However, we found that individuals with CVI had significantly lower VVIQ scores (mean=34.42 ± 18.57 SD) compared to controls (mean=63.83 ± 14.30 SD) [$t(20.650)=-4.347$, $p=0.0003$, $d=-1.775$]. We also found that verbal IQ was not significantly predictive of VVIQ in either group (CVI: [$F(1, 8)=0.090$, $p=0.772$, $R2=0.013$], Control: [$F(1, 10)=0.101$, $p=0.294$, $R2=0.011$]), indicating that reported visual imagery in CVI did not vary as a function of verbal ability. These results suggest that impaired visual mental imagery may also be a feature related to the complex clinical profile of CVI.

36.445 CAN WE PREDICT VIVIDNESS FROM THE CHARACTERISTICS OF IMAGINED IMAGES? A NOVEL DATABASE FEATURING VIVIDNESS JUDGMENTS OF THE NATURAL SCENE DATASET

Catherine Landry[1], Jasper JF van den Bosch[2], Vincent Taschereau-Dumouchel[3,4], Frédéric Gosselin[1], Ian Charest[1]; [1]Cerebrum, Université de Montréal, CA, [2]School of Psychology, University of Leeds, UK, [3]Département de psychiatrie et d'addictologie, Université de Montréal, CA, [4]Centre de recherche de l'Institut Universitaire en Santé Mentale de Montréal, CA

What makes an image easy to imagine? Previous research on mental imagery mostly used very limited samples, which prevented establishing a direct association between image characteristics and vividness ratings (i.e., the clarity and detail level of imagined images). Here, we present a large-scale database of vividness judgments associated with the natural scenes from the Natural Scenes Dataset (NSD; Allen et al., 2021), which consists of 73,000 annotated natural images. During each trial, participants sequentially view two NSD images and are then randomly asked to imagine one or the other (i.e., retro-cued target image) for 4 s. Participants then rate the vividness of their mental image (on a continuous scale from 0 to 100), followed by a test to ensure that they imagined the correct target. Participants (n = 1825), recruited from Prolific, are directed to the Meadows platform for online experiments. Each complete 120 trials of our vividness task, for a total of 219,000 vividness ratings across participants. They also complete the Vividness of Visual Imagery Questionnaire (Marks, 1973) to measure their visual imagery ability. Overall, preliminary data reveals excellent performance on the task (average accuracy of 95.83% correct target identification), as well as substantial interimage (M = 0.65 ± 0.25) and interindividual (M = 0.66 ± 0.17) variability in average vividness scores. This large-scale dataset of vividness ratings will offer invaluable insights into visual imagery by allowing to train predictive models of subjective imagery experiences. It will enable a deeper understanding of the visual and cognitive factors influencing mental image vividness, and serve as a guiding resource for future experiments. Furthermore, integrating these vividness judgments with neural data from the NSD will allow for an exploration of the relationship between subjective experience and objective brain responses.

36.446 DECODING SOUND CONTENT IN EARLY VISUAL CORTEX OF APHANTASIC INDIVIDUALS

Belen M. Montabes de la Cruz[1] (b.montabes-de-la-cruz.1@research.gla.ac.uk), Clement Abbatecola[1,2], Roberto S. Luciani[1,3], Angus T. Paton[1,2], Johanna Bergmann[4], Petra Vetter[5], Lucy S. Petro[1,2], Lars Muckli [1,2]; [1]Centre for Cognitive Neuroimaging, University of Glasgow, [2]Imaging Centre for Excellence (ICE), University of Glasgow, [3]School of Computing Science, University of Glasgow, [4]Max Planck Institute for Human Cognitive and Brain Sciences, [5]Department of Psychology, University of Fribourg

The retinotopic visual cortex of blindfolded and congenitally blind participants responds with decodable activity patterns to the sound of natural scenes (Vetter et al., 2014, 2020). Decoding accuracies increase from foveal to far peripheral retinotopic regions in V1 and V2, with stronger effects in the blind (V1, V2 & V3). There are direct projections from the auditory cortex to the peripheral visual cortex in non-human primates (Falchier et al., 2002) and we assume that in humans these projections are also more predominant in the periphery. In congenitally blind participants, visual feedforward processing is absent and top-down projections to the visual cortex might proliferate. In contrast, in aphantasia top-down projections might be less effective, leading to a loss of visual imagery experience. How about other forms of top-down projections? Is auditory top-down influence reduced in aphantasia? We presented auditory stimuli to 22 blindfolded aphantasic participants. We found that sound could only be decoded in foveal V3, reporting reduced decoding in V1 and V2 in aphantasics compared to controls and blind participants. When fitting a linear mixed effect model on data from the control, blind and aphantasic groups, we found a significant interaction effect between group and eccentricity and between area and eccentricity: in V1, the relative increase of decoding with eccentricity was strong only in the blind participants. The differences between groups increased in V2 and V3, with the largest differences being reported in V3, where decoding increased by eccentricity for the blind, did not change for controls, and decreased for aphantasic participants. Contrary to the enhanced feedback observed in blind participants, our findings suggest that the feedback of auditory content to the early visual cortex may be reduced in aphantasic participants. Reduced top-down projections might lead to both less decoding of sounds and reduced subjective experience of visual imagery.

The European Union's Horizon 2020 Framework Programme for Research and Innovation SGA 945539, the Biotechnology and Biological Science research Council (BBSRC) BB/V010956/1, and the college of Medicine, Veterinary and Life Sciences Doctoral Training Programme (MVLS DTP) at the University of Glasgow.

36.447 NEURAL ENTRAINMENT TO IMAGINED RHYTHMS IN INDIVIDUALS WITH PROFICIENT IMAGERY ABILITY

Fang Jiang[1] (fangj@unr.edu), Benjamin Sreenan[1], Canhuang Luo[1], Simon Whitton[1]; [1]University of Nevada, Reno

Sensorimotor synchronization (SMS) involves the temporal coordination of motor movement with an external or imagined stimulus. Finger-tapping studies indicate better SMS performance with auditory or tactile stimuli compared to visual. However, the ability to synchronize with a visual rhythm can be improved by enriching stimulus properties (e.g., spatiotemporal content) or as a result of individual differences (e.g., one's vividness of auditory imagery). We previously used a finger-tapping synchronization-continuation paradigm and showed that higher self-reported vividness of auditory imagery led to more consistent SMS when participants continued without a guiding visual rhythm. Here, we further examined the contribution of imagery to SMS performance by testing proficient imagers and including auditory or visual distractors during the continuation phase. While visual distractors had minimal effect, SMS consistency was significantly worse when auditory distractors were presented. Analysis of electroencephalography (EEG) revealed neural entrainment at the beat-related frequency, but only when visual or auditory distractors were present. Neural entrainment to the beat frequency positively correlated with SMS consistency in the presence of visual distractors, suggesting the potential utilization of auditory imagery and its role in supporting SMS performance. During continuation with auditory distractors, the neural entrainment showed an occipital electrode distribution suggesting the involvement of visual imagery. Unique to SMS continuation with auditory distractors, neural and sub-vocal (measured with electromyography) entrainment were found at the three-beat pattern frequency. In this most difficult condition, proficient imagers employed both beat- and pattern-related imagery strategies. However, this combination was not enough to restore SMS consistency to the same level observed with visual or no distractors. Combined, our results suggest that proficient imagers effectively utilized beat-related imagery in one modality when imagery in another modality was limited.

National Institute of Health (P20GM103650)

36.448 VISUAL CORTICAL CONTRIBUTIONS TO EMOTIONAL IMAGERY

Lily Cao[1], Lihan Cui[1], Maeve Boylan[1], Andreas Keil[1], Mingzhou Ding[1]; [1]University of Florida

Previous work has shown that mental imagery activates the visual cortex, including the primary visual cortex. Here, we examined the role of the visual cortex in representing the motivational significance of imagined events by recording fMRI data in an emotional imagery paradigm. In each trial, the participant began by silently reading a text varying in affective content (pleasant, neutral, or unpleasant) and then proceeded to imagine their active involvement in the described event continuously until the end of the trial. Applying both univariate and multivariate analysis methods to the fMRI data, we found that (1) the univariate analysis did not demonstrate valence-dependent activation in the visual cortex, which is consistent with findings from previous studies and (2) the MVPA analysis revealed above-chance decoding accuracy for pleasant vs neutral, unpleasant vs neutral, and pleasant vs unpleasant imagery in the entire visual hierarchy including the primary visual cortex. These results show that the visual cortex contributes to the representation of motivational significance of imagined events and multivariate analysis is essential in uncovering the valence-specific neural patterns in the visual cortex.

36.449 UNVEILING MENTAL IMAGERY: ENHANCED MENTAL IMAGES RECONSTRUCTION USING EEG AND THE BUBBLES METHOD

Audrey Lamy-Proulx[1] (audrey.lamy-proulx@umontreal.ca), Jasper van den Bosch[2], Catherine Landry[1], Peter Brotherwood[1], Vincent Taschereau-Dumouchel[3,4], Frédéric Gosselin[1], Ian Charest[1,2]; [1]Cerebrum, Département de psychologie, Université de Montréal, [2]School of Psychology, University of Leeds, England, [3]Département de psychiatrie et d'addictologie, Université de Montréal, [4]Centre de Recherche de l'Institut Universitaire en Santé Mentale de Montréal

The exact nature of the visual features that are brought to consciousness when one is engaging in mental imagery is still difficult to study empirically. The few studies that have attempted to reconstruct mental images obtained poor quality results due to a poor sampling of the "scene space". The aim of our study was to reconstruct better quality mental images using electroencephalography (EEG) and the Bubbles method, a technique that randomly samples visual information in an image. We hypothesize that the reconstructed mental images would reveal key visual features of the images and that verbal instructions (e.g., imagine the man and not the car in the image) could modulate the reconstructed image. We recorded the brain activity of participants (preliminary sample: N = 7, 4 males, mean age = 22.4) during two alternating tasks divided into 6 two-hour sessions. In the perception task, participants were presented with two images through different sets of randomly located Gaussian apertures or "bubble masks" (1,500 trials per image). In the mental imagery task, participants were shown the two stimuli successively and asked to imagine the first or the second one, in its entirety or in part (450 trials per image, including ⅓ object-specific trials). For each participant and for each image, we correlated the EEG activity patterns between the mental imagery and visual perception tasks. The bubbles masks, weighted by corresponding correlation coefficients, were then summed to generate "classification images" of mental images. Comparing these classification images between the object-specific imagery trials, we found that the content of mental images could, indeed, be modulated by instructions for some participants. This study not only contributes to the understanding of the neural mechanisms underlying imagery, but also offers a promising avenue for optimizing the communication methods through brain-computer interfaces.

36.450 BUILDING COMPOSITIONAL MEMORIES AND IMAGERY FROM DISENTANGLED LATENT SPACES IN AN AUTOENCODER

Brad Wyble[1] (bwyble@gmail.com), Ian Deal[1], Vijay Subedi[1]; [1]Penn State University

The visual system has a densely recursive hierarchy that integrates feedforward and feedback processing in the service of functions such as perception, attention, memory and imagery. This hierarchy presumably has the remarkable property that representations, whether perceived, or imagined, can be passed forward and backward to reach a given level of abstraction or dimensionality to allow different kinds of processing. We explore this idea using a generative, neurocomputational model called Memory for Latent Representations(MLR). First, an autoencoder is trained to disentangle features such as shape, color and location into distinct latent spaces. Then, a shared memory resource builds engrams from those latent spaces and binds them to tokens. By selecting which latent spaces are used for building these memories on-the-fly, an engram can focus on fine-grained visual details, compressed visual details, categorical codes, or any combination of these. Empirical demonstrations in human observers support the model by showing that engrams can be calibrated against task demands, to focus on the appropriate features and visual details. Through its decoder, MLR can reconstruct both recollected memories as well as arbitrary combinations of features according to top-down instructions, thereby providing an approximation of some aspects of compositional visual imagery. By combining a working memory system with an autoencoder, MLR provides a theoretical framework for understanding how visual memory and imagery work jointly to encode, decompose, modify and recode complex visual representations. In this expanded version of the MLR model we demonstrate the ability to separate location, color and visual form into disentangled latent spaces and then to modify and recombine those codes. Such codes can then be used in a generative fashion to create novel compositions of features according to top-down instructions. This work helps us to understand the mechanisms of visual imagery in a highly interpretable model context.

NSF grant 1734220

Sunday, May 19, 2:45 – 6:45 pm, Pavilion

Visual Memory: Working memory and objects, features

36.451 THE OBJECT AS THE UNIT FOR STATE SWITCHING IN VISUAL WORKING MEMORY

Shengnan Zhu[1] (shengnanzhupsy@gmail.com), Yongqi Li[1], Yingtao Fu[1], Jun Yin[2], Mowei Shen[1], Hui Chen[1]; [1]Zhejiang University, [2]Ningbo University

Visual working memory (VWM) efficiently maintains active representations and prevents interference by storing some information in a passive state. This study aims to investigate the unit responsible for switching representational states in VWM. The investigation focused on two opposing hypotheses: (a) the unit of switching as a feature (feature-based hypothesis), and (b) the unit of switching as an object (object-based hypothesis). Across three experiments involving 180 participants, participants were instructed to hold two features from either one or two objects in their VWM. In a retention-interval search task, participants searched for a tilted line among distractors and specified its direction. The memory-driven attentional capture effect of the first and second probed colors (indicated by a retro-cue) was examined. Results revealed that only the feature indicated to be probed first could elicit memory-related capture for the condition of separate objects. Importantly, features from an integrated object could guide attention regardless of the probe order. These consistent findings, observed across three experiments with features of different dimensions, same dimensions, or perceptual objects defined by Gestalt principles. They provide convergent evidence supporting the object-based hypothesis by indicating that features within a single object cannot exist in different states.

Science and Technology Innovation 2030-'Brain Science and Brain-like Research' Major Project (2022ZD0210800), National Natural Science Foundation of China (32171046)

36.452 CHUNKING AS AN OBJECT: WHAT COMES TOGETHER, GOES TOGETHER

Chattarin Poungtubtim[1] (chattarin2p@docchula.com), Chaipat Chunharas[1,2], Timothy Brady[3]; [1]Cognitive Clinical & Computational Neuroscience Lab, Faculty of Medicine, Chulalongkorn University, [2]Chula Neuroscience Center, King Chulalongkorn Memorial Hospital, [3]Department of Psychology, UCSD

When people have to remember several objects, they often employ grouping or chunking. The chunking does not only enhance memory precision but also introduces systematic biases for memorized items (e.g., attraction toward the gist of the display). While demonstrated for single-feature objects (e.g., orientation, shape, or color), it remains unclear how people chunk objects with multiple features. It could involve feature-level chunking, treating each feature separately, regardless of whether chunkable features belong to the same object. Alternatively, object-level chunking may occur, where grouping features within the same object improves chunking. Previous work investigated these hypotheses and found evidence of only feature-level chunking for multi-feature objects when both orientation and color were highly similar across all objects (Song & Chong, 2023). However, this work did not put feature and object cues into direct conflict to test their relative contributions. To test this, we ran a continuous report task where participants (N=44) had to remember 6 colored orientations and later report both the color and orientation of the probed item. In one condition, half of the stimuli were similar in colors but not orientations while the other half were similar in orientations but not colors. In another condition, half of the stimuli were similar in both colors and orientations, while the other 3 had random orientations and colors. Interestingly, when both features could be chunked for the same objects (i.e., the second condition), we found greater memory precision ($F(43) = 17.22$, $p < 0.001$) and stronger attraction bias ($F(43) = 9.58$, $p < 0.005$). Our results show that chunking can occur not only at a feature level but also at an object level. This prompts questions about the interaction of chunking at different levels and the potential existence of chunking at various levels of abstraction in working memory.

36.453 DISTRACTOR CONTROL FACILITATES AN INTEGRATION OF TARGET FEATURES IN VISUAL WORKING MEMORY

Hansol Lee[1] (jenyalee0401@gmail.com), Sun-Young Park[1], Jiyoun Jeong[1], Min-Suk Kang[1,2]; [1]Sungkyunkwan University, Seoul, Republic of Korea., [2]Center for Neuroscience and Imaging Research, Institute for Basic Science (IBS), Suwon, Republic of Korea

What is the unit of visual working memory (VWM) representation? Feature-based account posits that features comprising an object are represented independently, but object-based account posits that those features are represented as an integrated format. We sought to tease these two accounts apart by examining whether distractors produce feature specific disruption. If the features are stored independently, distraction against a specific feature should disrupt that feature of representation, keeping the other features intact. However, if the features are integrated in VWM, distraction against a single feature should disrupt the entire representation. Participants (N = 26) were informed to remember target colors in one of two visual hemifields for a short period of time and adjust the color of a single cued item. We manipulated distractor features (color and location) combinatorically, resulting in four trial types. In one, targets and distractors shared the location set (location-repeated) or not (location-distinct), and in the other, they shared the color set (color-repeated) or not (color-distinct). Further, we varied the distractor context from baseline to distraction block. In the baseline block, distractors were absent to establish the natural unit of VWM. The distraction block followed where the distractor-absent and the four distractor-present trials were given. We analyzed response errors of the distractor-present trials with a mixture of precision and guess rate over the two phases of the distraction block. In the first phase, immediately after the baseline block, we found a feature specific disruption. The guess rate was lower, but the precision was worse in the color-repeated than the color-distinct trials. On the other hand, the guess rate was higher, but the precision was superior in the location-repeated than the location-distinct trials. However, the feature-specific disruption becomes nearly absent in the late phase. This adaptive behavior suggests that the distractor control facilitates integrating features of VWM.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2022R1A2C2007363).

36.454 ICONIC MEMORY IS NOT A FEATURE SOUP

Mary Catington[1] (mfc120@msstate.edu), Saeideh Saffar[1], Michael Pratte[1]; [1]Mississippi State University

Visual sensory memory (also known as iconic memory) is typically characterized as a pre-attentive store, and attention is thought to move information from this rapidly decaying store into the more durable working memory system. If attention is the mechanism that binds visual features into coherent objects, as is assumed by prominent theories, then the contents of iconic memory should be comprised of a collection of visual features rather than bound objects. We recently found that iconic memory capacity for faces is no greater than that of working memory, which could reflect a failure to build holistic faces in the pre-attentive iconic store. In the following two experiments, recall performance was tested at retention intervals that measure iconic and working memory to directly determine whether errors in feature binding were more prevalent in iconic than working memory. In Experiment 1 participants reported the color and orientation of a cued line, and in Experiment 2 they reported the color and shape of a cued object. Standard analyses were applied to the joint distribution of color and orientation (or color and shape) to measure the rate of binding errors (e.g., accurately recalling the color of an object without any knowledge of its shape). In the working memory condition, results replicated previous findings that binding failures are somewhat common. Critically, binding failures were not more common in iconic memory than in working memory, as would be predicted if feature binding only occurred during an attentionally driven transfer into visual working memory. Therefore, we believe that iconic memory is not merely a soup of features that decay independently, and the binding failures observed in working memory are probably a limitation in perception rather than working memory storage.

This work was supported by National Institutes of Health Grant R15MH113075.

36.455 TASK-IRRELEVANT MOTION INFORMATION MODULATES WORKING MEMORY REPRESENTATIONS

Johan A. Gamba[1], Alexei Dawes[2], Hakwan Lau[2], Brian Odegaard[1]; [1]University of Florida., [2]Riken Center for Brain Sciences.

Previous studies have shown that working memory (WM) influences perception at both behavioral and neural levels (Kang et al., 2011, D'Esposito & Postle, 2015). For example, holding orientation information from a random dot kinematogram (RDK) in working memory influences subsequent perceptual judgments. However, most studies only require memorizing a single feature; whether multi-feature information held in WM influences perceptual decision-making, and vice-versa, is not fully understood. In this study, we explored both questions by exploiting the "motion repulsion" effect (Marshak & Sekuler, 1979). Thirty-one participants were presented with sequences of three RDKs, where the first (RDK1) and third (RDK3) shared two out of three properties (orientation, speed, or dot density). The task required subjects to report the orientation of the second RDK (RDK2), and also report (and rate confidence for) the property that differed between RDK1 and RDK3. Importantly, an irrelevant feature for the task—RDK2's speed—was systematically altered across three experiments (exp1: similar speed to RDK1, exp2: faster, exp3: slower). Here, we asked the following question: does manipulation of a task-irrelevant feature (RDK2's speed) influence the working memory judgment about RDK1 and RDK3? Our findings revealed that the motion repulsion effect, a perceptual phenomenon, was consistently observed at previously-reported angles (30° and 60°), despite the increased cognitive load of our task, which required maintaining multiple features in WM. Notably, in the experiment with the slowest irrelevant speed (exp3), task performance was adversely affected compared to the other conditions. Additionally, a unique pattern of metacognitive sensitivity emerged in the experiment with increased irrelevant speed (exp2): participants exhibited a higher proportion of "low confidence" responses for both correct and incorrect trials. Overall, these results demonstrate that motion repulsion occurs even when multiple features are held in working memory, and reveal an influence of task-irrelevant perceptual information on WM judgments.

J.A.G. was supported by funded internship at the RIKEN CBS summer program, where this research was initiated

36.456 OBJECT-LOCATION BINDING IN VISUAL WORKING MEMORY PREVENTS EFFECTIVE UPDATING

Eva Lout[1] (lout.2@osu.edu), William Narhi-Martinez[1], Blaire Dube[2], Julie Golomb[1]; [1]The Ohio State University, [2]Memorial University of Newfoundland

Despite the importance of updating visual working memory (VWM) representations in our dynamic visual world, memory performance has been shown to be reliably worse for spatially dynamic objects compared to static objects. This is possibly due to robust binding of object features to their original locations. Here, we conducted two experiments (preregistered) to test the effectiveness of intentionally updating object spatial information in a dynamic spatial context. Each trial began with the presentation of four different colored squares within white placeholders. In Experiment 1, subjects were instructed to maintain and update all four memory items; in Experiment 2, subjects were retroactively cued to maintain and update only two of the memory items, allowing subjects to drop the adjacent non-cued memory items. In half of the trials, during the subsequent memory delay, the placeholders would rotate one position clockwise or counterclockwise. The experiment consisted of two instruction blocks: Update (instructed to mentally update the items' locations when the placeholders rotated) and Ignore-Rotation (instructed to maintain the memory of the items in their original locations even when the placeholders rotated). At the end of the trial, participants were presented with a spatial probe and reported their memory of the color of the item in that location (original or updated, depending on instruction block) on a continuous color wheel. In Experiment 1, we observed large performance decrements when participants had to both intentionally update and, critically, attempt to ignore task-irrelevant rotations. In Experiment 2, even with a reduced load on VWM resources (only maintaining/updating two items), performance decrements persisted in both instruction contexts when the placeholders rotated. Analysis of systematic feature errors across both experiments revealed greater misreporting of nontargets when the display rotated, compared to static trials, regardless of instruction and even when those features were meant to be dropped from VWM.

NIH R01-EY025648 (JG), NSF 1848939 (JG), NSERC PDF (BD)

36.457 STRATEGICALLY CHOOSING BETWEEN RESETTING AND UPDATING PROCESSES IN VISUAL WORKING MEMORY.

Shani Friedman[1] (shanibuch@gmail.com), Roy Luria[2]; [1]Tel-Aviv University

Visual working memory (VWM) can hold representations of the objects in the world and modify them When they change. When such a change occurs, VWM can either update or reset its representation by relying on a pointer system associating each representation with the corresponding object in the environment. This study investigated the adaptability of the pointer system to different task contexts. Experiment 1 investigated whether VWM can adapt to a "resetting-mode", by performing resetting in a situation that previously triggered updating. We used the contralateral delay activity (CDA; an electrophysiological marker of VWM) as a marker of updating and resetting. Participants performed a shape change-detection task with polygons. In the One Polygon-Half condition, a single polygon-half appeared for 500ms, followed by a 50ms blank interval, and then the same polygon-half reappeared. In the Add condition, the original polygon-half reappeared together with a second polygon-half. In the Switch condition, the original polygon-half was abruptly replaced by an integrated polygon. Switch condition has been shown to trigger resetting in previous research. Importantly, 70% of the trials in the current experiment were Switch trials, encouraging resetting. The results indicated that VWM used resetting in the Switch condition (replication) but also in the Add condition, which suggests that VWM performed resetting in a condition that was known from a previous study to trigger updating. Experiment 2 used a similar design but now the Add condition appeared in 70% of the trials, to investigate whether VWM can adapt to an "updating-mode". The results showed evidence for resetting only in the Switch condition. This result indicated that VWM can't adapt to an updating mode. In conclusion, these results indicated that VWM can adapt to a resetting mode, but once the pointer is lost, a resetting process is triggered regardless of the context.

36.458 A CHANGE LOCALIZATION BENEFIT FOR MIXED ARRAYS OVER UNIFORM ARRAYS

Temilade Adekoya[1] (tyadekoya@uchicago.edu), Chong Zhao[1], Edward Vogel[1], Edward Awh[1]; [1]University of Chicago

Given relatively sharp capacity limits in visual working memory (WM), there has been sustained interest in whether these limits are based on the number of individuated items stored, or based on interference that varies as a function of inter-item similarity. Object-based models predict that the costs of concurrent storage will be determined only by the total number of items stored, while feature-based models of capacity often argue that competition for feature-specific resources is the key limiting factor. To examine this question we manipulated both the number of stimuli within each memory array, as well as the similarity between those items. Thus, we measured visual WM performance with uniform arrays with only one type of feature (e.g. color), or mixed arrays with two feature-types (e.g. color and orientation). Consistent with object-based models, we observed a significant cost of increasing load in both the mixed and uniform conditions. However, we also saw evidence for a modest advantage in the mixed array conditions, a difference that is not predicted by pure object-based models. Our follow-up work examined whether the advantage in the mixed condition reflects differences in whether items are stored, or differences in memory fidelity. We found that this advantage indeed reflected a change in memory fidelity, specifically a better precision in memory of items in the mixed arrays than in uniform arrays. Additionally, there was no evidence that participants were storing a larger number of items in the mixed condition. These results suggest that working memory performance is subject to item-based limits independent of item similarity, but the fidelity of those memories can be shaped by inter-item similarity.

We acknowledge the funding from National Institute of Mental Health (grant ROIMH087214); Office of Naval Research (grant N00014-12-1-0972).

36.459 COMPARING REPRESENTATIONAL STRUCTURES FOR SIMPLE AND COMPLEX STIMULI IN VISUAL WORKING MEMORY

Isabella Longoria-Valenzuela[1] (ilongoriavalenzuela@ucsd.edu), Timothy Brady[1], John Serences[1]; [1]University of California, San Diego

Different stimulus spaces have different representational structures: Gabor representations in perception and memory may depend primarily on tuning functions in early visual areas, while faces may depend on a whole hierarchy of face-selective and non-selective visual representations across the ventral stream. How do these differing representational structures affect memory errors when tasked to remember these stimuli? To test this, we examined working memory error distributions and neural representations for 3 stimulus spaces: gabors, and two face "wheels." To create these face wheels, a generative adversarial network was used by picking a random plane in the latent space and generating a circle with either a small or large radius. This gave us one wheel with more similar and one with much more distinct face stimuli. Target stimuli for the memory task were then sampled from these wheels uniformly every 10 degrees, creating 90 target stimuli in total across the 3 wheels. In a behavioral experiment, participants were asked to memorize either 2 or 4 stimuli from these wheels and report a target stimulus's location on the corresponding wheel. Memory error on each trial was calculated as the difference between the target stimulus and the recorded wheel response. Across all wheels, remembering 4 stimuli produced larger memory errors than trials showing 2 stimuli, and there were large differences in memory error distributions for the three stimuli wheels that allowed us to model the representational structure of simple and more complex stimulus spaces as revealed by behavior.

NSF BCS- 2146988 to TFB & R01-EY025872 to JTS

36.460 EXPLORING THE RELATIONSHIP BETWEEN FLUID INTELLIGENCE AND VISUAL WORKING MEMORY FOR SIMPLE FEATURES VS. REAL-WORLD OBJECTS

Kaira Shlipak[1], Yong Hoon Chung[1], Viola Störmer[1]; [1]Dartmouth College

Visual working memory performance has been shown to strongly correlate with measures of fluid intelligence. These correlations have been observed with working memory tasks that used simple visual stimuli such as colors or simple shapes (e.g., Fukuda et al., 2010). However, recent research has demonstrated that visual working memory performance is increased when people are asked to remember meaningful stimuli such as real-world objects relative to simple feature items (e.g., Brady & Störmer, 2022). How does performance on working memory tasks using real-world objects relate to measures of fluid intelligence? Here, we examined this question by measuring working memory performance for simple stimuli and real-world objects, and correlated individuals' performance on these tasks with measures of fluid intelligence. Participants (N = 103) completed different visual working memory tasks that used either colored circles or images of real-world objects as well as two different intelligence tests: the Raven's Advanced Progressive Matrices Test of fluid intelligence and the C-test of crystallized intelligence. First, we replicated the real-world object advantage in visual working memory, such that memory performance was higher for real-world objects relative to colors. Second, we found reliable correlations between all working memory measures (both objects and colors) and fluid intelligence ($r2 = 0.38$, $p < 0.0001$ for object condition; $r2 = 0.25$, $p = 0.01$ for color condition), and no significant differences between these correlations across stimuli types. Notably, working memory performance was not reliably correlated with crystallized intelligence in either object or color conditions, which indicates that crystallized knowledge is not linked to visual working memory performance even for meaningful objects. These findings suggest that visual working memory performance in tasks using both simple features and complex real-world objects relies on shared processes that relate to fluid intelligence.

36.461 HIGHLY MEMORABLE IMAGES ARE EASIER TO PERCEIVE

Will Deng[1] (wuyizhe2@illinois.edu), Kara Federmeier[1], Diane Beck[1]; [1]University of Illinois Urbana-Champaign

Image memorability, the likelihood that a person will remember an image, is an intrinsic property of the image that is distinct from many other visual and cognitive features. Research thus far has not identified any particular visual features that sufficiently explain this intrinsic memorability. We investigated the perceptual component of memorability by postulating a connection between memorability and statistical regularity, which refers to how well images match learned visual patterns. Statistical regularity affects detection time for images, such that stimuli with higher statistical regularity can be detected with shorter presentation. Therefore, we probed if memorability affects how quickly images are detected: we presented participants with high and low memorability images in an intact/scrambled task wherein they indicated whether they saw an intact image or noise, and we estimated the presentation duration necessary for participants to reach 70.7% accuracy. With two sets of stimuli, we observed and replicated that more memorable images have shorter detection thresholds than less memorable images, suggesting that more memorable images are more readily perceived. We further explored the perceptual processing evoked by memorable images using event-related potentials (ERPs) time-locked to stimulus onset. In a continuous recognition paradigm with delayed repetition, participants indicated whether they saw a new or repeated image. We focused our analyses on the N300 and N400 components, associated with high-level visual processing and long-term semantic memory access respectively. More memorable images evoked less negative N300 and N400 amplitudes, suggesting both facilitated perceptual and semantic processing. Overall, our results support the idea that memorability is an intermediate property related to both ease of perception and access to semantic memory, consistent with the idea that certain stimuli may better match stored templates used for image perception and recognition.

36.462 THE CREATION OF ARTWORK REVEALS TEMPORAL AND SPATIAL PROPERTIES OF MEMORABILITY

Trent M. Davis[1], Wilma A. Bainbridge[1]; [1]University of Chicago

By shifting focus away from viewers and instead toward stimuli, memorability research has made unique contributions to our understanding of human memory. Previous work has shown that in addition to stimuli including faces, naturalistic scenes, and objects, paintings have an intrinsic memorability (Davis & Bainbridge, 2023). This means that our memories of artworks are consistent despite art's connotation as a subjective field. We developed a database of over 2,500 still images collected from 50 artworks whereby their memorability was tracked over time as they were created. Here, we use this database to dive into the temporal and spatial properties that underlie changes in memorability. All changes to the artworks were tracked and labeled by artists as one of three categories: sketches, blockings, or details; details produce the smallest changes in memorability and blockings produce the greatest change. We also observed a significant, positive relationship between the number of pixels that changed and the change in memorability. Further, we found that visual features including color and clutter have no significant impact on an artwork's memorability. However, not all changes are created equal; by taking the difference between subsequent images and converting them to a binary mask, we determined that centrally-focused changes impact memorability the most, and this impact significantly decreases as distance from the center of the image increases. When looking at raw changes in memorability, paintings with localized changes higher up in the y-axis have a greater memorability. Additionally, artworks with a one-point perspective have a significantly lower memorability than those that do not. These results suggest that an image's memorability is strongly influenced by its composition. Artists, designers, and researchers alike are implicated by these findings that better our understanding of what makes an image memorable.

Sunday, May 19, 2:45 – 6:45 pm, Pavilion

Action: Reach, grasp, track

36.463 OBJECT CENTER OF MASS PREDICTS ENDPOINTS FOR FREE-HAND POINTING IN VIRTUAL REALITY

Immo Schuetz[1] (schuetz.immo@gmail.com), Leah Trawnitschek[1], Katja Fiehler[1]; [1]Experimental Psychology, Justus Liebig University Giessen, Giessen, Germany

Humans learn to point using their index finger at a very young age, intuitively directing other people's attention to objects in the environment. Typically, this is achieved by bringing the dominant eye, index finger tip of the dominant hand, and pointing target into alignment (eye-hand alignment; also termed "eye-finger-raycast" when used in human-computer interaction). This requires the human sensorimotor system to select an appropriate target location to plan the pointing movement. While free-hand pointing gestures have been extensively studied using small and well defined targets (e.g., numbers shown on a wall), it is currently less understood how humans select a specific target location on a more complex three-dimensional (3D) object. Eye movement research suggests that the center of an object has the highest likelihood to be fixated, and that this fixation location shifts to an object's implied 3D center of mass (CoM) if object characteristics such as shading imply a 3D shape, making the CoM a logical candidate. Here, we investigated whether the CoM also serves as a target for eye-hand alignment when pointing to 3D objects in virtual reality (VR). Participants pointed at custom objects ("vases") using their index finger, using different visual feedback and from different vantage points in a virtual room. Hand movements were tracked using a Vive Tracker. We computed pointing vectors for each trial using eye and hand tracking, and then estimated 3D endpoints within the object from averaged pointing vectors. Endpoints were best predicted by CoM, and manipulating 3D object shape to systematically shift the CoM induced corresponding shifts in participants' pointing endpoints. The type of visual feedback provided during pointing influenced overall accuracy, but did not influence the effect of CoM manipulation. These results suggest that object CoM plays a major role in eye-hand alignment when pointing to 3D objects in VR.

Cluster project „The Adaptive Mind“ funded by the Hessisan Ministry of Higher Education, Research, Science, and the Arts

36.464 MEASURING CONSCIOUS MONITORING AND METACOGNITION AT THE START, MIDDLE AND END OF A REACHING MOVEMENT

Gabriela Oancea[1,2], Craig S. Chapman[1,2]; [1]University of Alberta, Faculty of Kinesiology, Sport, and Recreation, [2]University of Alberta, Neuroscience and Mental Health Institute

The ability to monitor our arm position during goal-directed behaviour allows us to bring our limb to a target as accurately as possible. Despite our success in executing accurate movements, some work suggests that individuals have limited access to information about their limb position (Charles et al., 2020). However, contradictory evidence from metacognition research indicates that people are able to accurately monitor their movements. In these studies, individuals are asked to rate their confidence after making judgements about their movements and tend to give higher confidence ratings when they are correct, showing some capacity for self-monitoring (Arbuzova et al., 2021). These conflicting results suggest that we do not monitor an entire movement. Participants (n=43) made reaching movements toward targets on a screen. They were then presented with two movement paths: one being their actual trajectory and the other being a visually deviated version. Here, we manipulated the location that the deviation was implemented (i.e., start, middle, or end of the path). Participants were asked to determine which trajectory was their own, while also rating their confidence in their response. Overall, accuracy was lower than expected. Nevertheless, accuracy and confidence were higher when deviations occurred in the middle and end of the movement as opposed to the start, suggesting that participants were more aware of their true limb position at the middle and end of their reach. In addition, metacognitive sensitivity was greater during the middle and end implying that at these locations, individuals' confidence ratings better discriminated between correct and incorrect responses, indicating appropriate self-monitoring. We conclude that people have a remarkable blindness to the properties of their own movements. As well, monitoring of a limb is significantly reduced at the start of a movement suggesting reduced attention to limb position at this time, possibly due to movement programming demands.

CIHR, NSERC

36.465 EFFECTS OF SENSORIMOTOR ADAPTATION ON CONFIDENCE

Marissa Fassold[1] (mhe229@nyu.edu), Michael Landy[1,2]; [1]New York University, [2]Center for Neural Science, New York University

Humans are able to adapt to large and sudden perturbations of sensory feedback. For example, after plunging one's hand into the water to retrieve a shell there is an immediate mismatch between visual feedback and proprioceptively sensed location. Sensor-motor adaptation takes time, as anyone who has encountered an unexpected computer mouse sensitivity can attest. What sensory and motor-execution cues are used to determine confidence, and do the dynamics of confidence parallel those of ongoing sensorimotor adaptation? Participants made a slicing reach through a visual target with an unseen hand followed by a continuous judgment of confidence in the success of their reach. After the confidence response, visual feedback of hand position was shown at the same distance along the reach as the target. For the confidence judgment, participants adjusted the size of an arc centered on the target. Larger arcs reflected lower confidence. Points were awarded if the visual feedback was within the arc, and fewer points returned for larger arcs. This incentivized attentive reporting and minimizing feedback-target distance to maximize the score. A fixed, rotational perturbation (alternating clockwise/counterclockwise across blocks) was applied to the feedback on trials 20-70 within each 100-trial block. We used least-squares cross validation to compare four Bayesian-inference models of sensorimotor confidence adaptation based on prospective cues (e.g., knowledge of motor noise and past performance), retrospective cues (e.g., proprioceptive measurements), or both sources of information to maximize expected gain (i.e., an ideal observer) with additional parameters for learning and bias. All of the participants use proprioception when calculating sensorimotor confidence during motor adaptation. Most participants depended primarily on a recalibrated proprioceptive signal for confidence. Over repeated blocks of exposure to the perturbation, participants' confidence recovered exponentially to pre-adaptation levels, but at a different rate than motor learning.

Funding: NIH EY08266 (MSL)

36.466 RAPID ADAPTATION TO ACCELERATION DURING INTERCEPTIVE HAND MOVEMENTS

Philipp Kreyenmeier[1] (philipp.kreyenmeier@googlemail.com), Miriam Spering[1]; [1]University of British Columbia

Real-world objects in our environment rarely move at a constant speed but usually accelerate or decelerate. Yet, human perception is highly insensitive to visual acceleration. When manually intercepting moving objects, humans commonly ignore acceleration, resulting in systematic interception errors (Kreyenmeier et al., 2022, eNeuro). Here we ask whether humans' ability to manually intercept accelerating targets improves during repeated exposure to the same rate of acceleration. In a track-intercept task, observers (n=9) tracked the ramp of a small target either moving at constant speed (0 deg/s/s), accelerating (+8 deg/s/s), or decelerating (-8 deg/s/s). After 800 ms, the target disappeared behind an occluder and observers had to rapidly point at the target at the predicted time of reappearance from behind the occluder (time-to-contact; TTC). Observers performed blocks of twelve trials during which they were exposed to the same rate of acceleration. During the first eight trials, the occluder had a fixed width (reference), in the remaining four trials, the occluder was either narrower or wider than the reference (test). In the first trial of each block, observers systematically intercepted too late for accelerating targets and too early for decelerating targets, indicating that they did not take acceleration into account. Within the first four reference trials, they adjusted the timing of their hand movement to match veridical target TTC. In test trials, observers only partially accounted for acceleration and showed similar biases as in early reference trials. Our results show that humans can rapidly adjust the timing of their hand movement to intercept accelerating targets. However, their ability to transfer this adjustment to new TTC conditions is limited. These findings provide further evidence for the inability to decode accelerating motion and to accurately interact with accelerating objects.

36.467 "AUTOMATIC" ONLINE CORRECTIONS DURING A REACHING TASK ARE ASSOCIATED WITH INDIVIDUAL DIFFERENCES IN EXECUTIVE FUNCTION.

Christopher L. Striemer[1,2] (striemerc@macewan.ca), Branden T. Otte[1]; [1]MacEwan University, Edmonton, Alberta, Canada, [2]University of Alberta, Edmonton, Alberta, Canada

Previous research has demonstrated that the dorsal visual stream is able to execute rapid online movement corrections to sudden changes in target position. This "automatic pilot" can operate in the absence of visual awareness, and even under circumstances where participants are instructed to not correct their movements. In the current study, we examined the extent to which these "automatic" corrections might be related to individual differences in executive control. To examine this, healthy adult participants (n=80) completed two versions of the automatic pilot task (APT) on a touch screen: 1) a "correct" condition in which participants were instructed to correct their movement to the new target location on jump trials, and 2) an "ignore" condition in which participants were told to ignore any target jumps, and point to the initial target location. In addition to completing these two versions of the APT, participants also completed the Sustained Attention to Response Task (SART), in which they were asked to respond when a number was presented, except for the number 3. Finally, participants completed questionnaires indexing executive attention, impulsivity, and executive function including the Adult ADHD Self Report Scale (ASRS), the Cognitive Failures Questionnaire (CFQ), and the Behavioural Rating Inventory of Executive Function for Adults (BRIEF-A). Our results indicated that, similar to previous research, participants made significantly more corrections to target jumps in the "correct" condition, compared to the "ignore" condition. Importantly, "automatic" unintended corrections in the "ignore" condition were significantly correlated with poorer scores on the ASRS, the CFQ, and the BRIEF-A. However, unintended corrections were not correlated with errors or reaction times on the SART. These results suggest that the APT is sensitive to individual differences in executive function, and may be

useful as a visuomotor measure of response inhibition and cognitive control in both healthy and clinical populations.

This research was funded through a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to C.S.

36.468 MINIMAL INFLUENCE OF SUPPRESSED DISTRACTORS ON REACHING MOVEMENTS

Shinhae Ahn[1] (a.shinhae@wustl.edu), Richard A. Abrams[1]; [1]Washington University in St. Louis

Attention and action are tightly intertwined--action goes where attention goes. For goal-directed reaching, suppressing salient distractors is crucial to avoid unintended movements. Previous studies have shown that salient distractors influence reaching movement trajectories, causing deviations toward the distractors. However, those results were obtained under conditions where attentional suppression of the distractors was not possible. Little is known about how suppressed distractors may influence reaching movements. The present study examined the influence of distractor suppression on the trajectory of reaching movements toward targets. In a visual search task, participants searched for a target shape among non-target distractors in which one of the distractors contained a salient-but-irrelevant color singleton on some trials. Participants were instructed to reach to the target by moving a mouse cursor while ignoring the salient distractor. The search mode was manipulated in two experiments: singleton-detection mode, where the target was defined as a unique shape among homogeneous distractors, promoting attentional capture by salient singleton, and feature-search mode, where the target shape was specific among heterogeneous distractors, facilitating singleton suppression. The onset and duration of the reaching movements replicated the patterns observed in previous studies: Salient distractors produced attentional capture in singleton-detection mode, slowing the onset and completion of the movements, but yielded a slight benefit in feature-search mode, implying effective distractor suppression. Importantly, analysis of the movements revealed that the presence of salient distractors significantly affected reaching movement trajectories in singleton detection mode, showing a deviation toward the distractors consistent with attentional capture; however minimal impact was observed on the trajectories under feature-search mode, indicating successful suppression of the distractors. Overall, the findings suggest that suppression of salient distractors not only reduces their effects on attentional and perceptual processing, but can minimize their impact on goal-directed reaching actions.

36.469 AN INHIBITORY MECHANISM PREVENTS OUTDATED ACTIONS OF EYE AND HAND MOVEMENTS

Clara Kuper[1,2] (clara.kuper@student.hu-berlin.de), Martin Rolfs[1,3]; [1]Humboldt Universität zu Berlin, [2]Berlin School of Mind and Brain, [3]Bernstein Center for Computational Neuroscience Berlin

Sequences of saccadic eye movements are inhibited after salient changes, even when they are task-irrelevant [Reingold & Stampe, 2002]. This phenomenon of saccadic inhibition is considered an adaptive mechanism, preventing outdated movements while the change is evaluated [Stanford & Salinas, 2018]. Here, we show that inhibition extends to sequences of hand movements, implying that a global signal inhibits action execution across motor systems until the evaluation of changes is complete and new movement plans have been established. In two experiments executed online on participants' smartphone touchscreens, we evaluated hand movement rates after task-relevant and task-irrelevant changes. Participants collected a series of six randomly jittered movement targets by tapping on them. The task-relevant change (shown in 50% of all trials) was a displacement of the movement targets. The task-irrelevant change (shown in an independently chosen 50% of all trials) was a brief flash. The location and contrast of the flash varied between the experiments. Movement rates dropped below baseline after both task-relevant and irrelevant changes. Actions after the inhibition reflected a complete movement update, never an average of two responses. This suggests that the evaluation of the change continued throughout the inhibition. In a lab-based replication of the key results, we assessed the reliability of the timing of the online measurements and collected eye movement behavior. The combined data confirms that trials in which eye movements were likely to be inhibited were also more likely to show inhibited hand movements. We suggest that upon detection of a salient change, any overt motor output is suppressed – regardless of the motor system – to prevent the execution of outdated actions. This mechanism could reduce the number of incorrect motor actions without disrupting the cognitive processes evaluating the state of the world.

36.470 BAYESIAN INFERENCE OF PERCEPTUAL UNCERTAINTY, BEHAVIORAL COSTS, AND PRIOR BELIEFS FOR CONTINUOUS PERCEPTION-ACTION TASKS

Tobias F. Niehues[1,2] (tobias.niehues@stud.tu-darmstadt.de), Dominik Straub[1,2], Constantin A. Rothkopf[1,2]; [1]Institute of Psychology, Technische Universität Darmstadt, [2]Centre for Cognitive Science, Technische Universität Darmstadt

Bayesian observer and actor models have provided normative explanations for behavior in many perception-action tasks including discrimination tasks, cue combination, and sensorimotor control by attributing behavioral variability and biases to factors such as perceptual and motor uncertainty, prior beliefs, and behavioral costs. However, it is unclear how to extend these models to more complex tasks such as continuous production and reproduction tasks, because inferring behavioral parameters is often difficult due to analytical intractability. Here, we overcome this limitation by approximating Bayesian actor models using neural networks. Because Bayesian actor models are analytically tractable only for a very limited set of probabil- ity distributions, e.g. Gaussians, and cost functions, e.g. quadratic, one typically uses numerical methods. This makes inference of their parameters computationally difficult. To address this, we approximate the optimal actor using a neural network trained on a wide range of different parameter settings. The pre-trained neural network is then used to efficiently perform sampling-based inference of the Bayesian actor model's parameters with performance gains of up to three orders of magnitude compared to numerical solution methods. We validated our proposed method on synthetic data, showing that recovery of sensorimotor parameters is feasible. Importantly, individual behavioral differences can be attributed to differences in perceptual uncertainty, motor variability, and internal costs. We finally analyzed real data from a task in which participants had to throw beanbags towards targets at different distances. Behaviorally, subjects differed in how strongly they undershot and overshot different targets and whether they showed a regression to the

mean over trials. We could attribute these complex behavioral patterns to changes in priors because of learning and undershoots and overshoots to behavioral costs and motor variability. Taken together, we present an analysis method applicable to continuous production and reproduction tasks.

This research was supported by the European Research Council (ERC; Consolidator Award 'ACTOR'-project number ERC-CoG-101045783).

36.471 LIGHTING CONDITIONS AFFECT THE SELECTION OF CONTACT POINTS WHEN GRASPING COMPLEX 3D SHAPES

Zoltan Derzsi[1,2] (zd8@nyu.edu), Francesco Ceccarini[1,2], Robert Volcic[1,2,3]; [1]New York University Abu Dhabi, [2]NYUAD Center for Artificial Intelligence and Robotics, [3]NYUAD Center for Brain and Health

When humans grasp an object, there are several factors such as the natural grasp axis, force closure and minimum torque that impact the selection of the most appropriate contact points on the surface of the object and lead to natural and effortless grasping actions. For this process to be successful, constraints relative to 3D shape perception need also to be considered. It is indeed known that interactions between shape, material properties and lighting conditions can adversely affect 3D shape constancy. However, studies and models on grasp contact point selection have so far rarely paid attention to the effects of these factors. Here, we addressed this gap by focusing on how contact point selection in grasping might be affected by different lighting conditions. Participants were asked to grasp and lift a real 3D globally convex smoothly curved object created by distorting a sphere with a fractal Brownian motion displacement algorithm and presented at different orientations. Instead of presenting the object under diffuse ambient illumination, the object was illuminated with one directional light (cold-white pin-spotlight, 5000 lux) from one side at a time (top-left vs. top-right). This illumination altered the object's appearance such that the illuminated areas were very well visible, but the non-illuminated areas appeared as a very dark attached shadow in our otherwise dark lab environment, similar in appearance to photographs of asteroids. By considering the contact positions of the thumb and index finger just before lifting the object, we observed that the selection of these contact points changed systematically under variations of illuminant direction. Our results imply that the factors affecting the selection of contact points should necessarily include effects of lighting conditions and possibly material properties as well, as they can often influence the visual appearance of 3D objects.

We acknowledge the support of the NYUAD Center for Artificial Intelligence and Robotics and the NYUAD Center for Brain and Health, funded by Tamkeen under the NYUAD Research Institute Awards CG010 and CG012.

36.472 OBJECT AFFORDANCES THROUGH THE WINDOW OF VERB USAGE PATTERNS AND BEHAVIOR

Maryam Vaziri-Pashkam[1] (mvaziri@udel.edu), Ka-Chin Lam[2], Natalia Pallis-hassani[2], Aida Mirebrahimi[3], Aryan Zoroufi[4], Francisco Pereira[2], Chris Baker[2]; [1]Department of Psychological and Brain Sciences, University of Delaware, [2]National Institute of Mental Health, [3]Carnegie Mellon University, [4]Massachusetts Institute of Technology

When we see objects, we immediately know how to interact with them. Little research has been performed to understand what information people glean from objects about the interactions they support. Here, we first used language as a means to tap into humans' knowledge of what actions can be performed with an object. Using a large database of ~1850 object categories (THINGS database) and ~5000 verbs, we identified applications of each verb to each object in a large text corpus. We then used this data to embed each object in a space where dimensions correspond to verbs that apply to similar objects. We showed, in behavioral experiments, that these extracted embedding dimensions are meaningful to human observers. Next, to reveal people's understanding of potential actions towards objects, we conducted online behavioral experiments in which we presented images of individual objects from the THINGS database and asked people about the actions they associate with the objects and body parts they use to interact with the objects. Many objects, including both tool and non-tool items, had a strong action association. Although hand was the most common body part implicated, other body parts were also reported to be heavily involved in interacting with objects. Together, these results indicate strong object-action associations evident in both text corpora and in people's reports from viewing pictures of objects. They uncover the richness of object interactions and argue for moving beyond simple hand grasps and beyond the specific category of tools in future behavioral and neuroscientific experiments.

36.473 VISUAL UNCERTAINTY ABOUT TARGET AND CURSOR IN A CONTINUOUS PSYCHOPHYSICS TASK DIFFERENTLY AFFECT TRACKING PERFORMANCE

Dominik Straub[1] (dominik.straub@tu-darmstadt.de), Lukas Maninger[1], Constantin A. Rothkopf[1]; [1]Technische Universität Darmstadt

Continuous psychophysics is a new methodological approach gaining traction in vision science, because it enables measuring perceptual uncertainty significantly faster compared to classical psychophysics. Instead of performing forced choice trials, subjects track a target on a screen, promising broad applicability with untrained subjects. However, due to the naturalistic task, several perceptual, cognitive, and motor parameters contribute to behavior in nontrivial ways, which are not yet well understood. For example, because subjects track the target with a cursor, there is not only visual uncertainty about the target, but also about the cursor. Here, we address the question of how visual uncertainty about the cursor affects tracking behavior. Subjects tracked a target, which moved on a random walk, using a mouse cursor. In two conditions, we manipulated visual uncertainty about either the target's or the cursor's position. In the target uncertain condition, the target was displayed as a dot cloud pattern, while the cursor was displayed as a single dot and vice versa in the cursor uncertain condition. The dots in the pattern were drawn from a Gaussian distribution, and we manipulated the perceptual uncertainty by changing its standard deviation. For different levels of target uncertainty, we observed differences in both root mean squared error (RMSE) and temporal lag, in accordance with previous results. For different levels of cursor uncertainty, we observed a similar effect on the RMSE, but a smaller effect on temporal lag. We conclude by showing why visual uncertainty about the cursor and the target affect tracking behavior in different ways with our recently developed optimal control model of continuous psychophysics. In particular, because the subjects control the cursor, but not the target, the uncertain visual information is weighted in different ways against internal model predictions, resulting in the observed behavioral effects.

36.474 VIRTUAL REALITY, REAL INFORMATION FOR ACTION AND PERCEPTION? A VR STUDY.

Caterina Foglino[1], Tamara Watson[3], Niklas Stein[4], Patrizia Fattori[1,2], Annalisa Bosco[1,2]; [1]Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy, [2]Alma Mater Research Institute for Human-Centered Artificial Intelligence (Alma Human AI), University of Bologna, Italy, [3]School of Psychology, MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Australia, [4]Institute for Psychology, University of Münster, Germany

Distances are often misperceived in virtual environments but, while a brief period of interaction through physical walking can improve this, other body-based interactions, such as reaching, do not. In fact, research in the real world suggests that action planning affects perceptual processing by biasing the cognitive system toward response-related dimensions facilitating their perception, but the role of this action-perception interplay in the virtual space is far from being fully understood. To contribute to this area of research, this study investigates the perception of an object in a virtual environment by testing the performance of 15 healthy participants (Males=3; mean age=28.4 ± 3.62 years; all right-handed) using a size judgment task. Participants were instructed to interact with the virtual object by either grasping, reaching, or no hand-movement conditions using a virtual copy of their real hand. They were then asked to report the estimation of the object size by adjusting the dimension of a comparison-object before and after the interaction phase. The interaction phase was preceded by a walking simulation to approach the target positioned far away from the participant. The results show that, overall, the size estimation errors improve after the interaction phase (β= -0.33812 mm, SE= 0.14486, t= 2.334, $p<0.05$). However, the no hand-movement interaction condition leads to significant smaller errors compared to the grasping one (β= -0.78053 mm, SE= 0.25181, t= -3.100, $p<0.01$). These findings suggest that interacting with the object through hand movements, specifically grasping, which is known to facilitate the detection of size-related features in the real-world, does not lead to an improvement on the perception of the object's size in this virtual reality context.

MAIA project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 951910; Work supported by Ministry of University and Research (MUR), PRIN2022-2022BK2NPS; Work supported by Horizon Europe MSCA No 101086206-PLACES.

MONDAY MORNING POSTERS IN BANYAN BREEZEWAY

Monday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

Motion: Illusions

43.301 A NEW EXPLANATION OF THE FRASER-WILCOX ILLUSION

Patrick Cavanagh[1] (patcav1@yorku.ca), George Mather[2]; [1]Glendon College, CVR, York University, [2]University of Sussex

The Fraser-Wilcox illusion is seen in a circular pattern that is divided into sectors containing repeated luminance gradients. The pattern appears to rotate, especially when the observer makes saccadic eye movements or blinks. The illusion has been known since 1979, but its origin is still disputed. The most prominent theory, which proposes that the latency of neural responses varies across the pattern (Conway et al., J. Neuro., 2005), cannot explain the long duration of the motion or why the illusion is weaker in older participants, or under pinhole viewing. We propose that the illusion is caused by changes in the retinal luminance of the pattern, due either to changes in pupil diameter during eye movements or to the transit of the eyelid across the pupil during blinks. It is already known that luminance change can generate illusory motion in patterns containing luminance gradients (e.g., Mather, Vision Res., 1984), so we argue that the Fraser-Wilcox illusion is part of this illusion class. This new explanation can account for the importance of eye movements and blinks, the effect of age (reduced pupil mobility), and the duration of the motion impressions. A simulation of the illusion in which pattern luminance is modulated with the same time-course as that caused by blinks and saccades creates a markedly enhanced impression of illusory motion.

NSERC Discovery RGPIN-2019-03989 (PC), Leverhulme Trust Emeritus Fellowship (GM)

43.302 ARE ILLUSORY VISUAL PHANTOMS SEEN BY THE MOTION SYSTEM: INVESTIGATIONS UTILIZING THE MOTION AFTEREFFECT

Loic Daumail[1] (loic.daumail@vanderbilt.edu), Frank Tong[2]; [1]Vanderbilt University

The constructive nature of the brain, such as when visual percepts occur in regions with no direct stimulation, can be investigated through the motion processing system. Here, we investigated the perception of illusory visual phantoms, a form of completion that occurs when two low-contrast collinear drifting gratings are separated by a homogeneous region, leading to the ghostly impression of stripes that extend across the gap. Previous reports have suggested that a motion aftereffect (MAE) can be induced by visual phantoms (Weisstein, 1977; Science); however, our preliminary investigations with static test patterns failed to replicate these reports. We therefore explored whether visual phantoms might elicit an MAE using a dynamic counterphasing test pattern, as higher-level forms of motion perception can often elicit a dynamic MAE (dMAE) even if no static

MAE occurs. Our displays consisted of two collinear horizontal gratings (12° × 6°, 0.25 cyc/deg), separated by a vertical (5°) gap, that drifted (2 Hz) during an adaptation phase. A counterphasing grating appeared in the blank gap location during test. Either the darkest portion of the inducer gratings matched the background luminance to elicit vivid phantoms, or the mean luminance of the inducer gratings matched the background luminance as a control condition. Three contrast levels were tested: 2.73%,13.64%, and 63.64%. Participants reported whether the test grating moved upward, downward or whether motion was ambiguous. dMAE was strongly biased in the motion direction opposite to that of the adapting motion direction, but curiously, it appeared equally strong for the visual phantom condition and phantom control condition. These results suggest that even though salient visual phantoms are perceived in some conditions but not others, the perceptual representation of the phantoms does not seem to provide direct input to the visual motion processing system.

Supported by NIH grants R01EY029278 and R01EY035157 to FT.

43.303 INDIVIDUAL DIFFERENCES IN LONG-RANGE VISUAL APPARENT MOTION

Nicolaas Prins[1] (prins.nicolaas@gmail.com); [1]University of Mississippi

The correspondence problem in apparent motion occurs when motion tokens in one motion frame have multiple possible matches in the subsequent frame. Previously, Prins (2023, https://doi.org/10.3758/s13428-023-02061-0) observed a single outlier among six observers who, relative to the other observers, displayed a quantitatively and qualitatively different pattern of the effects of three variables on perceived motion correspondence. These variables were: the relative distances between tokens, whether motion paths preserved or violated the figural integrity of the tokens, and the attentional load of a secondary task. Here I test 33 observers to explore individual differences among observers further. Motion tokens positioned on diagonally opposed corners of an imaginary rectangle were alternated with tokens positioned on the other diagonal. Each frame contained a square and a round token such that only one of the possible solutions to the correspondence problem would preserve token integrity. During the motion sequence, observers performed one of two secondary attention-demanding tasks that required fixation at center of motion display. Observers can be divided into three categories: (1) those who tend to report motion along the shorter of the two possible motion paths (the nearest-neighbor principle), with an average bias towards reporting vertical motion and towards motion that preserves token integrity (n = 8), (2) those who (almost) exclusively report motion that preserves token integrity, regardless of relative distances between tokens, (n = 7), (3) those who (almost) exclusively report vertical (n = 14) or horizontal (n = 1) motion, or respond independent of any of the variables (n = 3). Strikingly, while all of (1) displayed strong trial-to-trial hysteresis, none of (2) did. Overall, results are consistent with individual differences being mediated by different relative roles of a low-level, competitive mechanism acting on the nearest-neighbor principle on the one hand and top-down cognitive control on the other.

43.304 MOTION-INDUCED SPATIAL SHIFTS AND MOTION-BASED CONTRAST ENHANCEMENT ARE LINKED.

Alan Johnston[1] (a.johnston@ucl.ac.uk), Dan Hu[1]; [1]University of Nottingham

Two lines of research on the effects of motion on spatial vision have developed largely independently. Motion induces a forward shift in the apparent spatial location of moving pattern (Ramachandran & Anstis, 1990; De Valois & De Valois, 1991). Also, moving pattern appears less blurred than might be expected from temporal summation or visual persistence (Burr, 1980). This might be explained by spatial summation along the motion trajectory (Nishida et al., 2007). Here we describe evidence that the two phenomena may be connected. We constructed a tripartite stimulus consisting of an array of square binary blocks (2 pixel; 0.039 × 0.039 deg) in a fixed window on a mean luminance grey background. The central rows translated by one or more blocks while the rows in the upper and lower third of the array randomly updated from frame to frame. Thus each frame appeared as an undifferentiated random block array and the temporal properties of the translating and updating blocks were identical. Participants reported whether the moving pattern or the flanking updating regions appeared to have the higher contrast. The contrast of the randomly updating pattern was varied by a staircase procedure to determine the matching contrast. We found the apparent contrast of the moving pattern was generally higher than the updating pattern. The effect was greater for slower speeds and for lower contrasts. We also measured the apparent spatial shift of the translating section relative to the updating section using a staircase method while varying the drifting section position. The spatial shift was greater for lower contrasts. Since the same participants contributed to both experiments, we used repeated measures correlation to compare scores over participants and conditions. We found a high correlation (0.69, p<0.001) between the spatial shift as a proportion of the pattern displacement and the apparent contrast enhancement.

Supported by the Leverhulme Trust

43.305 PREDICTIVE POSITION OF MOVING OBJECTS IS AFFECTED BY THEIR ORIENTATION WITH RESPECT TO THE MOTION TRAJECTORY

MASSIMO GIRELLI[1] (massimo.girelli@univr.it); [1]University of Verona, Dept. Neuroscience, Section of Psychology

The orientation of a moving object with respect to its motion trajectory is of fundamental importance for visual perception: It facilitates the predictions about the fate of the object we need to interact with and it mimics the multiple stimulation along the collinear axis to form borders. In the Beta motion, by manipulating the spatial gap between consecutive positions, it is possible to observe the K-effect i.e. The time of a larger gap appears to be prolonged with respect to the actual time. In the first experiment I manipulated the orientation of the elements leading to the Beta motion i.e. orthogonal or collinear with respect to the motion trajectory in a three step sequence with spatially different gaps to get the K-effect. The results from 7 observers showed that the collinear motion led to a significantly smaller K-effect in the final gap of the sequence, on average, up to 41% with respect to the orthogonal motion thus the collinear motion fills the gaps of the Beta motion. In the second experiment, the modulation of the motion extrapolation of moving elements with respect to static ones as in the Flash-Lag-Effect (FLE), was tested with orthogonal and collinear moving elements. Contrary to what it was observed in the K-effect, the collinear moving elements strengthened the FLE. In 15 observers the illusory gap between the collinear motion and the static flash was 61% larger with respect to the orthogonal motion. Thus, collinear motion better provides evidence of predictive activation along motion trajectory. All in all, objects moving collinearly to their motion trajectory embodies better the prediction of their future positions in space as in an anticipatory jet contrail that overlaps the motion trajectory.

43.306 THE DOUBLE-DRIFT ILLUSION IS DOMINATED BY THE FIRST-ORDER MOTION ENERGY OF THE INTERNAL TEXTURE

Eunhye Choe[1] (eunhye.choe.gr@dartmouth.edu), Nathan H. Heller[1], Patrick Cavanagh[2], Peter U. Tse[1]; [1]Dartmouth College, [2]Glendon College

In the double drift (also infinite regress and curve ball) illusion, a Gabor patch moves in one direction (external motion) while its internal texture moves orthogonally to the external motion. Instead of appearing to move in its physical direction, the Gabor patch appears to move in a direction midway between the external and internal motions. However, because of the effect of the Gaussian aperture on the carrier, the texture's first-order motion also moves in the illusory direction. It is not yet known whether the illusory motion is produced by a vector combination of two motion sources or simply tracks the apertured, first-order motion. Here, we test whether the illusion is driven by first-order motion or a combination of motion sources. To overcome the aperture problem, we used 1/f noise as the texture that moved back and forth along a path at 12° eccentricity. We varied the angular difference between the internal and external direction from 0° to 90° in steps of 15°. Pilot observations revealed that larger angle differences produced complex, non-linear percepts. Therefore, participants had two response options. If they perceived a linear trajectory, they matched its angle with a rotating bar. If they perceived a non-linear trajectory, participants pressed a key and moved on. For smaller angle differences (<45°), the illusory motion closely followed the direction of the internal motion. As the angle increased, the responses increasingly tended toward the direction of the external motion. These findings suggest that the illusion is dominated by first-order motion up to the 45° direction that characterizes the typical double drift Gabor stimuli. Beyond that, the internal and external motions begin to interact and the percepts become increasingly non-linear.

43.307 THE FLASH-LAG EFFECT IN BALL SPORTS PLAYERS

Reza Abbas Farishta[1] (reza.abbas@umontreal.ca), Emeline Delamarre[1], Xi Wang[2], Alexandre Reynaud[2]; [1]School of Optometry, Université de Montréal, [2]Dept. of Ophthalmology, McGill University

Estimating motion and trajectories is a core function of human visual system. It is particularly relevant for ball sports players such as baseball or tennis players, who must accurately catch or hit a ball to win their game. The flash-lag illusion is a motion-based illusion where one will see a transiently appearing "flash" lagging behind a moving object. It could be caused by predictive mechanisms which help to anticipate the position of the moving object. In this study, we wanted to test whether subjects who are trained at anticipating trajectories such as ball sports players are particularly sensitive to that illusion. We tested and compared three groups of participants, all students from the Université de Montréal: ball sports players, non-ball sports athletes, and controls on a standard flash-lag effect paradigm. A bar was horizontally moving on screen and a flashed bar was presented either above or below the moving bar. The participants had to report whether the flashed bar appeared left or right relative to the moving bar. We observed a typical flash-lag effect in all groups with no difference between groups. This suggest that the flash-lag effect could not be due to anticipation mechanisms (e.g. latency difference or mis localization) or that training those mechanisms has no effect.

MONDAY, MAY 20, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Visual Memory: Encoding, retrieval

43.308 STATISTICAL REGULARITIES DO NOT FACILITATE ENCODING BUT INDUCE BIASED GUESSING IN THE ATTRIBUTE AMNESIA TASK

Niya Yan[1] (yanniya@tamu.edu), Brian. A Anderson[1]; [1]Texas A&M University

While some studies have shown memory enhancement for items with statistical regularities, whether the effect persists when people are not expecting to report the probed information remains unclear. Here, we investigated whether an influence of statistical regularities on working memory encoding can be observed using the attribute amnesia task, a working memory task that employs a surprise test to probe people's memory for an attended item. Participants completed an epoch of 320 trials in which they reported the location of a target letter among three distractor digits. In 70% of those trials, targets were presented in one color (frequent color), while the remaining 30% involved targets displayed in one of three alternative colors with equal probability (infrequent colors, 10% each). Notably, for the 321st trial (surprise trial), the color of the target was unexpectedly probed prior to the location inquiry. In this trial, half of the participants had the target rendered in the frequent color (frequent group), while the other half had the target rendered in an infrequent color (infrequent group). Analysis revealed a significant difference in surprise trial color reporting accuracy, with the frequent group (45%) outperforming the infrequent group (16.7%). However, both groups exhibited a consistent bias towards selecting the frequent color as the target color during the surprise trial (frequent group: 45%, infrequent group: 50%). This suggests that, rather than an enhancement in working memory encoding specific to the frequent color, the observed between-group difference in surprise trial performance can be fully attributed to a bias to endorse the frequent color as the target color. In summary, our findings indicate that, while no enhancement in working memory encoding attributable to statistical regularities is observed, participants to some degree extract the summary statistics of the attributes of targets across trials and these summary statistics can inform behavior.

43.309 LOCOMOTION MODULATES VISUAL WORKING MEMORY CAPACITY: HIGHER CAPACITY FOR SWING-PHASE ENCODING

Cameron Phan[1] (cpha4652@uni.sydney.edu.au), Erik Van der Burg[2], David Alais[1]; [1]University of Sydney, [2]University of Amsterdam

We investigated visual working memory (WM) during locomotion, comparing capacity during two phases of stepping: stance phase vs swing phase. In a virtual reality setting, participants walked along a straight level path while viewing a circular array of 20 discs in a head-mounted display. Discs changed polarity every 150 ms (some changing from black to white, others white to black) and after a random duration, some of the discs temporarily became green (1, 3, 5, 7, or 9 discs). This cued the locations to be remembered (the encoding event) and at the trial's end one of the 20 disc locations was probed and participants judged whether this was one of the memorized locations or not. The task was completed whilst the participants walked at a comfortable natural pace or walked slowly. Importantly, the encoding event was timed to occur either in the swing or stance phases of the participants' step cycle and the test probe was timed to occur at the same phase two steps later. Visual WM capacity was greater for encoding in the swing phase of the step cycle than during the stance phase. This WM capacity difference did not depend on walking speed. Response accuracy was greater for swing phase encoding, but individuals were more likely to respond during the stance phase. These findings expand on the impact of active movement on visual abilities and add to recent work by our lab showing that perceptual performance and reaction times modulate over time at a rate locked to the step cycle

43.310 BREAKING BINDING: INTERRUPTING THE CONSOLIDATION OF SOME OBJECT FEATURES BUT NOT OTHERS

Saeideh Saffar Tabbassi[1] (st1701@msstate.edu), Mary Catington[1], Michael Pratte[1]; [1]Mississippi State University

Visual perception requires that multiple visual features be perceived and bound together into coherent objects. This consolidation of visual information occurs over a brief period of time, but it is unknown whether features are bound together during consolidation or if they are first consolidated and then bound. For example, we show that color is consolidated within 100 ms whereas shape requires at least 300 ms, and the critical question is what happens when the consolidation of a colored shape is interrupted by a pattern mask at 200 ms. If binding is integral to consolidation, the entire object should be masked. Alternatively, if binding occurs only after feature consolidation, then the mask should interrupt shape processing while the object's color will still be perceived. To investigate binding when a mask interrupts the consolidation of some features of an object but not others, six colored shapes were presented briefly, and participants simultaneously reported both the color and shape of a cued and masked item. The consolidation time for the target item was manipulated, using stimulus-mask onset asynchronies ranging from 100 to 700 ms. The extent of binding at each consolidation time was measured using standard analyses of the joint distribution of color and shape errors. At longer consolidation times, the results replicate previous observations that the color of an object is sometimes reported accurately despite having no knowledge of the object's shape. Critically, when consolidation time was around 200 ms, the rate of these color-without-shape recalls increased, suggesting that the mask interrupted shape processing but spared color. This result shows that feature binding can be broken by interrupting the consolidation of some object features but not others, implying that binding occurs after the consolidation of features.

This work was supported by National Institutes of Health Grant R15MH113075

43.311 MORE EFFECTIVE ATTENTIONAL ALLOCATION WITHIN VISUAL WORKING MEMORY LEADS TO BETTER SUBSEQUENT LONG-TERM MEMORY PERFORMANCE

Sisi Wang[1] (wangsisi2017@gmail.com), Freek van Ede[1]; [1]Vrije Universiteit Amsterdam

Classical memory theories propose visual working memory (WM) processing as a key pathway for long-term memory (LTM) generation. Yet exactly how WM processes, such as attentional prioritization of specific mnemonic contents, affect LTM remains elusive. To address this, we employed a two-stage visual WM-LTM task, while tracking observers' eye movements. In the initial WM stage, participants completed a WM-prioritization task with a retrocue indicating the relevant visual object for an upcoming search task. Following a break, participants performed a surprise LTM-recognition task, identifying the object seen in the earlier WM stage. Results at the WM stage revealed more accurate and faster searching for informatively cued objects versus neutrally cued ones, replicating cue-induced attentional prioritization in WM. Crucially, at the LTM stage, we found significantly higher recognition accuracy for objects that were cued during WM (but absent in the search array) compared to both uncued objects and trials with a neutral cue (which showed similar recognition accuracy). These findings show that attentional prioritization during WM improves subsequent LTM performance for the prioritized item without incurring an LTM cost for unprioritized items (while having equal sensory exposure). In addition, through observers' eye-movements we could track attentional allocation after the retrocue. This revealed how attentional allocation during WM affected LTM, with trials with faster attentional allocating in WM having a larger subsequent benefit in LTM. Overall, these results show that attentional prioritization during visual WM affects LTM, and how this is mediated by the effectivity of attentional allocation – offering a mechanistic interpretation that bridges the interaction between WM and LTM.

This work was supported by a NWO Vidi Grant (14721) and an ERC Starting Grant from the European Research Council (MEMTICIPATION, 850636) to F.v.E..

43.312 PERSISTENT RESAMPLING OF EXTERNAL INFORMATION DESPITE TWENTY-FIVE REPETITIONS OF THE SAME SEARCH TEMPLATES

Alex J. Hoogerbrugge[1] (a.j.hoogerbrugge@uu.nl), Christoph Strauch[1], Tanja C. W. Nijboer[1], Stefan Van der Stigchel[1]; [1]Experimental Psychology, Helmholtz Institute, Utrecht University

We commonly load visual working memory minimally when to-be-remembered information remains available in the external world. In visual search, this is characterised by both encoding novel (externally available) templates during search and revisiting previously encoded templates. When all search templates have been encoded, subsequent rehearsal should lead to even stronger representations in memory, eventually making the (re)inspection of templates redundant. To test whether repetition indeed leads to decreased external sampling, participants (n=15) searched for 1, 2, and 4 search templates which remained available throughout trials. Critically, each unique set of search templates was repeated 25 trials consecutively. When presented with a novel template set, participants primarily encoded and searched templates one-by-one. Although the number of inspections and inspection durations initially decreased significantly as template sets were repeated, behaviour largely stabilised between the tenth and last repetition and participants still frequently resampled templates. In Experiment 2, participants (n=14) performed the same task, but could not inspect templates in the last 10 repetitions. Strikingly, accuracy remained high when templates could not be resampled, although the number of inspections decreased faster than in Experiment 1. This suggests that participants prepared for template removal by encoding templates more elaborately before the cut-off point. Across both experiments, resampling was linked to increased response times and no benefit to accuracy. Furthermore, participants could draw from long-term memory for the search task, as indicated by above-chance level recognition of template sets ten minutes after the main experiment finished. We thus argue that external sampling is not only used to offload effortful working memory maintenance, but also to boost confidence in existing (long-term) memory content. These findings demonstrate the persistence with which we sample information from the external world rather than rely on memory – even when we repeatedly need that same information.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 863732)

43.313 GAZE SCAN-PATHS AFFECT RECALL STRATEGY IN CONTEXT DEPENDENT MEMORY.

Neomi Mizrachi[1] (neomi.mizrachi@weizmann.ac.il), Ehud Ahissar[1]; [1]Weizmann Institute of Science

Motor-sensory dynamics is an essential component of visual perception. During an episode, the scanning patterns of the eyes affect the acquired information. To reckon eye movements' role in context dependent memory, we designed a recall task in virtual reality (VR) environments. Two groups were asked to explore a virtual room with 15 virtual daily objects; 20 minutes later they were asked to retrieve objects' names in either an environment that was similar (SIM group) or different (DIFF group) from the encoding environment. Only under similar context (SIM group), gaze scan-paths at recall were spatially-temporally similar to encoding scan-paths, such that the same locations were visited at the same relative times. Gaze spatial-temporal similarity dictated also recall strategy: SIM group participants retrieved objects' names according to the order of their gaze fixations at encoding and not by random or semantic connections. These results suggest that gaze scan-paths dynamics play a significant role in context dependent memory.

43.314 FORGETTING IN LONG-TERM MEMORY: RECOGNITION DOES NOT INDUCE THE FORGETTING OF SIMILAR OBJECTS

Jamal Williams[1] (jrwilliams@ucsd.edu), Timothy Brady[1]; [1]University of California, San Diego

Recent work has proposed that testing items in memory (e.g., a specific mug) causes the forgetting of related items (other mugs) when compared against a baseline of items from untested categories. Using behavioral studies and computational modeling, we challenge the view that active inhibition or suppression is responsible for the accuracy difference observed in recognition induced forgetting (RIF) studies. Across 6 experiments, we had participants encode items from 12 categories, with half of the items from half of the categories being restudied during a 'practice' session where participants had to discriminate an item from the initial encoding session against a new foil from the same category in a 2AFC task. Across all experiments, we replicate the RIF effect, although we find that the standard analyses in this literature inflate the effect size of RIF by conflating response bias with the true effect. Critically, we also find that participants have strong memories for the foils used in the practice tests, and that memory for these items drives the purported RIF effect. We demonstrate that the classic REM model of memory predicts RIF without any modifications—with no notion of inhibition or suppression— and can predict this effect simply based on the finding that memory for foils increases the set size of studied categories and induces cue overload. Importantly, using this insight we can create a reversal of the RIF effect: Showing worse memory for baseline items compared to non-studied items from studied categories. Our results suggest that differences observed in RIF studies do not stem from inhibition or suppression but are instead due to an inaccurate baseline, as restudy-foils are encoded and inadvertently increase the item count in studied categories thereby making them larger categories than baseline, non-studied categories, and impairing performance for all items—studied or not—in those categories.

43.315 CHARACTERIZING REPRESENTATIONAL DRIFT VIA A CONTEXT-DEPENDENT VISUAL WORKING MEMORY TASK

Yixin Yuan[1], Mikio Aoi[1], John Serences[1]; [1]UC San Diego, [2]UC San Diego, [3]UC San Diego

Recent studies with longitudinal neural recordings in mice suggest that neural activity associated with repeated stimuli gradually deviates from its initial pattern over time, even when task performance stays constant. This phenomenon, termed 'representational drift', challenges the common assumption that after fully learning a task, neural responses stabilize to support robust representations of task-relevant stimuli. Here we sought to determine whether representational drift occurs in human subjects while they are engaged in a visual working memory task, and to determine if drift is context specific. We recruited 4 participants (3F, 1M) to undergo a 5-session long fMRI study spread across a span of 3-5 weeks. During each session, the participants performed a two-alternative forced choice task that required recalling whether a probe object matched the target object shown prior to a 3-10 second delay period. Importantly, we manipulated the relevant stimulus components so that there were two task contexts: on half of the trials, the participants were instructed to remember the location and color of the object, while on the other half they were instructed to remember the location and shape. The participants were behaviorally trained and the task difficulty level was staircased so that accuracy stabilized at ~75% prior to the first scan. We then ran ridge-regularized circular regression on activation patterns in functionally determined regions-of-interest to predict the spatial location of the target object across trials. On average, within-session decodability of the target object's spatial location is higher than that of cross-session decoding. These results suggest that representations of spatial locations encoded in working memory drift from session-to-session, and this effect can be readily detected in the human cortex via fMRI.

43.316 CONTEXT AS A SCAFFOLD AND DETAILS AS BRICKS: NARRATIVE UNDERSTANDING AND UPDATING INFORMATION

Jayoon Choi[1] (sincerely95@g.skku.edu), Seongyun Kim[2], Minjae Jo[1], Min-Suk Kang[1,2]; [1]Sungkwunkwan University, [2]Center for Neuroscience and Imaging Research, Institute for Basic Science

When individuals perceive the real world, they actively maintain and update a representation of the current event as an event model. The event model can then be updated as those individuals take in and handle new information. We investigated how the brain serves the maintenance and modification of the event model while participants understand narratives of four short visual-audio clips in an fMRI scanner. In the initial session, participants watched only the visual stimuli of the four clips where sound was removed (visual encoding). In the second session, participants listened to only the sound extracted from the same, original clips (auditory encoding) and were instructed to integrate the new auditory information with the visual stimuli from the previous session. After completing the narrative comprehension task, participants were surveyed outside the scanner about their personal experience with the tasks. The survey indicated that the second encoding and recall were comparatively easier than the first encoding and recall across all stories. To identify brain regions sharing a common neural response among participants, we compared the inter-subject correlation of BOLD responses for the visual and auditory encoding conditions, respectively. Across all stories, the neural responses of the TPJ are similar across participants. More important, to identify regions maintaining information of the event model, we calculated intra-subject correlations between BOLD responses of the visual and auditory encoding conditions within each participant. We found a positive correlation for most stories in TPJ and PCC, indicating that the regions within the DMN play a key role not only in story integration but also in updating event models. In summary, participants demonstrated constructing a robust model during auditory encoding, aided by the event model formed during visual encoding. Together, neural results suggest that maintaining necessary information in the TPJ is instrumental in forming a richer event model.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2022R1A2C2007363).

43.317 DRAWINGS AND WORDS SHOW A RECALL ADVANTAGE OVER OBJECT PHOTOS

Anne Gilman[1] (anne.gilman@gmail.com), Cole Norberg[1]; [1]College of Saint Rose

People demonstrate an enormous recall capacity for photographs they have seen only once (Brady et al., 2008; Shepard, 1967; Standing, 1973). Using mostly drawings, Ally & Budson (2007) found a recall advantage for images over words in their groundbreaking EEG study of the picture superiority effect. Curran & Doyle (2011) replicated their findings also using line drawings, which were better recalled than words when using only words as test stimuli as well as when testing with images. Working memory research has shown bias differences in memory responses between line drawings and photographs (Gilman, Ware, & Limber, 2010). The present study used words at test, following the behavioral portion of Curran & Doyle (2011)'s Experiment 1, and contrasted photo, drawing, and word stimuli as study items. Forty-three undergraduates completed the study for course credit, with 40 providing responses. Stimuli consisted of IMABASE (Bonin et al., 2020) drawings of objects with naming agreement over 88%, related photos primarily from Konkle and colleagues' sets (2012, 2013), and the English name for each item. Roughly half of the stimulus items were randomly selected to be "new" items. Participants engaged in three randomized blocks of 36 study and 60 test trials. At study, stimuli were shown for 2000ms followed by a liking question; at test, words for items were shown for 1500ms to elicit recall ratings. Recall accuracy decreased by block (94%, 91%, 85%) and showed a significant advantage by TukeyHSD for drawings (94%) and words (91%) over photos (86%), $F(3,461) = 6.82$, $p<.001$. Neither sensitivity nor bias differed significantly by stimulus study format. The overall recall advantage for studied words and drawings over photographs conflicts with Curran & Doyle's 2011 behavioral findings. Image nameability may confer a recall advantage, similar to the memory advantage of nameable category differences found by Olsson & Poom (2005).

43.318 THE TEMPORAL DYNAMICS OF VISUAL SHORT-TERM MEMORY RETRIEVAL

Tianye Ma[1] (tma039@ucr.edu), Weiwei Zhang[1]; [1]University of California, Riverside

Recent research on short-term memory (STM) has revealed a combination of continuous and categorical reports in color recall, indicating a complex interplay between these memory formats. Nonetheless, the temporal dynamics of recruiting these STM representations during retrieval have remained unclear. The current study aims to elucidate how memory retrieval of continuous and categorical information unfolds over time by examining participants' computer mouse cursor movement trajectories during the recall phase of a continuous estimation task. First, with empirically derived color categories from an independent color naming task, the data from the recall task exhibited a between-trial mixture of categorical and continuous recall of colors, replicating some previous findings. Second, we partitioned the entire time course of the recall trajectories into different bins and fit a categorical-continuous mixture model to each bin. The estimated model parameters showed a gradual increase in the categorical memory over time, accompanied by an incremental shift towards categorical centers in the recall trajectories. Third, a formal model comparison indicated that a model with no categorical memory responses outperformed other models in the early stages of recall, whereas the categorical mixture model peaked towards the end of the trajectory. This pattern was replicated in another continuous estimation task with a different response method, ruling out an alternative account based on bias in the decisional stage. Overall, our data suggests that the retrieval of continuous color short-term memory precedes categorical color memory retrieval. This observation aligns with the notion that during memory encoding and retention, continuous representations undergo a gradual transition towards discrete stable attractor states, correcting accumulated errors across time.

43.319 CHARACTERISTICS OF SEQUENTIAL LEARNING AND MEMORY IN NON-HUMAN PRIMATES

Xuefei Yu[1] (yxdhhaa@126.com), Okihide Hikosaka[1]; [1]NIH

Sequential learning and memory are fundamental aspects of various daily tasks, including communication, navigation and executing a series of actions. Previous studies have demonstrated that monkeys can acquire motor sequence through trial-and-errors. However, the specific characteristics of how monkeys learn and maintain these sequences in memory across days remained unclear. In the current study, we focused on the learning and memory process involved in oculomotor sequences in non-human primates using a novel sequential saccade task. Two monkeys were trained to identify the 5-stages sequences by making consecutive saccades to one of the two targets in each stage. A 200ms blank screen period before each target's onset was introduced to assess anticipatory saccades to the next future location in the sequence. Throughout the training, both monkeys exhibited clear signs of sequence acquisition. They gradually learned the correct sequence within hundreds of trials, evident in increasing correct rates, and decreasing target acquisition times. Repetitive training across days led to increased proficiency in directing oculomotor sequences, reflected in a plateaued correct rate and decreased target acquisition time. Examing the time for each order in the sequence suggested changes of the sequence representation. During the early training days, the target acquisition time increased with the steps within one sequence, implying a serial representation of each step within the sequence. Such increasing trend gradually decreased after a few days' training, indicating the incorporation of the sequence as a whole. Anticipatory eye movements, signifying predictability of the next step in the sequence, emerged in most well-trained sequence. In summary, our results demonstrate the signs of sequence learning and memory during repetitive training. The study provides a paradigm with the potential to unveil the neural mechanisms underlying sequential learning and memory in non-human primates.

43.320 COMPARING ARTIFICIAL NEURAL NETWORK MODELS WITH VARIED OBJECTIVES TO PROBE THE ROLE OF SENSORY REPRESENTATION IN PRIMATE VISUAL MEMORABILITY

Ram Ahuja[1] (ramahuja@my.yorku.ca), Kohitij Kar[1]; [1]Department of Biology, York University

Imagine a typical experience of scrolling through numerous photos on social media. While most images blend into obscurity, certain ones, like a playful kitten tangled in yarn, linger in our memories. This selective retention, known as image memorability, raises intriguing questions about its neural basis, particularly how the sensory representations in the cortex facilitate this behavior. In this study, we investigate this process, leveraging recent strides in computer vision through artificial neural network (ANN) models. These models, designed to mimic the ventral visual pathway in primates, offer a vast hypotheses space for the brain function that are critical for determining the memorability of an image. We compare two distinct types of ANNs: models geared toward basic object categorization (e.g., ResNet-50, AlexNet, GoogLeNet) and those tailored to predict image memorability (e.g., MemNet, ViTMem) on a set of 200 images (20 images from 10 distinct object categories) from the MS-COCO dataset. Consistent with previous results, we observed that both these model classes can predict which images humans find most memorable. However, our results show that they produce significantly different internal representations as assessed by representational similarity analysis (comparing representations across model classes in architecture-specific and non-specific manner). In addition, here, we used neural data recorded across the macaque inferior temporal (IT) in 6 monkeys to address two primary questions. 1) Can both model classes accurately predict variance in the neural data? 2) Do they predict distinct variances in the responses of the individual neurons? Surprisingly, our results show that memorability models (MemNet) predict significantly more variance in the neural data than object categorization models with matched architecture (AlexNet) – emphasizing a) the need to further probe these model classes as encoding models of the ventral stream, b) provide a new normative framework to think about the evolution of sensory representations.

Google Research, CFREF, Brain Canada, SFARI

43.321 THE RISE AND FALL OF MEMORIES: TEMPORAL DYNAMICS OF VISUAL WORKING MEMORY

Andre Sahakian[1] (a.sahakian@uu.nl), Surya Gayet[1], Chris L.E. Paffen[1], Stefan Van der Stigchel[1]; [1]Utrecht University

Visual working memory (VWM) temporarily stores task-relevant visual information to enable interactions with the environment. VWM is typically studied in temporally rigid paradigms in which memory arrays (viewing time) and retention intervals (the delay to the probe) are determined by the experimenter, and are often kept constant. In everyday VWM use, however, there is huge variation in (1) how long we look at things to put them in memory, and (2) how much time passes before we act on the memory. The temporal variations in both viewing times and delays depend on many internal (e.g., strategy, resources) and external factors (e.g., stimulus availability, physical constraints). Here, we ask how the recall performance of VWM content develops across these two orthogonal temporal dimensions: how do memories build up over viewing time and how do they decay over increasing delay periods? We employed a copying task, in which participants were tasked to recreate an "example" arrangement of items in an adjacent empty "workspace". We tracked their unconstrained viewing and copying behavior at the level of individual items, recording how long items were viewed, and how much time passed before they were placed. Our results show that performance monotonically increased for viewing times up to one second (per item), but plateaued afterwards. Interestingly, while inspections exceeding one second did not improve performance for short (two second) delays much, inspections beyond one second did improve performance for longer delays. Our findings suggest that usable representations are produced quickly, while longer inspections make representations more resilient against decay. Likely, the natural variations partly arise due to strategic use of time and VWM resources in everyday behavior: when information needs to be applied immediately, a quick look should suffice, but when the use will be delayed, looking for longer might be worth the time investment.

This project was supported by an ERC Consolidator Grant [grant number ERC-CoG-863732] to Stefan Van der Stigchel, and a Veni grant from Netherlands Organisation for Scientific Research (NWO) [grant number: VI.Veni.191G.085] to Surya Gayet.

43.322 WHERE'S WALDO IN THE MIND: ACCESSING PERCEPTUAL AND SEMANTIC ATTRIBUTES IN PERCEPTION AND WORKING MEMORY.

Edyta Sasin[1] (edytasasin@gmail.com), Ying Zhou[1], Aytac Karabay[1], Sulav Shrestha[1], Daryl Fougnie[1]; [1]New York University Abu Dhabi

During perception, low-level features (such as color) are processed faster than high-level features (such as semantic properties). But what about accessing information from working memory? Recent work (Kong & Fougnie, 2021) has shown that search in working memory may be distinct from visual search regarding which features are most efficient. Further, research on long-term memory (Linde-Domingo, Treder, Kerrén, & Wimber, 2019) has shown that semantic information is retrieved more rapidly than perceptual information. However, it is not yet known whether semantic properties are accessed faster from working memory than perceptual attributes. In two experiments, participants were shown four images that were either animate or inanimate objects (semantic property) and which could be in the form of a photograph or drawing (perceptual property). Participants were pre-cued (perception – Experiment 1) or post-cued (working memory – Experiment 2) to the location of one of these objects. The cues were accompanied by either a semantic ("animate or inanimate?") or perceptual ("drawing or photograph?") question. Unsurprisingly, perceptual aspects were discriminated faster than semantic aspects when the information was available to visual perception. However, when the task required accessing no longer presented information from working memory, participants took less time to respond to semantic than perceptual queries. These experiments, together with other recent findings, point to a reversal of the processing hierarchy for perception and memory. While visual perception is feed-forward, retrieving information in memory might first involve accessing high-level properties such as semantic categories, followed by access to lower-level visual properties.

NYUAD Research Institute Grant CG012

Monday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

Visual Memory: Working memory and development, individual differences, capacity, resolution

43.323 DOES FORGETTING BENEFIT REMEMBERING IN WORKING MEMORY?

Daryl Fougnie[1] (darylfougnie@gmail.com), Ying Zhou[1,2], Anek Rajbhandari[1]; [1]New York University Abu Dhabi, [2]New York University

Only a limited amount of information can be maintained in working memory (WM). We have control over how this capacity is allocated–retrocue paradigms demonstrate that participants can prioritize a subset of information. However, the other side of this coin is less clear–can participants deprioritize or drop cued items from memory and does doing so improve memory for the remaining items? Previous studies have observed benefits of forgetting, but critically, in these paradigms deprioritized information is accompanied by increased priority in the other items, raising the possibility that participants are not benefiting from forgetting but translating forgetting cues into 'prioritize other' cues. For example, in Williams and Woodman (2012), participants were instructed to retain or forget half of the items. Both conditions improved performance, but the forget instruction took many trials to yield effects, consistent with participants learning to translate 'forget' into 'prioritize other'. We developed a novel paradigm that disentangles prioritize and forget cues. Specifically, subjects had to remember three orientations and report one of them after a brief delay. In each trial, they were either provided with only a prioritize cue indicating an item more likely to be tested, or both prioritize and forget cues (the latter indicating an item that would never be tested). Unlike previous attempts, forget cues did not increase the probability of other items being tested, due to the inclusion of no-report trials equal to the proportion that the dropped item would have been tested. In two experiments (differing only if cues were blocked or intermixed) we observed benefits of forgetting, but curiously these benefits were found for the neutral and not the prioritized item. These results are consistent with the idea that forgetting can benefit the remaining items in WM, but tentatively suggest that such benefits are not spread evenly among items.

NYUAD Research Institute grant CG012

43.324 REDUCING FAILURES OF VISUAL WORKING MEMORY WITH TAILORED FEEDBACK

Olga Kozlova[1] (olga.kozlova@mail.utoronto.ca), Kirsten Adam[1]; [1]Rice University

Although individuals can hold up to 3-4 items in visual working memory, they frequently experience severe lapses of performance (e.g., 0-1 items correct). An initial study found that performance feedback can improve working memory performance if the feedback encourages participants to perform consistently and avoid lapses (i.e., get 3+ items correct or lose points; Adam & Vogel, 2016). However, past work used a single performance goal that matched group-averaged capacity (3 items). Here, we hypothesized that performance feedback would be most effective when the performance goal is aligned with individuals' capacity. In Experiment 1 (N = 28), participants performed a whole-report memory task with performance feedback and either a well-aligned goal (get 3 items correct) or a poorly aligned goal (get 1 or 6 items correct). During each trial, participants first saw a reminder of the performance goal. Then, a memory array composed of six different colored squares was presented, and participants were instructed to remember the items over a blank delay. Finally, participants recalled each item's color and either gained or lost points depending on if their performance exceeded the performance goal (e.g., at least 3 correct). In Experiment 2 (N = 25), participants completed an identical task, but with a more fine-grained goal range (varying from 1–5). Across both experiments, we found that a performance goal tailored to working memory capacity (3 items) is most effective. Compared to sub-optimal goals, an optimal goal improved performance by 0.5 items. Moreover, in Experiment 2, tailoring the performance goal to an individual's specific capacity improved performance by 1 item relative to a sub-optimal goal. Performance was particularly poor when the goal was too high, with lapse rates increasing from 3% to 22%. Together, our results suggest that individualized goals can encourage optimal performance by targeting the frequency of lapses.

43.325 ACCESSING INDIVIDUAL DIFFERENCES ACROSS DIFFERENT DOMAINS OF SERIAL DEPENDENCE

Patrick kelly[1], Jefferson Ortega[1], David Whitney[1]; [1]University of California Berkeley

Many studies have found serial dependence in perceptual judgments of all kinds of features, objects, and abstract impressions. Fewer studies have investigated serial dependence for compound features and it remains less clear whether serial dependence can occur independently for different dimensions within a given stimulus. Here, we investigated this with combinatorial stimuli consisting of a skin lesion with a superimposed orientation texture. The two dimensions of the compound stimulus were orthogonal and randomly distributed. These stimuli were chosen because they are not configural objects like faces, are unfamiliar, can be approximately equated in terms of discriminability, and are translationally relevant to clinical settings. We first approximately equated discriminability of the two features. We then measured serial dependence in both domains using a 2 alternative forced choice task consisting of two separate feature judgments across two blocks of trials in which participants viewed the same set of stimuli. We found an overall significant level of serial dependence in both orientation ($p < 0.001$, permutation test) and malignancy ($p = 0.036$, permutation test) feature domains. However, there was not a clear relationship between the magnitude of serial dependence in the orientation dimension and that in the malignancy dimension at the level of individual observers ($r = 0.113$, $p = 0.616$). The results suggest that serial dependence can occur in different dimensions within the same object and that these effects may be independent of each other.

43.326 THE INFLUENCE OF VISUAL PERCEPTION ON WORKING MEMORY THROUGHOUT THE ADULT LIFESPAN

Shira Tkacz-Domb[1] (shira.tkaczdomb@email.gwu.edu), Sarah Shomstein[1], Dwight J. Kravitz[1,2]; [1]George Washington University, [2]National Science Foundation

Previous studies show bidirectional interference between visual working memory (VWM) and ongoing perception, consistent with the common recruitment of perceptual cortical areas (Teng & Kravitz, 2019) in tasks requiring the maintenance of visual detail. Here, we hypothesized that the strength of the maintained representation in perceptual areas will degrade with age. In two experiments, five groups of participants, divided by decade (ages 20 to 69), were asked to maintain a memory cue (Experiment 1: circle color; Experiment 2: Gabor orientation) while performing a detection task in which a sequence of letters appeared on a task-irrelevant background (Experiment 1: color distractor; Experiment 2: Gabor distractor). Following a white letter detection task, participants reported the maintained memory cue using a continuous report. Similarity between the maintained memory cue and the task-irrelevant background (0º, 15º, 35º, 55º, or 75º) was manipulated. For each similarity condition, we measured the bias of the response from the memory cue towards the task-irrelevant background and the standard deviation of the error distribution that conveys the quality of the representation. We replicated previous findings in both experiments, showing that perceived information affects memory representation. As the similarity between the memory cue and the distractor background increased, so did the bias and standard deviation. Additionally, in Experiment 1, which examined memory for color, we found a greater amplitude of bias toward the task-irrelevant background and a larger standard deviation with increasing age. Memory for orientation, however, did not interact with age. Together, these findings identify a very specific way in which VWM degrades with age, pointing to a stronger decrease in the fidelity of memory representations for color than orientation.

Research supported by NSF BCS-1921415 to SS, and BCS-2022572 to DJK and SS

43.327 HOW ATTENTIONAL CONTROL AND WORKING MEMORY CAPACITY PREDICT NATURAL MEMORY USAGE.

Dejan Draschkow[1] (dejan.draschkow@gmail.com), Levi Kumle[1], Joel Kovoor[1], Amy Li[1], Amelia Rock[1], Ellen Jones[1], Emily Hesselink[1], Anna C. Nobre[2]; [1]University of Oxford, [2]Yale University

Working memory (WM) has been studied extensively using highly controlled laboratory tasks. In typical WM tasks, to-be-remembered information is briefly flashed on a screen, followed by a delay period and eventually a memory probe. In contrast, recruiting WM in natural behaviour requires the continuous coordination of sampling information externally from the environment and relying on information internally from memory. That is, we are given the choice to self-determine when and how much to rely on memory to guide behaviour. Very little is known about the factors that influence how participants engage WM during natural behaviour. Here, we focused on two strong candidates for potentially predicting natural WM use: 1) the capacity of WM and 2) the ability to direct one's attention in a purposeful and goal-oriented manner (attentional control). The capacity of WM has been shown to vary substantially across individuals. Individual differences in WM capacity have also been linked to attentional control and both have been recognised as important determinants in performance across a wide range of complex behaviours. Using a large sample size of 100 participants, we demonstrate how WM capacity and attentional control influence the natural reliance on memory during a temporally extended virtual reality task. Our approach constitutes the first systematic exploration of the determinants of natural memory reliance during complex behaviour. The findings showcase how benchmarking the limits of cognitive functions relates to how these functions are being engaged in everyday life.

This research was funded by a Wellcome Trust (104571/Z/14/Z) and James S. McDonnell (220020448) Award to A.C.N. The Wellcome Centre for Integrative Neuroimaging is supported by Wellcome Trust (203139/Z/16/Z) funding. The work is supported by the NIHR Oxford Health Biomedical Research Centre.

43.328 WHITE MATTER MICROSTRUCTURE AND WORKING MEMORY OF MACAQUES IN ADOLESCENCE

Anna Machado[1], Junda Zhu[1], Clement Garin[1], Matthew McPhail[1], Adam Anderson[1], Finnegan Calabro[2], Beatriz Luna[2], Christos Constantinidis[1]; [1]Vanderbilt University, [2]University of Pittsburgh

White matter in humans and non-human primates continues developing from infancy through adolescence in parallel with gains in cognitive abilities. To investigate the association of white matter growth and working memory with developmental trajectories, we utilized diffusion tensor imaging (DTI) in eight monkeys by tracking changes in fractional anisotropy (FA) - a DTI metric measuring the diffusivity of water molecules in white matter tracts. A total of 86 DTI images were acquired using a Siemens 3T MRI scanner. Animals were scanned periodically over three years with identical acquisition parameters. Throughout their development, subjects were trained to perform an oculomotor delayed response task, requiring a memory-guided saccade after a 3 s delay period. DTI data was preprocessed and analyzed using MRtrix3 and FSL. Diffusion images from both phase encoding directions were combined, denoised, and susceptibility-induced distortion was corrected. After an affine transform registration of diffusion and structural images to a standard macaque template, FA values were extracted from 57 tracts using manual region-of-interest labeling. General additive mixed models with penalized smooth plate regression splines quantified non-linear associations between age/maturation and FA measures. We defined maturation age as the age of ossification of distal tibias in each of our subjects. The models included per-subject random intercepts and development slopes. Hierarchical differences of white matter development were shown, with developmental effects across almost all tracts, and no regions showing significant decreases in FA measures. Cerebellar tracts showed significant white matter growth from late childhood into adulthood (p<.05). During adolescence, association and projection tracts had yet to reach peak, significantly increasing after the maturation age (p<.01). In parallel, precision of memory guided saccades improved during this period of maturation. Collectively, our results indicate white matter maturation progressively develops from childhood through adolescence, reflecting gains in cognitive abilities.

Supported by NIH grants R01 MH117996 and R01 MH116675

43.329 CROSS RECRUITMENT OF HAPTIC REPRESENTATIONS DURING A VISUAL WORKING MEMORY TASK

Leah J Ettensohn[1] (lettensohn@ucsd.edu), Sunyoung Park, John T Serences; [1]University California San Diego

When looking at a familiar object, we can often imagine what it might feel like to the touch. Despite the multisensory nature of our interactions with common objects, working memory (WM) research has focused almost exclusively on unimodal stimuli. The current study investigated the recruitment of somatosensory representations during pure visual WM tasks as a function of haptic training with two classes of novel objects, Greebles and Fribbles. In a pilot experiment (N = 66), subjects completed a visual delayed match-to-sample (DMTS) task to confirm that memory performance for these novel objects was below ceiling (μ = 0.77, σ = 0.07). Next, 7 subjects completed a pre-test on the visual DMTS task, followed by haptic and visual training, followed by a post-test on the same visual DMTS task. During training sessions, the objects were divided into three sets: one set for haptic training without visual feedback, one set for visual training without haptic feedback, and one 'control' set of objects not trained. The results reveal improvement after training in both the visual condition and the cross-modal haptic condition, relative to the control condition. These results suggest that haptic training may improve performance even on a purely visual WM task, perhaps by recruiting additional representations of the remembered items in somatosensory cortex

43.330 SHARED POINTERS FOR BIOLOGICAL AND NON-BIOLOGICAL OBJECTS IN VISUAL WORKING MEMORY

Xinchi Yu[1,2] (xcyu@umd.edu), Ellen Lau[1,2]; [1]Program in Neuroscience and Cognitive Science, University of Maryland, [2]Department of Linguistics, University of Maryland

Behavioral and neural evidence suggests that human visual working memory (VWM) is implemented by a limited set (~ 3) of pointers, binding features for each object together respectively (Thyer et al., 2022). What has been more controversial is whether different visuo-cognitive domains – for example, systems engaging with biological entities vs. inanimate objects – share the same limited set of pointers, or maintain their own distinct sets. In the current study, we tested whether biological motions (BM) and complex shapes share the same set of pointers, by asking participants to memorize 3 complex shapes, along with 1 or 3 biological motions. We find that, holding more biological motions in the 3- vs. 1-BM condition does not affect Cowan's K for complex shapes when the probe is a new object (object-change trials) across two experiments (Experiments 1 and 2, N=24 each), but does reduce Cowan's K for complex shapes when the probe is an old object at a different location (location-change trials; Experiment 2, N=24). The interference effect in the location-change trials, which require binding object features to locations, supports the hypothesis that biological and non-biological objects share the same set of VWM pointers. The contrasting non-interference effect in the object-change trials suggests that similar previous results can be accounted for by the use of a simple feature familiarity strategy that doesn't require pointers for binding.

This work is supported by NSF #1749407 (to E.L.).

43.331 VISUAL WORKING MEMORY FOR CONFIGURAL INFORMATION

Inik Kim[1] (ikim124@ucr.edu), Weiwei Zhang[2]; [1]University of California, Riverside, [2]University of California, Riverside

Natural vision integrates both local and global environmental information, which finds representation and retention in visual working memory (VWM). Previous research has illustrated the nuanced manipulation of global and local representations within VWM. Notably, a configural cue's presence has been shown to enhance global VWM representations while leaving local representations unaffected. However, it remains unclear whether an invalid configural cue could also detrimentally impact global VWM representations. This study employs Xie & Zhang's (2017) orientation VWM change detection task to explore the effects of configural cues (invalid versus neutral) on VWM performance. The task involves connecting the centers of each orientation bar in a coherent (invalid cue condition) or random (neutral cue condition) manner, forming simple and complex polygons, respectively. We independently measure local and global VWM representations using the Xie & Zhang Dual Trace Signal Detection (DTSD) model (2017). Our hypothesis posits that the invalid cue will impair global VWM representations compared to the neutral cue. Preliminary data aligns with this prediction, underscoring the pivotal role of global representation beyond item-based representation in VWM. This research adds to the growing literature on the impacts of irrelevant contextual information on the dynamics of visual working memory.

Visual Memory: Capacity, long-term memory

43.332 STORING DYNAMIC RELATIONS INDUCES CONTRALATERAL DELAY ACTIVITY

Xu Jianzhe[1] (12139020@zju.edu.cn), Jiang Fei[1], You Chenyun[1], Xu Jiyu[1], Zhou Jifan[1], Chen Hui[1], Shen Mowei[1]; [1]Zhejiang University

The human mind can effortlessly construct mental representations defined by the intricate web of visual relations among its constituent objects and hold them in working memory for subsequent perceptual processing. To identify an online monitoring method for relation memory, a suitable event-related potential (ERP) index is required. Contralateral delay activity (CDA) stands out as a robust, feature-independent index of the content stored in visual working memory (VWM), and our hypothesis posits its ability to track the memory of visual relations. The current work adapted dynamic relation stimuli from Shen et al. (2021, Experiment 2) to test our hypothesis. Each dynamic relation was demonstrated by the movement of two objects governed by a specific kinematic equation. Up to four dynamic relations were sequentially displayed on either side of the screen, each lasting 1000ms, with a 200ms interstimulus interval (ISI) during the sequential display and a 1000ms ISI preceding the test phase. Participants were instructed to memorize one side of dynamic relations based on a memory cue and determine if the dynamic relation of a specified color pair changed in the test phase. We replicated prior behavioral findings indicating that the average upper limit of relation memory (measured with K-max) is two dynamic relations. The ERP data revealed a different pattern from the behavioral results: CDA progressively increased with the number of displayed relations, instead of reaching a peak after storing the second relation. Furthermore, we observed a positive correlation between the difference in CDA amplitude and individual working memory capacity. We concluded that: (1) the maintenance of dynamic relations induces CDA; (2) the VWM is capable of storing more than two dynamic relations, a result not evident in the behavioral findings, likely due to insufficient memory quality in the high load condition for performing the change-detection task.

This research was supported by National Natural Science Foundation of China [Grant 32071044; 32371088; 62337001], Science and Technology Innovation 2030-"Brain Science and Brain-like Research" Major Project (2022ZD0210800).

43.333 SPATIAL PROXIMITY AND OBJECT-BASED GROUPING EFFECTS ON VISUAL WORKING MEMORY

Naphapa Panyanirun[1] (naphaprin@docchula.com), Chaipat Chunharas[1,2], Timothy Brady[3]; [1]Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, [2]King Chulalongkorn Memorial Hospital, Thai Red Cross Society, Bangkok, Thailand, [3]University of California, San Diego

Performance in visual working memory is strongly influenced by how visual features are chunked across and within objects. For example, previous studies have demonstrated improved performance when items are presented spatially close to each other (Xu, 2006), as well as when multiple features are combined into the same object (Luck and Vogel, 1997; Oberauer & Eichenberger, 2013). However, limited evidence exists about the interaction between spatial and object-based grouping. To address this gap, we conducted a delayed match-to-sample experiment (N=71) to investigate the interaction of these two grouping cues on visual working memory. Participants had to remember the colors of 6 items grouped into pairs by their spatial arrangement or connecting lines. There were 4 grouping conditions: no grouping, 3 pairs grouped by spatial arrangement only, 3 pairs using object-based grouping only, and both cues. Sub-conditions for the latter included a congruent condition (when the two cues grouped the same pairs) and an incongruent condition (when the two cues conflicted). After a 1000 ms delay, one of the items was probed and participants had to choose between the correct color and a foil, which was always one of the adjacent items. We found that performance was generally better with either spatial or object-based grouping cues than with no cues. Interestingly, the combination of these cues did not yield any additive benefit, with spatial cues exerting a stronger influence in conflict situations. This finding aligns with the idea that attending nearby locations is more efficient than spreading attention through an object. Moreover, errors occurred more frequently when foils were outside the grouped pairs across all conditions, suggesting that the underlying mechanism involves reducing confusion within the group rather than between groups.

43.334 INFLUENCE OF INTRINSIC REWARDS ON WORKING MEMORY ALLOCATION

Ivan Tomic[1,2] (ivn.tomic@gmail.com), Paul Bays[2]; [1]University of Zagreb, [2]University of Cambridge

Visual working memory (VWM) has been characterized as a limited resource shared between visual stimuli. Previous studies demonstrated that this resource can be flexibly allocated depending on task demands, e.g. to prioritize an item that is more likely to be tested for recall. In these cases, the observed resource allocation has been interpreted as evidence for an optimal strategy that minimizes expected errors. In the present study, observers memorized two motion stimuli of different colours and later reproduced the motion direction of one cued item. In the first experiment, we manipulated the feedback of the true motion given at the end of each trial, artificially increasing the error for items of one colour and decreasing it for the other. Modelling resource sharing in VWM showed that the error-minimizing strategy was to allocate more resources to the items for which feedback error was magnified. Surprisingly, we found strong evidence that observers instead allocated more resources to the stimuli for which the feedback error was reduced, i.e. those perceived as easier to remember. This preference for 'easier' items was confirmed in a second experiment where we manipulated the true difficulty by using different motion coherences for the two colours. Two further experiments tested optimal allocation in contexts with previously documented unequal allocation, specifically, variable cueing probability and variable rewards. Despite replicating previously observed patterns, modelling again revealed suboptimal resource allocation. Altogether, these results challenge previous assumptions about how resources are shared in VWM, and suggest an implicit bias in allocation towards stimuli associated with lower estimation uncertainty, which observers may experience as intrinsically rewarding. We show that such a bias provides an alternative account of seemingly error-minimizing strategies observed in previous studies, while also being consistent with the broader literature on human preferences for rewarding stimuli.

43.335 SHARED VISUAL MEMORY RESOURCES FOR DYNAMIC AND STATIC STIMULI

Garry Kong[1], Lisa Miyahara[1]; [1]Waseda University

Visual working memory stores the current state of the world for use in other cognitive processes. As the world is a dynamic place, visual working memory therefore needs to be able to store dynamic information. Despite this, visual working memory is typically studied with only static stimuli. Are dynamic stimuli simply stored as series of static stimuli? How many visual working memory resources does a dynamic stimulus take compared to a static one? Are they even stored in the same system? Here, we used a dual-task paradigm to investigate these questions. Participants saw either a 100% valid cue indicating whether the test item would be dynamic and static, or a neutral cue. They then viewed a static stimulus, consisting of either 2-6 circles randomly located on the circumference of a circle (Experiment 1) or 1-4 colored circles (Experiment 2), followed by the dynamic stimulus, consisting of a 5 second sequence where a trail of white dots moved between two random points on the perimeter of a large circle, changing directions 3 times along the way. Finally, they were asked to reproduce either the static stimulus (location in Experiment 1, color in Experiment 2), or the dynamic stimulus, by reproducing the path that trail of white dots took using a touchscreen. Overall, we found having to remember a dynamic stimulus decreased memory recall performance for a static stimulus, and vice versa. Importantly, although the dynamic stimulus was predominantly spatial in nature, it was still impacted by the static color stimulus. Interestingly, despite the difference in amount of information inherent in a dynamic stimulus, the effect on static memory was equivalent to having to remember two extra locations, or one extra color. We conclude that the same visual working memory system is used to remember both static and dynamic stimuli.

This work was supported by the Japanese Society for Promotion of Science Kakenhi Grant Number JP22K13874 and Waseda University Grant for Special Research Projects (2023C-582).

43.336 SIMILARITY-DEPENDENT MEMORY INTEGRATION OF SCENE IMAGES

Simeng Guo[1,2,3,4,5], Sheng Li[1,2,3,4,5]; [1]School of Psychological and Cognitive Sciences, Peking University, Beijing, China, [2]PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China, [3]Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China, [4]Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China, [5]National Key Laboratory of General Artificial Intelligence

People often encounter novel events similar to their previous experiences. An intriguing question is how the similar representations interact during learning. Previous studies suggested that feature-based similarity resulted in systematic memory distortions. The present study examined the effect of memory integration due to learning similar scenes. We used generative adversarial networks (GANs) to generate scene wheels from which the to-be-remembered scenes were selected. In an online experiment (n = 59), we evaluated the perceptual similarity of images from the scene wheels and selected scene-pairmates (A1 and A2) with varying perceptual similarities. In three main experiments (n = 27, 27, 28), A1 and A2 were paired with different images (B1 and B2) to form competitive associations. Subjects learned these associations with explicit knowledge that scenes paired with different images were always different (even though they might look similar). Importantly, learning of competitive associations was temporally separated ("A1-B1" preceded its competitor, "A2-B2"). Across three experiments, we found robust attractive memory distortion of A2 towards its highly-similar competitor (A1). In Experiment 1 and Experiment 2 (with increased training on A2-B2), the attraction effects were asymmetric: memory of A1 was not biased relative to A2. Interestingly, in Experiment 3 in which training on A1-B1 was increased, the asymmetry disappeared: memories of A1 and A2 were biased towards each other. Moreover, we examined the consequences of the distortions. As expected, attractive distortions decreased discriminability between highly-similar associative memories. We unified these findings using a Hebbian learning framework and suggested that (1) greater coactivation between B2 and A1 as compare to the coactivation between B1 and A2 caused asymmetric integration, and (2) the balanced coactivations eliminated such asymmetry. Collectively, we showed that similarity-dependent integration of complex visual experiences might cause asymmetric memory distortion. The degree of the asymmetry depends on the level of coactivation during integration.

43.337 CAN YOU ENHANCE VISUAL LEARNING WITH STIMULATION OF THE MEDIAL-FRONTAL CORTEX?

Gengshi Hu[1] (gengshi.hu@vanderbilt.edu), Geoffrey Woodman[1]; [1]Vanderbilt University, Department of Psychology, Vanderbilt Vision Research Center

Previous research suggests that applying anodal transcranial direct-current stimulation (tDCS) to the medial-frontal cortex can improve how quickly subjects learn to make simple discriminations (i.e., red from blue). Here, we tested the idea that this superior learning is due to superior encoding into the long-term memory of images. Thirty subjects completed an anodal stimulation session and a sham session, with order counterbalanced across subjects, before performing recognition-memory tasks using pictures of real-world objects and visually presented words or nonwords. These tasks allowed us to detect potential memory differences across types of memoranda. Contrary to the hypothesis that the medial-frontal cortex helps control encoding veracity, we found that 20 minutes of tDCS at 2.0 mA did not significantly improve participants' memory, regardless of stimulus type, JZS Bayes Factors < 0.4. Our findings show that although medial-frontal cortex stimulation can change how quickly we learn stimulus-response mappings, this is not simply due to superior memory for the items.

Grants were provided by the National Science Foundation (BCS-2147064), and NEI (P30-EY08126).

43.338 ALPHA/BETA OSCILLATIONS TRACK ORIENTING ATTENTION TOWARDS LONG-TERM MEMORY REPRESENTATIONS

Rhianna Watt[1] (rhianna.watt@psy.ox.ac.uk), Bo-Cheng Kuo[2], Nahid Zokaei[1], Anna C. Nobre[1,3]; [1]University of Oxford, [2]National Taiwan University, [3]Yale University

Many studies have investigated the orienting of internal attention towards information maintained in short-term memory (STM). However, how internal attention can be oriented towards information stored in long-term memory (LTM) remains unclear. Here, we investigated whether orienting attention to selective contents of LTM representations modulated oscillatory activity in the alpha (8-14 Hz) and beta bands (15-30 Hz) in an analogous fashion to internal attention orienting in STM. Twenty-nine participants learned a series of spatial-contextual associations between images of scenes and everyday objects over two days. On the third day, we recorded EEG from the participants while they performed delayed-response tasks with LTM or STM information. Spatial retro-cues after the initial scene array indicated the location of the object that would be probed at the end of the trial, or neutral retro-cues provided no spatial information. At the end of the trial, a probe object appeared, and participants decided whether it was associated with that scene. In the LTM task, participants viewed empty scenes they had previously studied and oriented their attention to previously associated objects. In the STM task, participants encoded new scenes containing two new objects and oriented their attention to scene objects maintained in STM. Our behavioural data showed a significant cueing benefit for valid cues over neutral cues for both tasks. When comparing the alpha and beta power between valid cues versus neutral cues, we found a similar pattern of alpha/beta power attenuation between STM and LTM tasks, with a stronger alpha power attenuation for the LTM task than the STM task. Finally, we showed a similar alpha-power lateralization for both tasks, with more decreased power contralateral than the power ipsilateral to the cued hemifield. Our results showed that alpha/beta oscillations track internal attention not only in STM but also when information is reactivated in LTM.

This research was funded by a Wellcome Trust Investigator Award (104571/Z/14/Z) to A.C.N.

43.339 INFLUENCE OF FAMILIARITY, AESTHETIC VALUE, AND CHANGE TYPE ON VISUAL MEMORY OF AI-GENERATED PAINTINGS IN A VIRTUAL REALITY CHANGE/NO CHANGE PARADIGM

Courtney G. Nutt[1] (cgnutt0361@ung.edu), Tanner L. Lumpkin[2], Patsy E. Folds[3], Erin A. Conway[4], Ralph G. Hale[5]; [1]University of North Georgia

Human visual memory capacity and limitations of these systems have been highly investigated, often highlighting the impact these factors have on our cognitive abilities and biases. Many visual memory errors relate to a reliance on gist information instead of precise details. These errors can have real-world implications (e.g., eyewitness testimony). Previous research has suggested familiarity and aesthetic value of visual information are likely to affect semantic engagement, strength of encoding, and veridical memory accuracy. In the present study, participants were immersed in a VR environment and asked to study 36 AI-generated paintings. All stimuli consisted of paintings displayed on the wall of an art gallery, with half mimicking famous paintings and half representing novel paintings. After each image, participants rated the item's familiarity and aesthetic value on an 8-point scale. A VR distractor task followed the study block to prevent items from being held in working memory. The test block consisted of half identical and half changed trials. Painting change options were color changes, inversions, and new images of the same style. Change/no-change data was collected at test. As expected, participants found the paintings that mimicked familiar paintings to be more familiar than the novel paintings. Interestingly, the novel paintings were found to be more aesthetically pleasing overall. However, neither familiarity nor aesthetic value significantly impacted the overall accuracy of detecting changes. Also, no significant differences in accuracy were found between change types. While inconsistent with previous research stating the importance of top-down factors like familiarity, these findings are consistent with models suggesting differences in visual memory ability for, and errors related to, gist versus verbatim information. By testing these factors in a more ecologically valid VR environment, this study improves our understanding of visual LTM and factors that contribute to accurate encoding, particularly related to familiarity and aesthetics.

43.340 INFLUENCE OF FAMILIARITY, AESTHETIC VALUE, AND CHANGE TYPE ON VISUAL MEMORY OF REAL-WORLD SCENES IN A VIRTUAL REALITY CHANGE/NO-CHANGE PARADIGM

Tanner L. Lumpkin[1] (tllump9771@ung.edu), Courtney G. Nutt[2], Patsy E. Folds[3], Erin A. Conway[4], Ralph G. Hale[5]; [1]University of North Georgia

The capacity and limitations of visual long-term memory (VLTM) has been investigated for decades. The process of VLTM encoding is not perfect, with a bias toward gist over verbatim information exacerbated due to time and other factors. For instance, familiarity with a visual stimulus has been shown to improve STM capabilities, thereby improving consolidation and LTM storage. Aesthetics also may play an important role in VLTM encoding. Subjective feelings regarding the aesthetic value of a stimulus could influence ability to remember certain details. In this study, we investigated how VLTM of indoor, cityscape, and landscape scenes were impacted by familiarity, aesthetic value, and change type in a virtual reality environment. Participants were exposed to 36 scenes. For each scene, participants used an 8-point scale to rate familiarity and aesthetic value. Afterward, a distractor task was used to ensure items were not being held in working memory. At test, participants were shown 36 scenes; half were identical to study and half were changed. Changes included color changes, inversions, or new images of the same style. Change/no-change responses were collected after each scene. A significant main effect of scene type was found for familiarity and for aesthetics, indicating notable and consistent differences in familiarity and aesthetic judgement between indoor, cityscape, and landscape scenes. Additionally, a main effect of change type was found, with a significant interaction between change type and scene type. Interaction appears to be driven by individual stimulus characteristics rather than being caused by change type. Surprisingly, accuracy was not systematically impacted by familiarity, aesthetics, or scene type. This study is the first of its kind to study VLTM in a immersive VR setting, providing new insight into factors that influence VLTM mechanisms in a more real-world setting.

43.341 THE BRAIN KNOWS MORE IS STORED IN VISUAL LONG-TERM MEMORY THAN WE CAN REPORT

Geoffrey Woodman[1] (geoff.woodman@vanderbilt.edu), Chong Zhao[2], Keisuke Fukuda[3]; [1]Vanderbilt University, [2]University of Chicago, [3]University of Toronto Mississauga & University of Toronto

If your visual long-term memory storage is virtually perfect, then when you happened to press the wrong button during a visual recognition memory experiment you would detect that motor error as such. That is, your brain should know when you are about to commit an error during a visual long-term memory experiment, even before you press the button. This is exactly the prediction we tested in this study. Electroencephalograms were recorded from 50 subjects who viewed 500 photographs of real-world objects for a subsequent recognition memory test in which the confidence of their judgments were also collected. Here we show that subjects' brains knew they were making mistakes during the recognition memory test at the time of button press, indicating that sufficient memory representations were available to executive control mechanisms to know that errors were being committed. Our findings show that people's choice behavior in visual long-term memory experiments misses information that is available to other high-level mechanisms of the brain.

Grants were provided by the National Science Foundation (BCS-2147064), and NEI (P30-EY08126 and T32-EY007135).

43.342 PERCEPTUAL COMPARISONS ARE NECESSARY AND SUFFICIENT FOR THE PERSISTENCE OF MEMORY BIASES ACROSS TIME

Joseph M. Saito[1] (joseph.saito@mail.utoronto.ca), Susanne Ferber[1], Morgan D. Barense[1], Keisuke Fukuda[1,2]; [1]University of Toronto, [2]University of Toronto Missisauga

Interactions between visual memories and percepts have been shown to induce systematic biases in observers' memory reports. When memories are explicitly compared with percepts, the resulting biases are potent enough to persist long after they are initially reported. However, it remains unclear whether bias persistence is attributable to processes that occur during perceptual comparisons themselves or those that occur when the bias is read out during initial memory reporting. To test this, we asked observers to encode colored object silhouettes into their long-term memory in anticipation of memory testing that occurred immediately after the encoding phase and 24 hours later. At each test, observers were cued to recall a target object from memory by presenting the uncolored target silhouette. Following recognition, the probe silhouette was then re-presented in a color that was sampled proximal to the encoded target color and observers were instructed to either ignore this colored probe or to judge its similarity to the remembered target. Observers then completed the trial by either reporting the remembered target color from a continuous wheel or by completing a search for an uncolored letter. Critically, all objects tested during the immediate test were tested again during the delayed test in a report condition where colored probes were omitted. Within each test, observers' target reports showed attractive biases towards the colored probes, with larger biases following comparisons than passive viewing. More importantly, comparison-induced biases persisted into the delayed test with comparable magnitude across reported and unreported targets, while biases induced by passively viewing the probe dissipated by the delayed test, even for targets that were initially reported. These findings suggest that processes tied to perceptual comparisons are both necessary and sufficient for the formation of report biases that carry over from one retrieval episode to the next.

This work was supported by an NSERC Discovery Grant awarded to KF (RGPIN-2017-06866)

43.343 UNVEILING THE EARLY IMPACT OF STIMULUS MEMORABILITY ON VISUAL SHORT-TERM MEMORY FORMATION

Chaoxiong Ye[1,2] (chaoxiong.c.ye@jyu.fi), Lijing Guo[1], Nathan Wang[3], Qiang Liu[2], Weizhen Xie[4]; [1]University of Jyvaskyla, Finland, [2]Sichuan Normal University, China, [3]John Hopkins University, [4]University of Maryland

Human observers consistently exhibit remarkable recall for certain stimuli, such as, face and scene images. This phenomenon is attributed to the concept of stimulus memorability, which is thought to represent an ensemble perceptual quality that enhances the long-term retention of visual information. Nevertheless, a comprehensive understanding of how this stimulus property contributes to the formation of visual memory remains unclear. In this study, we aim to investigate the hypothesis that memorability imparts a constant perceptual advantage in the establishment of visual short-term memory (VSTM). Given the inherently probabilistic nature of visual perception, this hypothesis posits that memorable information should increase the likelihood of an item being encoded into VSTM during the critical period when fragile sensory information is transferred into durable VSTM. To test this hypothesis, participants were tasked with memorizing images of three unfamiliar faces, each characterized by a specific level of memorability (memorable vs. forgettable). Memory performance was assessed following a brief delay using the change detection paradigm. Memory stimuli were consistently presented for 150 milliseconds, and the memory-and-test stimulus onset asynchrony (SOA) remained fixed at 1,600 milliseconds. However, we introduced a 200-millisecond consolidation mask at various memory-and-mask SOAs to manipulate the time available for further encoding of these stimuli into VSTM. Our findings reveal that participants consistently exhibited better recall for memorable faces when compared to forgettable ones across all masking conditions. These results suggest that the facilitation of VSTM by memorability manifests early (≤150 ms), distinguishing it temporally from the impact of stimulus familiarity on later stages of VSTM formation, as observed in some prior studies (>330 ms). Future research should aim to differentiate this early memorability effect from the later familiarity effect, with the ultimate goal of unraveling the intricate processes underlying VSTM formation within the context of naturalistic vision.

This study received funding support from the Academy of Finland (355369) to Chaoxiong Ye. Weizhen Xie is supported by the NIH Path to Independence Award (K99NS126492).

Color, Light and Materials: Appearance, categories

43.344 CATEGORIZATION AND NAMING OF SURFACE TEXTURE AND COLOR

Angela M. Brown[1] (brown.112@osu.edu), Delwin T. Lindsey[1]; [1]Ohio State University

According to the Sapir-Whorf hypothesis ("Sapir-Whorf"), language controls categorical perception. Thus, Sapir-Whorf predicts that terms should be associated easily and fluently with the categories they name, and well-understood categories should have high-consensus terms. ***We studied 99 physical samples of surfaces in an incompletely crossed design. Each sample had a color (red, orange, yellow, green, turquoise, blue, purple, black, white, or polychromatic) and a surface texture (smooth, transparent, metallic, ridged, nubbly, bumpy, rough, felt, burlap, vinyl upholstery, glitter or holographic). Sixteen subjects performed unconstrained sorts of the samples, based on sample appearance, then they provided a single term for the shared quality of the samples in each group. Four people sorted by color, eleven sorted by texture, and one sorted by both texture and color. ***People who sorted by color agreed what the 9–10 color categories were (90% agreement within categories), and they provided color terms for their categories quickly and easily. The color categories and the 14 color terms were nearly 1:1, and pairs of individuals who sorted by color agreed on the color term for 87.4% of samples (SEM=±3.4%). ***People who sorted by texture also agreed what the 11–16 texture categories were (95% agreement within categories), but they provided 100 different texture terms slowly and with difficulty. Words like "shiny" and "woven" named several different textures, and pairs of individuals who sorted by texture agreed on the texture term on only 11.7% ±3.3% of samples. ***There is a long literature on the difficulties with applying Sapir-Whorf to color terms, and this study does not strongly confirm or refute it for color. However, the Sapir-Whorf hypothesis clearly cannot apply to the understanding of the surface texture of physical samples, because texture naming was difficult, the association between texture terms and categories was complex, and observers' texture textures showed low consensus.

43.345 VISUAL SEARCH FOR WARM AND COOL COLOURS

Jake Manalansan[1] (jmanalansan@nevada.unr.edu), Camilla Simoncelli[1], Michael Webster[1]; [1]University of Nevada Reno

Distinctions of warm vs cool are considered fundamental aspects of colour experience, though their status as a perceptual dimension remains unclear. Using a colour search task, we examined the relative salience of warm and cool colours relative to other hues. The search task was similar to McDermott et al. (JOV 2010) and involved searching for a 0.5° circular "fruit" target with variable chromaticity, superimposed on a dense 11.6° by 20° "foliage" background of random ellipses. The ellipses' chromaticity varied along fixed axes in the cone-opponent colour plane that corresponded approximately to warm-cool (orange to greenish-blue), blue-yellow, or the two orthogonal chromatic directions (magenta-greenish). Targets on each background spanned a fixed range of hue angles and contrasts, and could be compared across backgrounds by how far the target colour was from the background axis. Search times for the targets were faster for detecting chromatic deviations from the warm-cool or blue-yellow backgrounds than the orthogonal directions. This suggests that both the warm-cool and blue-yellow dimensions are less salient than the orthogonal directions, consistent with biases found for a single bluish-yellowish condition tested previously by McDermott et al. However, the biases trended toward stronger asymmetries for the warm-cool axis. Interestingly, there were no differences in search times for warm vs cool targets or for blue vs. yellow, even though cool/blue hues may be more associated with shading or backgrounds while yellow/warm hues with objects. Overall the results suggest that sensitivity is weaker for the warm-cool dimension than for orthogonal dimensions of colour space, possibly because of selective adaptation to warm-cool and/or blue-yellow biases in the colour environment. This is also consistent with our recent analyses showing that the warm-cool dimensions closely align with the asymmetries in sensitivity predicted by uniform colour systems (Manalansan and Webster VSS 2023).

Supported by EY-010834

43.346 THE "SUPER-IMPORTANCE OF HUE" IN PSYCHOPHYSICS, PHYSIOLOGY, AND AI

Laysa Hedjar[1] (laysa.hedjar@psychol.uni-giessen.de), Takuma Morimoto[1,2,3], Matteo Toscani[4], Arash Akbarinia[1], Mandy V. Bartsch[5,6], Hendrik Strumpf[7], Jens-Max Hopf[6,7], Karl R. Gegenfurtner[1]; [1]Justus-Liebig-Universität Gießen, Germany, [2]University of Oxford, [3]Physics Center of Minho and Porto Universities (CF-UM-UP), Braga, Portugal, [4]Bournemouth University, Poole, UK, [5]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Netherlands, [6]Leibniz-Institute for Neurobiology, Magdeburg, Germany, [7]Otto-von-Guericke-Universität Magdeburg, Germany

Hue is thought to play a more important role for color discrimination than chroma. However, whether this is true appears to depend on the region of color space. In physiologically-based color opponent spaces, previous literature has shown that for orangish and bluish colors, hue thresholds are much smaller than chroma. But for purplish and greenish colors, hue and chroma thresholds are nearly equal. It is unclear why the visual system prioritizes some mechanisms over others. We first present neural evidence of this effect: we took magnetoencephalography measurements while subjects performed a simple discrimination task with either purplish or orangish uniform discs. We found greater amplitude modulation – indicating better discriminability – when the orangish stimuli differed in hue compared to chroma, and compared to both hue and chroma for purplish stimuli. Behaviorally, we explored whether the psychophysical differences would also arise in color discrimination of single-hue rendered objects, which elicit a distribution of points in color space. We found thresholds were merely elevated for the rendered objects compared to single patches of light. Hue superiority was present only for orangish colors, not purplish. We then analyzed multiple image databases to see whether the color distribution of objects found in the environment was biased and found that they overwhelmingly plot in the orangish regions of color space. We used the linear probe technique to interpret the internal representation of several deep neural networks trained on such biased image sets with high-level visual tasks such as object recognition and text-image pairing. The pattern of thresholds estimated from the networks, in particular the hue-chroma asymmetry, was similar to humans. We conclude that hue is indeed of superior importance for color discrimination, and that the peculiarities of this are shaped by the color statistics of our environment.

LH, KRG: ERC Advanced Grant Color 3.0 (No. 884116). TM: Sir Henry Wellcome Postdoctoral Fellowship (Wellcome Trust: 218657/Z/19/Z) and Junior Research Fellowship (Pembroke College, University of Oxford). AA, KRG: DFG, Germany SFB/TRR 135 (No. 222641018) TP Z. MVB, HS, JMH: DFG, Germany CRC1436/B05.

43.347 COMPARISON OF HUE DIFFERENCES MEASURED BY PERCEPTION VERSUS VISUAL EVOKED POTENTIALS

Jesse R. Macyczko[1] (jrmacyczko@gmail.com), Mohana Kuppuswamy Parthasarathy[1], Osman B. Kavcar[1], Michael A. Crognale[1], Michael Webster[1]; [1]University of Nevada, Reno

Despite the continuous variation in wavelength or cone excitations, colors tend to group into a small number of categories, and the ranges of these categories can vary widely. For example, "yellow" occupies a relatively narrow range of wavelengths or stimulus angles in a linear cone-opponent space compared to "green" or "blue." The nature of the size differences between color categories and where they arise in visual processing remains poorly understood. We compared the magnitude of hue differences as measured behaviorally or with steady-state Visual Evoked Potentials (VEPs), recorded from a single electrode placed on Oz. The stimuli were hue angles of equal chromatic contrast in a scaled LvsM or SvsLM cone-opponent plane, and differed over a range from 0 to 50 deg (corresponding to hues ranging roughly from blue-green to yellow-green for the green set and yellow-green to red for the yellow set). Psychophysical measurements confirmed that the difference between two hue angles that straddled an observer's unique green is perceptually smaller than the same angular difference around their unique yellow – i.e. equivalent stimulus differences in the green region appear more similar. For VEPs the hue angles were presented as 1 c/deg gratings (from gray to the maximum chromaticity) in a tapered 5 deg window. The two angles forming a pair were alternated at a rate of 6 images/sec, with the response amplitude extracted from the 6 Hz component. Responses increased with increasing difference in the alternated angles, but unlike the perceptual responses, the amplitudes were similar for the differences in the green and yellow regions. These results are consistent with the perceptual range of different hue categories arising at later stages of cortical color processing, beyond primary visual cortex.

Supported by EY-010834

43.348 SURROUND EFFECTS ON COLOR CONSTANCY IN VIRTUAL REALITY

Raquel Gil Rodriguez[1] (raquel.gil-rodriguez@psychol.uni-giessen.de), Laysa Hedjar[1], Matteo Toscani[2], Dar'ya Guarnera[3], Giuseppe Claudio Guarnera[3], Karl R. Gegenfurtner[1]; [1]Justus-Liebig University Giessen, [2]Bournemouth University, [3]University of York

The degree of color constancy can be influenced by the color of an object's surroundings. We used virtual reality (VR) to observe how different color surroundings affect a target object under different illuminations. VR technology allowed us to easily modify the illumination and the reflectance properties of objects in a photorealistic indoor scene. The scene was rendered with Unreal Engine under five illuminants: a neutral illuminant metameric to D65, two more illuminants also on the daylight locus and two orthogonal to it. A color calibrated HTC Vive Pro Eye headset was used for display. The target object was presented on top of a colored surround, whose color was chosen to lie in between two of the illuminant directions in CIELab color space: kaki, purple, rose and teal. We measured the degree of color constancy for each surround color under each one of the five illuminants. Participants were asked to select the target that best matched an achromatic one, from a set of five differently colored competitors present in the scene. These five were chosen to include one that represented perfect constancy (matching the achromatic reflectance), one that represented zero constancy (matching the tristimulus value), two that were sampled in between reflectance and tristimulus, and one that represented overconstancy. We found that surround and illumination color interacted in their effects on constancy. For example, the purple surround under a greenish or yellowish light source – surround and illumination color in nearly opposite directions – led to a decrease in CCI. Similar effects were observed for other combinations of illumination and surround. Our results indicate that illumination and surround jointly affect color constancy and appearance.

43.349 SWIPING COLORS IN VIRTUAL REALITY: COLOR CATEGORIES IN ACTION

Avi M. Aizenman[1] (avigael_aizenman@berkeley.edu), Zoe Goll[2], Karl R. Gegenfurtner[3]; [1]Giessen University

We adapted a paradigm from animal learning to investigate the stability of color category borders in humans using a VR videogame task. Observers held a colored saber in each hand and swiped approaching cubes which contained a colored stripe. Observers were instructed to use the saber whose color best matched the colored stripe. Saber colors were green and blue, or pink and purple, and the cube colors varied smoothly in fixed multiples of discrimination threshold. In a baseline block, observers were tested on a predetermined set of colors, where three of the in-between hues were ambiguous and close to the category border. We fit the saber choices with a psychometric function to determine the location and sharpness of the category border. Subsequent blocks shifted the tested color range toward one endpoint, and if observers' color category borders were stable, there would be no difference between the baseline and shifted borders. Alternatively, observers could base their responses on the color difference between the cube and the saber only. In that case, the PSE would shift in the same direction as the shift in the colors tested. Our results show that observers exhibit a halfway shift of their category borders in the direction of the saber color shift. In follow-up studies, we found that this partial range effect persists even when equalizing the proportion of responses made with each saber color. We also found a comparable adaptation to the range when using green hues without a category border. This work suggests a very limited role of color categories for our task. We speculate that observers learn the task and quickly become adept at performing the match to sample task. This expertise may allow observers to respond automatically instead of focusing on the categorical distinction of the hues.

Alexander von Humboldt Fellowship to author AMA

43.351 SEEING THROUGH ANOTHER'S EYES: MODELING AND CORRECTING FOR INDIVIDUAL DIFFERENCES IN COLOR APPEARANCE

Camilla Simoncelli[1] (camilla.simoncelli@gmail.com), Michael Webster[1]; [1]University of Nevada

Individual differences in color vision arise at many levels, from the spectral sensitivities of the cones to how individuals judge or label color appearance. Peripheral sensitivity differences are routinely corrected (e.g. to control for equiluminance) and there are growing efforts to calibrate displays and create standards that account for the spectral sensitivities of individual users. However, these sensitivity differences do not predict and therefore cannot correct for the substantial differences that also occur in color appearance. As examples, it is well established that differences in color appearance are not dependent on factors such as the density of preretinal screening pigments or the cone ratios, which strongly impact sensitivity. We developed a procedure that directly adjusts images for the varied color percepts of different observers, based on previous measurements of variations in hue scaling (Emery et al. PNAS 2023) and on a new task where we measured unique and binary hues as well as the achromatic point. Chromaticities in the image are first mapped onto the average scaling function. The corresponding hue percepts are then used to estimate the chromatic axis that would produce the same hue percept in an individual, based on their individual scaling function. Such images should have the property that two observers – each looking at different images tailored to their specific hue percepts – should describe the colors in the images in more similar ways.We use procedure to visualize the range of phenomenal color experience when different observers are looking at the same physical stimulus. The correction we developed is similar in principle to correcting for low-level visual differences in sensitivity (e.g. for observer metamerism) but instead compensates for high-level differences in color perception, and could be used to factor out potential perceptual differences for application and analyses of tasks like color communication or data visualization.

Supported by EY-010834

MONDAY, MAY 20, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Undergraduate Just-In-Time 2

43.352 ALTERED REACH AND GRASP STRATEGIES IN DORSAL CEREBRAL VISUAL IMPAIRMENT

Mia W. Nevin[1], Isabel B. Moody[1], Claudia C. Gonzalez[1], Uma Shahani[2], Isobel Hay[3], Laura M. Ward[4], Jenni M. Karl[1]; [1]Thompson Rivers University, [2]Glasgow Caledonian University, [3]NHS Dumfries and Galloway, [4]University of Dundee

Cerebral Visual Impairment (CVI) is the leading cause of visual impairment in children in developed countries. Ventral stream CVI, which affects perception, is more readily recognized than dorsal stream CVI, which affects visually-guided action, and is frequently mislabelled as Attention Deficit Hyperactivity Disorder (ADHD) or Autism Spectrum Disorder (ASD). A comprehensive description of reach-to-grasp behaviour in dorsal CVI is not currently available but could aid its diagnosis. We hypothesized that children with dorsal CVI would alter their movements to compensate for a dorsal stream impairment when performing visually guided reach-to-grasp actions. Participants reached to grasp blocks of increasing size and complexity from a pedestal while their movements were video recorded from the lateral and frontal perspective. They completed the task with, then without, vision. The results revealed that sighted neurotypical children consistently employed a visually-guided strategy characterized by preshaping their hand to the size of the block, contacting it with their distal fingertips, and efficiently gripping it with minimal adjustments. In contrast, dorsal CVI participants approached the block with inconsistent trajectories and less accurate preshaping. The way they contacted the blocks was more varied than controls. They made contact with a variety of digit and/or hand configurations; this seemed to influence the atypicality of the grasp types they used, as well as the number of post-contact adjustments they were required to make to achieve a more optimal grip. When vision was occluded, the two groups did not differ from one another – they consistently transported an open hand towards the block, contacted it with the proximal fingers, and used post-contact adjustments to eventually achieve an efficient grip. The results suggest that children with dorsal CVI adopt unique multisensory control strategies that incorporate haptic feedback to compensate for impaired visual guidance when performing visually guided, but not unsighted, reach-to-grasp actions.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number RGPIN-2017-05995] and Thompson Rivers University.

43.353 ARE FACIAL MOTION CUES SUFFICIENT FOR RECOGNIZING FACIAL EXPRESSIONS?

Natalia Pallis-Hassani[1], Shruti Japee[1], Maryam Vaziri-Pashkam[2], Christopher Baker[1]; [1]National Institutes of Health, [2]University of Delaware

Our ability to extract meaning from facial expressions is critical for our social interactions. Recent work has proposed a third visual pathway specialized for motion processing, which may be vital for understanding facial expressions. While prior research on facial expression perception has typically used static facial expressions, dynamic facial motion, particularly of eyes and mouth, has recently been shown to improve recognition of facial expressions. Thus, in this study, we examined whether facial motion cues—in the absence of underlying facial features—are sufficient for recognizing facial expressions. To answer this question, we converted dynamic video stimuli of various facial expressions into random dot kinematograms (RDKs) using the underlying optic flow information in the videos. The resulting RDK videos included seven facial expressions—happy, sad, angry, disgusted, fearful, surprised, and neutral—across 22 actors. Participants were shown each of the 154 intact and RDK videos and asked to label the facial expression. Analyses of the labeling data revealed that while average accuracy rates for RDK videos were lower than those for intact videos, participants achieved above-chance accuracy for all expressions—happy: 62%, surprise: 47%, sad: 46%, disgust: 45%, anger: 28%, fear: 26%, and neutral: 98%, with chance accuracy being 14.3%. These results support the idea that motion cues in the absence of facial features are sufficient for identifying expressions. Analysis of eye-gaze data acquired while participants performed the labeling task will help identify facial motion cues that are most salient for the recognition of different expressions. In a future functional magnetic resonance imaging (fMRI) study, these RDK videos will be used to examine whether regions in the third visual pathway process facial motion cues in the absence of the underlying facial features.

43.354 CONSERVATION OF CORTICAL CROWDING DISTANCE ACROSS ECCENTRICITIES IN HUMAN V4

Zelin (Linda) Zhao[1] (zz3130@nyu.edu), Noah C. Benson[2], Jan W. Kurzawski[1], Denis Pelli[1], Jonathan Winawer[1]; [1]New York University, New York, NY, USA, [2]University of Washington, Seattle, WA, USA

Crowding is the failure to identify a visual target due to nearby clutter. It is a major bottleneck for recognition, affecting reading and visual search. Despite its importance both clinically and practically, the neural basis of crowding is not well established. Crowding distance is the minimum separation between a target and surrounding objects that allows recognition. It increases with eccentricity. This has been approximated as a linear function (Bouma law). Our recent study, however, found that the increase is nonlinear, captured by a polynomial with a quadratic term (Kurzawski et al, 2023, Journal of Vision). Here, we take advantage of this finding to ask whether the increase in crowding distance with eccentricity is related to cortical magnification of retinotopic maps. More specifically, we hypothesized that crowding distance, when projected to mm spacing on a cortical map, is a constant. We test this conservation hypothesis separately in 4 maps: V1, V2, V3 and hV4. To do so, we measured crowding distance at 2.5, 5, and 10 deg eccentricity using psychophysics, and measured retinotopic maps with functional MRI. We then calculated cortical crowding distance, measured in mm of cortical spacing, as the product of crowding distance, in deg, and cortical magnification, in mm per deg. We find that in hV4, cortical crowding distance is conserved with eccentricity: 1.5±0.08 mm, 1.6±0.15 mm, and 1.8±0.15 mm for 2.5, 5, and 10 deg targets. In contrast, in V1 to V3, the cortical crowding distance increases systematically with eccentricity. Our finding that cortical crowding distance is conserved across eccentricities in the hV4 map complements our recent finding that it is also conserved across observers in the hV4 map. Together, the results show that crowding is closely linked to the cortical magnification of the hV4 map.

43.355 DEEP LEARNING AND THE ALLOCATION OF COVERT EXOGENOUS SPATIAL ATTENTION: A NEURAL NETWORK PREDICTS THE PRESENCE OF AN ABRUPT ONSET FROM TRIAL-LEVEL PUPIL DATA

Isshori Gurung[1], Matthew Parrella[1], Nicholas Crotty[1], Michael Grubb[1]; [1]Trinity College

An abrupt onset in the visual periphery (a task-irrelevant “disk”) elicits the reflexive allocation of spatial attention without observable eye movements. Transient, peripheral disks also modulate pupil size, which is observable with an eye-tracker. It can be difficult, however, to disentangle the pupillary response to the disk, from the pupillary response to subsequently presented stimuli (ie, those needed to observe the impact of attention). In the service of an ongoing investigation about the interaction between expectation and exogenous attention, we used a 2000ms stimulus-onset-asynchrony (SOA) between a briefly-presented peripheral disk and a simple visual target to isolate the pupillary response to the disk alone. Across ~10,000 trials, a disk was presented half the time. For each trial, we collected 1000 samples of pupil area during the SOA and recorded whether a disk was presented or omitted. Motivated by the popularity of deep learning, here we asked: Can a fully connected neural network (NN) categorize a trial’s disk status (presented or omitted) using the pupil timeseries alone? We built a NN that takes in trial-level pupil data and passes them through three fully connected hidden layers, each with 512 neurons. The NN returns a single value indicating its prediction of whether a disk was presented or omitted on that trial. The network underwent a 3/4:1/4, training:validation split with an early stopping procedure. After 18 epochs, the NN predicted disk status well above chance (72% vs. 50%) for the ~2500 validation trials. In sum, a NN can reliably predict the presentation/omission of a task-irrelevant peripheral disk using trial-level pupil data alone. This project demonstrates the utility of using a NN as a complementary analysis technique for pupil data, and a shareable, interactive python notebook will make our pipeline accessible beyond our lab.

NSF-2141860 CAREER Award to Michael Grubb

43.356 EFFECTS OF MODIFIED VISUAL ENVIRONMENTS ON QUIET STANCE

Aliza Maqsood[1], Lisa Lavalle[2], Taylor Cleworth[1]; [1]School of Kinesiology and Health Science, York University, [2]School of Medicine, Queen's University

INTRODUCTION: Falls caused by balance deficits are a major concern among older adults. Visual feedback provides a crucial source of sensory information to maintain postural control. Optimizing the visual environment could help support postural control among populations with an increased fall risk. Previous work has shown that both contrast and spatial frequency changes can contribute to postural instability in patients with cataracts. Therefore, the aim of this study was to explore the effect of modifying contrast and spatial frequency intensities in a virtual environment on upright stance. METHODS: 28 healthy participants stood quietly on a force plate with feet together while wearing a virtual reality (VR) headset. Four conditions modified the level of contrast (low or high) and/or spatial frequency (high or low) of the wallpaper surrounding a virtual room. Each trial was 60s and each of the four conditions were completed on both a firm and foam surface. Anteroposterior (AP) and mediolateral (ML) center of pressure (COP) was calculated from ground reaction forces and moments, and AP head displacement (HD) was captured from the VR headset. Root mean square (RMS) was calculated to quantify amplitude for all variables. RESULTS: There was a significant main effect of contrast on COP AP RMS and HD AP RMS for firm and foam conditions, and COP ML RMS on foam only, where RMS increased as contrast decreased. There was also a significant main effect of spatial frequency on COP ML RMS for foam conditions only, where RMS increased as spatial frequency decreased. No other interactions or main effects were observed. CONCLUSIONS: Overall, participants postural stability decreased when exposed to lower levels of contrast and spatial frequency. Therefore, visual cues in the environment should be taken into consideration when designing environments for older adults, such as residential and long-term care facilities.

Vista at York University and Natural Sciences and Engineering Research Council of Canada (NSERC)

43.357 EMOTIONAL JUDGMENTS DEPEND ON PERCEIVED GENDER.

Sheida Mirzaei Domabi[1], Ellie Leahey, Jason Haberman; [1]Rhodes college, [2]University of Alabama, Birmingham

Recent studies show that the perceived emotion of a face may be influenced by its perceived gender — female faces are often seen as sadder than they actually are, while male faces are viewed as angrier. However, a major concern in how these studies are designed is that the stimuli (male and female faces) come from different identities, which introduces a confound: Are the differences in emotional perception driven by differences in gender, or by other unidentified differences between the stimuli. We address this concern by using physically identical, androgynous faces along with a rigorous psychophysical approach. We biased perceived gender by introducing non-face cues to androgynous faces (e.g., long hair). We generated morph sequences of androgynous faces (perceptually in-between male and female) from sad to angry, then feminized or masculinized them by modifying their hairstyles. Critically, the facial content was identical in both sets. In a 2x2 design, observers (N=46) viewed either an individual face or a set of four faces composed of either 'female' or 'male' morph(s), followed by a single, neutral test face (no hair). The mean expression of the sets was chosen randomly on every trial, and sets comprised faces varying in emotion. Observers adjusted the emotional intensity of the test face to match the perceived emotion of the previous set. There was a significant bias to view the 'male' faces as angrier and the 'female' faces as sadder, consistent with the stereotyped view of how perceived gender influences perceived emotionality. This, in spite of facial content being identical. There was no effect of set type (individual or ensemble). Although some of our previous work did not show an effect of perceived gender, the method-of-adjustment employed here provides a more sensitive measure to detect differences.

43.358 EXAMINING TASK-DEPENDENT CHANGES OF VISUAL WORKING MEMORY REPRESENTATIONS

Safiya grant[1], Chunyue Teng[1]; [1]Lawrence University

Task-relevant information maintained in visual working memory (WM) may be used flexibly to address different task demands. This study investigates how the functional relevance of information within visual WM – distinguished as task-relevant content (e.g., actively used and reported) and context (e.g., cues guiding memory retrieval) – may differentially affect visual processing and consequently be influenced by distractor processing. In a dual-task paradigm, participants performed a visual search task during the delay period of a visual WM task. They memorized both the location and orientation of a grating stimulus, followed by a delay, after which they were prompted to recall either the orientation or location of the memory sample. The other stimulus dimension served as context, signaling whether the recalled content required mental manipulation. The visual search task included distractors that matched the content of WM sample, its context, both, or neither, allowing for an examination of WM-perception interaction. Results showed that both content and context within visual WM guide visual processing, evidenced by increased response time in the visual search task when the distractor matched with either dimension of the WM representation. Further, interference to memory recall was modulated by the information's task-specific function. Recall was most precise when the search distractor was consistent with the WM item in both dimensions, and suffered the most when the distractor matched the context but mismatched the content dimension. Conversely, distractors with a mismatched context caused more swap errors, indicating a disruption to the content-to-context binding. This pattern was consistent across different feature dimensions. Altogether, these results suggest that both content and context of WM representations interact with visual perception bidirectionally. They further underscore the importance of understand the mechanisms of visual WM from a functional perspective.

43.359 FAMILIAR SIZE CREATES A DEPTH EFFECT THAT GENERATES ILLUSORY MOTION WHEN THE OBSERVER MOVES.

Amoolya Chengalasetty[1] (amoolyac@gmail.com), Jhera Darefsky[1], Albert Yonas[1]; [1]Arizona State University

Familiar size is a pictorial depth cue for which an object's known physical size and visual angle influence observations of apparent distance. There is controversy over the type of process that accounts for the reports of depth. There is also controversy on the question of the effectiveness of this cue (. According to Gogel (1976), a non-perceptual cognitive process occurs when the observer experiences a display larger or smaller than a familiar object, such as off-size. The viewer judges that the object is closer or farther away than its perceived distance. To test this notion, Google had his participants move their heads to see if the apparent depth generated motion parallax. He found that the displays generated slight illusory motion parallax when observers moved laterally. Using realistic larger and smaller versions of Rubik's Cubes and dice, Culham has recently reported that familiar size affects the actual distance of the display even in the presence of conflicting information from stereopsis. We created sets of novel objects that varied in size but were identical in shape and color. Twenty-two college students were familiarized with a smaller or larger version of three objects and asked to judge its apparent distance by moving a rod viewed with both eyes. After apparent depth was reported, the rod was placed at the same distance as the familiarized object, and the observer moved their head back and forth. Perception of motion of the object toward and away from the rod that was concomitant with the motion of the head was found on trials in which a large depth illusion was reported. On trials with little depth effect was observed, subjects reported no apparent motion. Apparent motion supports the view that familiar size is a perceptual illusion, and when it is sufficiently compelling, it can create apparent motion.

43.360 OPTIMIZED LAYER-SPECIFIC FMRI METHODS TO DISSOCIATE FEEDFORWARD AND FEEDBACK INFORMATION ACROSS LAYERS OF THE VENTRAL VISUAL STREAM

Taylor L. Li[1], Andrew S. Persichetti[1], Sam Audrain[1], Jiayu Shao[1], Laurentius Huber[1], Alex Martin[1]; [1]NIMH (National Institute of Mental Health)

Layer-specific fMRI promises to dissociate feedforward and feedback information across cortical laminae from V1 to downstream category-selective visual regions in ventral occipitotemporal cortex (VOTC). However, using a cutting-edge functional MRI method called vascular space occupancy (VASO) to measure fMRI signals at submillimeter resolution comes with major methodological challenges. Thus, we introduce two methodological advances that allow us to measure

layer-specific fMRI signals in VOTC. The first is a forward model that can predict the optimal flip angle regime for the VASO sequence in the brain region to be studied. The second is an anatomical segmentation routine that cleanly segments the cortical ribbon from white matter and cerebral spinal fluid for precise definition of cortical layers. We used this optimized VASO fMRI routine in a study on perceving and imaging faces and places. Participants saw the names of famous faces and places followed by either a picture (perception), or a white frame (mental imagery) during separate task blocks. After independently localizing the fusiform face area and parahippocampal place area, we found preliminary evidence that mental imagery elicits the strongest responses in the superficial and deep layers of the corresponding category-selective region that receive feedback signals from higher-order brain regions but not in the middle layers that receive feedforward signals from early visual cortex. In contrast, viewing pictures of famous faces and places elicits the strongest responses in the middle (and superficial) layers. Thus, our methodological advances allow us to accurately dissociate feedforward and feedback information across layers of VOTC.

43.361 PERCEIVED SELF-MOTION DURING A DYNAMIC BALANCE TASK WITH CONTINUOUS VISUAL ROTATION

Atara Lipson[1], Sara Weinberg[1,2], Taylor Cleworth[1,2]; [1]School of Kinesiology and Health Science, York University, Toronto ON, Canada, [2]Centre for Vision Research, York University, Toronto ON, Canada

INTRODUCTION: Dynamic visual cues, such as circular vection stimuli (CV), influence both balance control and self-motion perception. Furthermore, visual cues impact balance behaviour during dynamic balance tasks. Currently, there is limited work examining self-motion perception when exposed to dynamic visual cues during a dynamic balance task. This study aims to assess actual and perceived self-motion during a dynamic balance task with continuous visual rotation. METHODS: Twenty healthy young adults (mean age 21, 14 female) stood on a platform that continuously tilted forwards (3º) and backwards (3º) for 30 seconds. Participants wore a virtual reality headset displaying white dots on a black background. 3D motion was recorded from markers on 6 bony landmarks. Four balance tasks were randomized, including two static stance (SS) conditions (quiet stance, quiet stance with CV) and two dynamic stance (DS) conditions (platform tilt, platform tilt and CV). Participants completed all balance tasks while tracking their movement using a handheld tracking device, and again while not tracking their movement. Trunk angular (TAD) and linear (TLD) displacement and tracked displacement (TRACK) were assessed using root mean square (RMS). RESULTS: During SS and DS conditions, CV influenced actual movement, where TAD and TLD RMS were larger during CV. CV also influenced tracked movement, where TRACK RMS increased with CV in SS conditions only. Therefore, the ratio between perceived and actual movement increased when exposed to CV during SS conditions, while there was no change during DS conditions. During DS trials, tracking increased TAD and TLD RMS compared to non-tracking trials. No other interaction or main effects were observed. CONCLUSIONS: Exposure to dynamic visual feedback increases postural sway. However, perception of self-motion is disrupted when simultaneously exposed to a dynamic balance task with CV. Further work needs to examine cortical regions involved in perception of self-motion during dynamic balance tasks.

This project was funded by NSERC and VISTA at York University (TC).

43.362 TASK INSTRUCTION AFFECTING VISUAL SALIENCE AND EYE MOVEMENTS

Xiangteng (Kai) Zhao[1] (kaizhao102@gmail.com), Alejandro Lleras[1], Simona Buetti[1]; [1]The University of Illinois Urbana-Champaign

Visual salience is thought to be the key determinant of eye movements under free-viewing conditions. The MIT Salience Benchmark is a resource that compares the success of computational toolboxes at predicting salience (and thus, eye movements) in digital images. However, a close look at the MIT Benchmark paper reveals that investigators used “memorization” instructions rather than a true “free viewing” instruction, when collecting the ground truth eye movement data. It is less clear how salience ought to impact eye movements under memorization conditions. The impact of the memorization instructions on eye movements could be problematic if subjects’ eye movements are different under the two instruction conditions. We hypothesized this would be particularly pronounced when indexing eye movements to low salient objects because whereas in free-viewing conditions, participants have little reason to visit these locations, under memorization conditions, participants might feel a need to visit those locations to better commit that information into working memory. The current study explored how eye movement behaviors towards low salience stimuli differ as a function of these two instruction conditions. Subjects were exposed to 96 varied scenes across two blocks of trials, engaging in either a memorization or a free-viewing task. Eye gaze was tracked using an Eyelink1000 Plus and fixations were categorized by the item they landed closest to. Results across both conditions revealed significant task-driven differences in fixation selectivity. A 13% increase in fixation count and a 42% increase in fixation duration on low salience objects were observed during memory tasks, indicating that memory instructions tend to produce fixation maps that overemphasize low salience regions of the scenes. These results question the validity of the MIT Salience Benchmark rankings. Further examinations and research are needed to validate these findings.

This material is based upon work supported by the National Science Foundation under Grant No BCS1921735 to SB.

43.363 THE EFFECT OF LANDMARKS ON VISUAL STABILITY IN NATURALISTIC SCENES.

Garrett Hensley[1], Jessica L. Parker[1], A. Caglar Tas[1]; [1]University of Tennessee, Knoxville

The present study investigated the effect of landmarks on visual stability in naturalistic scenes. We have previously that spatial position shifts in saccade target (ST) and saccade source (the fixated object prior to saccade) are better detected during the saccade compared to shifts in either background or whole scene, suggesting a special role of ST and source objects in establishing visual stability (Parker & Tas, VSS22). The present study expanded these findings by including a close (2dva) landmark object either above or below the ST. If presaccadic shift of attention spreads to the landmarks, then we expect to find similarly accurate displacement detection for ST and landmarks. If, however, STs are preferentially processed then we expect better displacement detection for STs than landmarks. On each trial, ST was cued with a red highlight, and participants were instructed to execute a saccade to it. During the saccade, one of the five possible horizontal shifts could occur: ST, landmark, whole image, background, no-shift (control). The shifts could be in the same (forward) or opposite direction (backward) as the saccade. Further, we manipulated visual stability with target blanking paradigm. Blanking significantly improved displacement detection, replicating previous work. We also found a significant effect of condition where shifts in ST (.91) were better detected than landmark (.58) and background shifts (.47). Importantly, shifts in landmark objects were better detected than background shifts, but not as good as whole image shifts (.88). We also found a significant interaction: Blanking significantly improved displacement detection for landmarks and whole image shifts, but not background or target shifts, possibly due to floor and ceiling effects respectively. These results show that the visual system uses various sources of information when establishing visual stability in naturalistic scenes, including the landmark objects with the priority given to the ST.

43.364 THE HEMIFIELD ASYMMETRY FOR CROWDING IS STRONGER FOR LETTERS THAN VISUALLY-MATCHED SHAPES

Anishka Yerabothu[1], Nicole Oppenheimer[2], Alex White[1]; [1]Barnard College, Columbia University, [2]Rutgers University

Crowding refers to a phenomenon where neighboring objects interfere with the perception of a central target. An interesting property of crowding is that it is asymmetric across the left and right hemifields for certain stimuli. Previous research, including our own, has found that letters crowd each other less in the right visual field than the left. We investigated how the hemifield asymmetry for letters compares to that of pseudo-letters with matched visual features. In one experiment, participants identified a central stimulus within a set of three stimuli of the same type (letters or pseudo-letters), positioned at 4° eccentricity. Crowding was assessed by fitting full psychometric functions to extract critical spacing thresholds. In terms of those thresholds, there was a more pronounced hemifield asymmetry (left > right) for letter recognition compared to pseudo-letter recognition. In a separate experiment, central target letters were flanked by either other letters or by pseudo-letters. Crowding was less pronounced when the flankers were letters than pseudo-letters, but the hemifield asymmetry was strong in both cases. These results suggest that there is a unique advantage for processing strings of letters in the right visual field. That advantage could reduce interference between the letters within words, when they are presented in locations processed by the language-dominant left cerebral hemisphere.

43.365 TRANSIENT TWINKLE PERCEPTION IN EQUILUMINANT RED-GREEN STIMULUS

EUN CHO[1] (choeun@unist.ac.kr), Chang-Yeong Han[1], Hyosun Kim[2], Oh-Sang Kwon[1]; [1]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, [2]R&D center, Samsung Display

Transient Twinkle Perception (TTP) is the detection of the frequency transition between two sequentially presented flickering stimuli, which occurs even when the two flickering stimuli appear steady by themselves (Nakajima and Sakaguchi, 2015). A recent study showed that the TTP in luminance flickering stimuli can be explained by applying a biphasic filter model (Han et al., 2023, VSS). Here we investigate the properties of TTP in equiluminant stimuli flickering between red-green colors. In Experiment 1(N=11), we measured the magnitude of TTP in equiluminant red-green flicker stimuli, which changed the frequency from 40Hz to 120Hz. Both 40Hz and 120Hz flickering appear steady by themselves; however, participants experienced TTP. To further characterize the TTP, we inserted 'in between' frequency epochs (48, 60 and 80Hz) so that the frequency gradually changes from 40Hz to 120Hz. The numbers of the 'in between' frequency frames varied across eight conditions (one to eight). The magnitude of TTP did not increase or decrease monotonically but fluctuated as the frame number of 'in between' epochs increased. We applied a biphasic filter model and Gaussian filter model to fit the data. The results showed that the Gaussian filter model fits the data from all participants except one (AIC difference: $t(10) = -3.45$, $p=0.0062$). In Experiment 2 (N=4), we directly measured the temporal interaction between two flashing stimuli using Rashbass' experimental paradigm (1970). We presented red and/or green stimuli successively with varying SOA (11, 18, 27, 36, 43, 52, 68, 77, 94ms) and measured detection threshold. Results showed that the suppressive interaction between flashes are negligible, replicating an existing study (Watson & Nachmias, 1977). The results demonstrate that the temporal processing of equiluminant chromatic stimuli can be explained by applying a monophasic filter model, unlike the luminance stimuli, and the TTP is no exception.

This work was supported by the Samsung Display Research Center

43.366 VISUAL WORKING MEMORY IMPAIRS VISUAL DETECTION: A FUNCTION OF SHARED ATTENTIONAL OR SENSORY RESOURCES

A Hamkari[1], D Melcher[1,2], X Liu[1]; [1]Division of Science, Department of Psychology, New York University Abu Dhabi, [2]Center for Brain & Health, NYUAD Research Institute, New York University Abu Dhabi

The sensory recruitment hypothesis posits that the same cortical areas used for visual perception are also used for maintaining visual working memory (VWM). Support for this idea comes from evidence that information stored in VWM can influence perception. To test whether there is competition for sensory processing resources between perception and VWM, we measured the sensitivity (d') of detecting a near-threshold flash during the delay period of an oculomotor delayed response task. We compared detection performance for participants (N=30) under three conditions: (1) an oculomotor delayed matched-to-sample task, in which participants prepared to make a saccade to a remembered target location, (2) an oculomotor delayed non-matched-to-sample task, in which they made a saccade to a new target that was not part of the memory set, and (3) a detection-only condition in which participants ignored the VWM display and only focused on detecting the flash. The visual input was matched across conditions up to the point of the oculomotor response. Based on previous studies, we expected that the matched-to-sample task would only require maintenance of a motor plan, while the non-matched-to-sample version would require maintenance of a sensory code. In a separate session, participants also repeated a simplified version of the study in the MRI. We found that detection varied significantly between the three conditions. Participants had a significantly lower sensitivity (d') for the nonmatched condition, when a sensory code was maintained, compared to the matched condition where a motor code is maintained. Highest sensitivity was found when the VWM stimuli were ignored (condition 3). This pattern of results suggest that maintaining a sensory code in VWM influences basic visual processing, potentially due to the recruitment of sensory processing areas in the active maintenance of visual memory representations.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYU Abu Dhabi Research Institute grant CG012. This research was partially conducted using the Brain Imaging Core technology platforms resources at NYUAD.

MONDAY MORNING POSTERS IN PAVILION

Monday, May 20, 8:30 am – 12:30 pm, Pavilion

Data Visualization

43.401 PERCEPTUAL BENEFITS OF ANIMATION ARE TASK-DEPENDENT: EFFECTS OF STAGING AND TRACING IN DYNAMIC DISPLAYS

Songwen Hu[1] (shu343@gatech.edu), Ouxun Jiang[3], Jeffrey Riedmiller[2], Cindy Xiong Bearfield[1]; [1]Georgia Institute of Techonology, [2]Dolby Laboratories, [3]Northwestern University

Our visual system has evolved to process a dynamic world. We can rapidly summarize the average speeds and trajectories of moving objects at high accuracies (Haberman & Whitney, 2012; Watamaniuk & Duchon, 1992; Williams & Sekuler, 1984). Dynamic data visualizations can convey large amounts of information across time, such as using movements to depict changes in data values for multiple entities. Such dynamic displays demand our visual processing capacities, yet our visual system is limited when perceiving motion. When tracking multiple objects across space and time, we can typically track up to four objects, and the capacity is even lower if we also need to remember the history of the objects' features (Horowitz et al., 2007). Several techniques have been shown to improve the processing of dynamic displays. Staging the animation to sequentially show steps in a transition and tracing object movement by displaying trajectory histories can increase processing by reducing the cognitive load (Fisher, 2010; Robertson et al., 2008). We examine the effectiveness of staging and tracing in dynamic displays. We showed participants animated line charts depicting the movements of four lines and asked them to identify the line with the highest mean and variance. We manipulated the animation to display the lines sequentially or synchronously, either with or without tracing, and compared the results to a static chart as control. Preliminary results showed that, for the mean comparison task, displaying lines sequentially without tracing is associated with increased accuracy, despite participants expressing low preferences for this animation style. For the variance comparison task, tracing is associated with higher accuracy, with the static chart outperforming most dynamic displays. Overall, the benefit of dynamic displays can be task-dependent and misaligned with human preferences. These findings help inform real-world best practices for constructing dynamic displays to harness our visual processing.

This work is sponsored in part by NSF awards IIS-2237585, IIS-1901485 and Dolby.

43.402 BEYOND THE CLOUD: A PERCEPTUAL ILLUSION IN OVERLAID BAR CHARTS

Wenxuan Zhao[1] (xiaodoubizwx@gmail.com), Karen B. Schloss[1]; [1]University of Wisconsin-Madison

A general challenge in information visualization is to represent multiple datasets in a way that supports comparison. One approach is to use overlaid bar charts, in which multiple sets of bars, each representing different measures of the same categories, are layered using colors with varying opacities. One set of bars is sorted from low to high, leaving the other set(s) unsorted. Although such charts may support comparison, we observed an illusion when comparing means—the mean of unsorted bars appears higher than the mean of sorted bars when the means are equal. To study this illusion, we presented participants with overlaid bar charts representing the popularity of two flower types (two sets of bars) across 50 counties (categories), sorted by "flower 1" or "flower 2" (balancing which flower was represented by opaque/translucent bars). Participants reported which flower was more popular overall (i.e., which set of bars had a greater mean). Numerically, the mean popularity of the two flowers was the same. We varied bar color across trials; participants saw all pairs of eight bar hues controlled for lightness and chroma. Overall, participants reported the mean of the unsorted bars was higher than the mean of the sorted bars ($p<.001$), despite the means being equal. A potential explanation is that taller bars in the unsorted set may appear especially salient as they extend beyond the overlapping region, or "cloud," leading participants to weight taller bars (beyond the cloud) more than shorter bars when estimating the mean. Supporting this account, making the smaller bars more salient within the "cloud" by increasing color contrast reduced the illusion ($p<.001$). Thus, we call the illusion the "beyond the cloud" illusion. These results emphasize the importance of understanding how low-level perceptual features influence the ability to perform accurate statistical estimates from information visualizations.

43.403 CORRELATIONS ARE ESTIMATED WITH BIAS IN A 2-CLASS SCATTERPLOT

Yuka Omae[1] (oomae@cv.jinkan.kyoto-u.ac.jp), Jun Saiki[1]; [1]Kyoto University

The current study examined in a multi-class scatterplot, containing multiple bivariate datasets represented by colors, whether humans can visually separate datasets based on color differences and accurately estimate the correlation of each dataset. Previous studies investigating the discrimination threshold for correlations have reported an increase in JND with scatterplots having two overlapping colors, (Elliott & Rensink, 2015 VSS; Elliott, 2021). These studies predict that magnitude of correlations might also be estimated differently in 1-class and 2-class scatterplots. The current study investigated observers' efficiency in filtering out the irrelevant subset in estimation of target correlation in 2-class scatterplots by psychophysical experiments. In each trial, two scatterplots, one 2-class and one 1-class, were presented. Observers compared a correlation of 1-class scatterplot (Comparison) with a correlation of one sub-dataset (Target) in a 2-class scatterplot with the same color as the 1-class, and judged the stronger correlation. The correlation coefficient of Target was constant ($r = 0.6$), and that of the other sub-dataset (Distractor) in a 2-class was set to 4 levels ($r = 0.0, 0.3, 0.6, 0.9$). Using psychometric functions, we estimated the point of subjective equality

(PSE) for the Target's correlation strength. The result showed that PSE for the correlation strength of Target was biased toward that of Distractor, suggesting that it is difficult to filter out irrelevant data points in a 2-class scatterplot. Furthermore, we investigated the robustness of the bias. Manipulation of the stimulus duration (short or unlimited), and of the color and luminance difference between two datasets did not modulate the magnitude of biases at all, suggesting that the estimation of the correlation is biased robustly in a 2-class scatterplot.

This work was supported by JSPS KAKENHI Grant Number 23H04349.

43.404 DRAWN CORRELATIONS CONSISTENT WITH UNDERESTIMATION OF PERCEIVED CORRELATIONS FROM SCATTERPLOTS

Lucy Cui[1] (lucy.cui.phd@gmail.com), Medha Kini[2], Zili Liu[2]; [1]Rensselaer Polytechnic Institute, [2]University of California, Los Angeles

Previous research assessing people's sense of correlation used scatterplots as stimuli and asked participants to estimate the correlation of said scatterplots. This method has consistently shown that people tend to underestimate the correlation of a scatterplot (e.g., guessing the correlation is 0.25 when the actual correlation is 0.5). However, it is unclear whether this underestimation is perceptual or reflective of having a poor internal representation of different correlations. We investigated this question by flipping the task: instead of estimating the correlation from a scatterplot, participants drew a scatterplot based on a given correlation. They drew 20 points onto a tablet to represent the correlation coefficients: 0, 0.25, 0.5, 0.75, and 1 in counterbalanced order. We then calculated the drawn correlations based on x, y coordinates. While the drawn correlations of 0, 0.75 and 1 were quite accurate, the drawn correlations of 0.25 and 0.5 were significantly higher than the requested correlations. The drawn correlations of 0.25 and 0.5 were consistent with the estimated correlations of 0.25 and 0.5 in the literature where participants viewed scatterplots and estimated correlations. Taken together, this suggests that people's underestimation of correlations of < 0.5 stems from not having a good sense of what correlations 0 – 0.5 look like.

43.405 GLOBAL MEAN POSITION PERCEPTION OF MULTIPLE SPATIALLY-SEPARATED CLUSTERS

Yang Wang[1] (wangy0802@hotmail.com), Timothy Brady[1]; [1]UCSD

The visual system is remarkably efficient at extracting summary statistics from the environment, which drive the overall perception of a scene and inform judgments about individual objects within it. Most ensemble research focuses on the perception of one specific characteristic (e.g., average size) from a single group of stimuli, but natural visual environments usually consist of many groups of objects including outliers that are distributed over space. We evaluated how people perceive the aggregated ensemble mean position (i.e., center of mass) perception in the presence of multiple spatial clusters using a visualization task with scatterplots. Consistent with previous results, when there are distinct two clusters of dots, we find that people reliably overweight the group of smaller cardinality, including if it is a single outlier. For example, for two clusters, made up of 8 vs 16 dots, people report the average position too close to the cluster of 8. However, when the larger cluster is equally partitioned into multiple clusters (e.g., the 16-item cluster is split into two 8-item clusters), the overweighting of the small cluster dampens significantly, despite the total cardinality and center of mass of the large cluster being held constant. Furthermore, the bias towards the partitioned clusters increases at larger cardinality and increasing number of partitions. These results suggest that people are affected by the presence of distinct clusters, and partially rely on the average position of these clusters rather than solely the average position of the individual objects. Thus, Gestalt organization of clusters significantly alters the aggregated mean perception. Overall, we show that people make large systematic errors when judging the aggregated mean of multiple clusters, which is a realistic task that occurs in domains like calculating the average value in a scatterplot or identifying the balancing point of a set of objects.

43.406 MAPPING INFECTIOUS DISEASE AS A TEST OF TOP-DOWN EFFECTS ON ENSEMBLE PROCESSING

Jessica Witt[1] (jessica.witt@colostate.edu), Amelisa Warden[1], Danielle Szafir[2]; [1]Colorado State University, [2]University of Norh Carolina - Chapel Hill

Choropleth maps are a critical tool for communicating geospatial data, such as the spread of an infectious disease. How effectively a map communicates information depends on its design. Insights concerning map design principles can be gleaned from research on ensemble perception—the visual system's ability to extract summary statistics from a scene, such as the mean number of disease cases. Ensemble perception suggests that showing more data points on a map and encoding data with colors that minimize variability may improve people's ability to estimate statistical information from maps. We tested these recommendations and the effects of the size of geographic regions and data semantics to understand ensemble perception processes when viewing maps. People made yes/no judgements about whether the mean number of disease cases on a map exceeded a critical threshold. We found that people were better at estimating the mean of the map when presented with more data points (i.e., individual counties instead of regions) and lower variability across data points. That aggregating data into larger regions did not improve mean estimation suggests that while ensemble perception processes may bear some resemblance to statistical processes, they are incapable of performing the visual equivalent of a meta-analysis. We also found that the perception of ensembles in maps was influenced by the content of the map: people were more sensitive to the map's mean when the map depicted the number of people with an infectious disease rather than when the map depicted the number of people with immunity to the disease. This reveals a top-down effect on ensemble processing. Our results have implications for theories of ensemble perception as well as for design recommendations for choropleth maps. Understanding how people interpret and derive meaning from geospatial data informs design guidelines for effectively communicating critical information to the public.

NSF BCS-1632222; NSF SES-2030059

43.407 THE IMPACT OF COGNITIVE DIFFERENCES ON PROCESSING DATA STORIES THROUGH INFOGRAPHICS: ADVANCING TOWARD INCLUSIVE DESIGN

Kristine A. Zlatkovic[1] (kristine.zlatkovic@northwestern.edu), Pavlo Antonenko[2], Do Hyong (Ryan) Koh[2], Andrea Ramírez-Salgado[2]; [1]Northwestern University, [2]University of Florida

This study explores how individuals with diverse cognitive abilities process data stories through infographics that integrate visual or verbal data representations, including seductive details. Specifically, it investigates the interactive effects of individual cognitive differences (working memory, inhibitory control, and visual search abilities) and data representation mode (visual vs. verbal) on the visual processing and comprehension of infographics. Fifty-one undergraduates performed tasks involving data infographics, with six communicating data verbally and six visually, alongside completing Inquisit tests assessing cognitive differences. Infographics were segmented into four regions of interest (ROIs), containing task answers. Comprehension was assessed by task accuracy. Visual processing was assessed by the number of fixations and their average durations on ROIs, utilizing the dispersion-threshold identification method and moving window algorithm. The author's custom Python program, available on GitHub (Kantonyan, n.d.), facilitates this analysis. Multiple mixed-effects regression analyses predicted visual processing and comprehension, incorporating data representation mode and individual cognitive differences. The study reveals that individuals demonstrate more effective engagement and comprehension with verbally represented data stories, influenced by the emphasis on improved reading comprehension in K-12 mathematics instruction. However, individuals with superior visuospatial working memory and goal-oriented visual search abilities exhibit enhanced engagement and comprehension across both data representation modalities, emphasizing the importance of spatial information processing. Furthermore, individuals effectively suppress unnecessary seductive details primarily when exploring verbally represented data stories, driven by their perception of data visualizations as images. This finding reflects their behavior to seek supportive information in other displayed images. The findings carry twofold implications: firstly, they provide valuable insights for designers to enhance data infographic design, and secondly, they underscore the effectiveness of utilizing eye-movement data to explain infographic processing. Future research is encouraged to explore real-time eye-movement data for classifying individuals based on cognitive differences and developing adaptive data stories for inclusive engagement.

43.408 VIEWERS RARELY NOTICE OBJECTS THAT ENTER OR LEAVE DYNAMIC INFORMATION DISPLAYS WITHOUT SALIENCE-BOOSTING TRICKS

Ouxun Jiang[1] (ouxunjiang2026@u.northwestern.edu), Fumeng Yang[1], Matthew Brehmer[2], Steven Franconeri[1]; [1]Northwestern University, [2]Tableau Research

One primary motivation for studying object tracking is to inspire better designs of visual information displays that contain multiple moving points, such as air-traffic control displays, educational simulations, or dynamic data visualizations. Most studies focus on improving one metric within these moving displays: how many 'targets' can we track under different display conditions. But realistic displays also demand another task: noticing whether a moving object (e.g., an airplane, a particle, or a data point) enters or leaves the edges of that display, even when it is not actively tracked. How well would viewers spontaneously notice that event, and how can we help them improve? We set up dynamic displays where tracking two of 20 objects was the primary task, and noticing whether one object entered or left was a secondary task. When tracked objects left the display, people easily noticed them (94%), though they were far less likely to notice when untracked objects left (46%). We expected that noticing untracked objects entering or leaving would be difficult, so we tested a salience-boosting trick inspired by real-world information displays: including buffer areas around the display where entering or leaving objects would be temporarily spatially salient. For untracked objects, people were far better at noticing objects enter or leave when they passed through the outer buffer (75%) compared to when they did not (37%). Intriguingly, both with and without the buffer, it was easier to notice untracked objects entering the display (66%) than leaving (46%). In summary, the buffer around the edges of the display helped people notice entering or leaving untracked objects, validating a technique used by professional information designers. If our visual system provides an illusion of perceived detail by relying on heuristics and statistics, then real-world information displays must leverage the right tricks to guide viewers to important changes.

43.409 LEARNING TO DETECT PATTERNS IN 2X2 GRAPHS

Nestor Matthews[1] (matthewsn@denison.edu), Megan Broderick[1], Samantha Kozlowski[1]; [1]Denison University

Introduction: Factorial research designs play an important role in many academic fields. Accordingly, diverse academic disciplines commonly require students to develop skills in comprehending factorial graphs, often starting with simple 2x2 graphs. Prior studies on 2x2 graph comprehension had participants either "think aloud" (Ali & Peebles, 2013; Peebles & Ali, 2015) or provide written descriptions (Shah & Freedman, 2011) about 2x2 main effects and interaction effects. Those methods privilege declarative knowledge (what can be stated), and potentially miss perceptual patterns that participants might learn non-declaratively. To fill this gap, we measured perceptual learning in 2x2 graph-pattern-detection using a trial-and-error task that does not require declarative knowledge. Method: Participants viewed black and white 2x2 bar graphs or line graphs and classified each graph into either of two initially unknown categories. The categories corresponded to significant versus non-significant effects in one of three randomly assigned 2x2-target-factors. These included Factor A Main Effects (left vs right height differences), Factor B Main Effects (black vs white height differences) or Interactions (slope differences). Across trials, 2x2 graphs with significant effects had effect sizes that varied randomly among Cohen's d values of 0.2 ("small"), 0.5 ("medium"), 0.8 ("large"). Results: Significantly more perceptual learning occurred for detecting significant effects in line graphs than for detecting significant effects in bar graphs ($p<0.001$). For line graphs and bar graphs alike, significantly more perceptual learning occurred for detecting significant interactions than for detecting significant Factor A Main Effects ($p<0.001$). Perceptual learning for detecting significant Factor B Main Effects fell between those two extremes. Surprisingly, perceptual learning did not depend on the effect sizes in the 2x2 graph stimuli. Discussion: Significant biases exist in how viewers perceptually organize 2x2 graphs. These biases have implications for effective data visualization (e.g., visualizing public health information), and statistics education (e.g., exploiting perceptually salient patterns).

Monday, May 20, 8:30 am – 12:30 pm, Pavilion

Visual Search: Neural mechanisms, clinical, applied

43.410 SACCADIC LATENCY DIFFERENCES BETWEEN C-SECTION AND VAGINALLY-DELIVERED INFANTS AND ADULTS IN A VISUAL SEARCH TASK

Scott Adler[1] (adler@yorku.ca), Maryam Rahimi[1]; [1]York University

Search asymmetry occurs when an efficient search is exhibited for feature-present targets, but an inefficient search is exhibited for feature-absent targets. Both 3-month-old infants and adults have been found to display a search asymmetry when assessed with saccade latencies (Adler & Gallego, 2014) Additionally, C-section delivered infants exhibit slower attention and saccadic latencies than those born vaginally in an attentional cueing paradigm, implicating a detriment to bottom-up attention (Adler & Wong-Kee-You, 2015). This study was designed to confirm the relative effects of different birth experiences on attention with a different attentional task, namely, search asymmetry, and whether any differences persist to adulthood. Two visual circular arrays were presented: feature-present target among feature-absent distractors (R among Ps) or feature-absent target among feature-present distractors (P among Rs) with array set sizes of 1, 3, 5, 8. Results indicated that infants' and adults' saccadic latencies to feature-present targets were unaffected by set size and C-section infants and adults both had slower saccadic latencies when compared to vaginal groups. Interestingly, infants born via emergency C-section were faster when compared to a planned C-section but not different from vaginal infants. There were no differences, however, in saccadic latencies between emergency and planned C-section adults. For feature-absent targets, both infants and adults exhibited increasing saccadic latencies with set size. Further, both infants and adults across all types of birth experiences exhibited equivalent saccadic latencies. These findings further suggest that C-section birth influences bottom-up attention even into adulthood. The discrepancy between the relative difference between planned and emergency C-sections in infants compared to adults suggests that multiple mechanisms, birth process versus bacterial baptism, are likely responsible for differential pattern across ages. Regardless, this study demonstrates that the development of attentional mechanisms can be influenced by early birth experiences and that this birth impact persists into adulthood.

The Hallward Fund of the Toronto Foundation

43.411 A COMPARATIVE ANALYSIS OF THE USE OF COMPUTER-AIDED DETECTION SYSTEMS WITH MOCK DIGITAL BREAST TOMOSYNTHESIS SETS

Katharine E Sabo[1] (sabo.katie@gmail.com), Dawn M Sarno[1]; [1]Clemson University

Computer-aided detection (CAD) systems are essential tools for aiding radiologists in assessing breast images for cancer. Despite their widespread use, limited research has compared the various types of CAD systems, particularly for digital breast tomosynthesis (DBT) imaging, which is anticipated to surpass mammography in use. Additionally, calls for including global classifications in CAD systems, particularly the Breast Imaging Reporting and Database System (BI-RADS), suggest such ratings may enhance diagnostic accuracy (Qian et al., 2015). This study examined three existing CAD system types (binary, analog, and interactive); a novel CAD system with a global classification (BI-RADS); and a control with no CAD system. Undergraduate students (N = 50) were randomly assigned to these conditions and tasked with identifying a target (T) among distractors (Ls) in mock DBT displays. For each trial, stimuli were randomly distributed across 13 layers to mimic the pseudo-3D images of DBT sets. Assistance provided to participants in the automation conditions varied by CAD type: binary CAD circled likely targets; BI-RADS CAD supplied a rating and descriptor indicating the presence of a target; analog CAD circled potential targets and provided probability estimates; and interactive CAD (iCAD) required participant activation on each trial to reveal an analog CAD overlay. Results revealed that participants in the binary and analog conditions exhibited an average 39% increase in hit rate and a 16.59s decrease in response times compared to the control, BI-RADS, and iCAD conditions. Notably, participants in the BI-RADS and iCAD conditions did not significantly differ from the control condition or each other. While iCAD and BI-RADS are considered advancements in CAD technology, this research suggests that further investigation into their influence on search performance is necessary before replacing the widely used binary and analog CAD systems.

43.412 YOUR DERMATOLOGY OFFICE VISIT: THE EFFECTS OF TARGET VISIBILITY, TARGET-DISTRACTOR SIMILARITY, SATISFACTION OF SEARCH, AND TIME PRESSURE ON A VISUAL SEARCH FOR CANCEROUS MOLES

Katherine S Moore[1] (moorek@arcadia.edu), Devon Fithian[1], Jenna Pedrick[1]; [1]Arcadia University

Basic and applied studies of visual search have demonstrated a number of factors that influence search accuracy and speed. Among these factors include target-distractor similarity, visibility of a target against the background, and subsequent miss errors. While some applied settings for visual search have been extensively studied, dermatology has received little attention from visual search research. Here, participants (Arcadia students) completed a visual search task for cancerous moles (melanoma and carcinoma) among benign distractor moles, simulating a dermatological search. First, participants were trained to identify melanoma, carcinoma, and benign moles. Once they reached criterion, they performed a speeded skin search for these moles on a skin background. We varied (1) target visibility with skin tone, (2) target-distractor similarity (melanoma is more visually similar to benign moles than is carcinoma), and (3) the number of targets on a trial, to measure subsequent misses. Since dermatologists are often under great time constraints when seeing patients, we investigated the effect of time pressure by displaying a dynamic clock on half of the trials. Consistent with visual search literature, skin tone (target visibility) and mole type (target-distractor similarity) affected performance, but there was not a significant subsequent miss error effect. As predicted, participants were both faster and less accurate in the presence of a clock than when no clock was present. Our findings may inform how dermatologists can be trained to perform the most accurate and efficient skin cancer screenings. Reliable and valid screenings can not only save lives by catching cancers early, but also save on expensive and invasive procedures by avoiding false positives.

Parts of this study were supported by the Barbara Nodine fund for undergraduate research at Arcadia University

43.413 ADVANTAGES AND DISADVANTAGES OF SEQUENTIAL VS. SIMULTANEOUS SEARCH IN SIMULATED BREAST CANCER SCREENING

Injae Hong[1] (ihong1@bwh.harvard.edu), Chirag Maheshwari[2], Jeremy Wolfe[1,3]; [1]Brigham & Women's Hospital, [2]Cypress Ridge High School, [3]Harvard Medical School

Radiologists, screening mammograms for breast cancer, are required to search for different signs of cancer such as masses, calcifications, and structural distortions. This search should be fast, accurate, and complete. Search for more than one type of target is known as “hybrid search”. Hybrid search can impose a cost on performance that increases with the number of different types of possible targets. Accuracy might be improved by splitting a hybrid search into multiple simple searches for a single target type. This study investigated whether splitting search might be a useful intervention to improve target detection in breast cancer screening. Non-experts searched for either masses or calcifications or both in simulated 2D mammograms (Experiment 1) or while scrolling through 3D volumes of simulated digital breast tomosynthesis images (Experiment 2). Masses and calcifications were independently present with 60 % prevalence. If the target was present, participants clicked on the target and labeled the item as a mass or a calcification using a key press. They received feedback after pressing the space bar to complete the trial. There were four types of task: search for calcifications alone, search for masses alone, search for both types of targets simultaneously, or both types but one after the other, sequentially. The results showed that sequential search was advantageous compared to simultaneous search. In particular, there was a reduced level of “satisfaction of search” errors in rare cases when more than one target was present. SoS errors declined from 60% to 22% (Experiment 1) and 67% to 32% (Experiment 2). However, reduction of SoS errors on a few trials comes at substantial 20% and 28% increase in time for all trials. The standard simultaneous method of searching might be adequate, but splitting hybrid search into several simple searches may reduce errors in some, important cases.

CA207490

43.414 ASSESSING FUNCTIONAL VISION IN CEREBRAL VISUAL IMPAIRMENT WITH WEARABLE EYE TRACKING

Madeleine Heynen[1] (mheynen@meei.harvard.edu), Claire Manley[1], Serena Micheletti[2], Elisa Fazzi[2], Lotfi Merabet[1]; [1]Massachusetts Eye and Ear, Harvard Medical School, [2]ASST Spedali Civili of Brescia

Functional visual deficits are often present in individuals with cerebral visual impairment (CVI). We characterized gaze behavior while CVI participants performed the Star Cancelation Task (SCT) and Developmental Test of Visual Perception (DTVP) as assessments of visuospatial processing and visuomotor integration, respectively. Gaze behavior was recorded using the Tobii Pro Glasses 3 in five CVI subjects (mean age=21.20 years ± 4.66 SD) and eight controls (mean age=22.13 years ± 4.16 SD). For the SCT, both groups performed at ceiling (i.e. no targets missed). However, CVI subjects had higher fixation (81.05%) and saccade counts (67.31%), shorter average fixation durations (42.62%), and shorter average saccade amplitudes (18.68%) compared to controls. These findings suggest that while there was no evidence of visuospatial neglect, search behavior in CVI was comparatively different from that of controls. Regarding the DTVP, the CVI group showed evidence of impaired copying abilities and had overall higher fixation counts (23.18%) and shorter average fixation durations (40%) compared to controls. Both groups had higher fixation counts (CVI: 9.68%; Control: 31%) and longer average fixation durations (CVI: 20.51%; Control: 97.78%) while viewing the drawings compared to targets as well as a trend of increasing fixation counts with increasing target complexity. In general, we found that gaze behavior in CVI was characterized by a greater number of fixations and decreased fixation duration, as well as an increased number of saccades and decreased saccade amplitude. These findings provide insight into the gaze behavior of CVI subjects during visuospatial processing and visuomotor integration tasks that characterize functional vision impairments in this population.

43.415 ASSESSING VISUOMOTOR ABILITIES IN CEREBRAL VISUAL IMPAIRMENT WITH EYE AND HAND TRACKING

Lotfi Merabet[1] (lotfi_merabet@meei.harvard.edu), Madeleine Heynen[1], Claire Manley[1], John Ross Rizzo[2], Helen Lindner[3]; [1]The Laboratory for Visual Neuroplasticity, Department of Ophthalmology, Massachusetts Eye and Ear, Harvard Medical School, Boston, USA., [2]Department of Rehabilitation Medicine, New York University Grossman School of Medicine, New York, USA., [3]School of Health Sciences, Faculty of Medicine and Health, Örebro University, Sweden

Cerebral visual impairment (CVI) is a brain-based visual disorder, and visuomotor impairments are often observed. Using wearable eye-tracking technology and hand-focused motion-capture analysis, we characterized gaze and hand movement behaviors during a wooden block-matching puzzle test. 5 subjects with CVI (mean age: 21.20 years ± 4.66 SD, mean MACS: 1.20 ± 0.45) and 9 controls (mean age 21.00 years ± 3.61 SD, mean MACS: 1.00 ± 0.00 SD) participated in the study. For the task, 9 blocks were placed in a starting position, and subjects were required to grasp and place them in their corresponding target position based on their shape and orientation. To examine the effect of cognitive demand, two versions of the task were performed (presented in counterbalanced order), namely, a free task during which subjects selected the blocks at random and a sequenced task during which the blocks were selected in a predetermined order. Gaze behavior was initially recorded using Tobii Pro3 eye-tracking glasses, and videos were then analyzed using a customized MATLAB script that automatically tracked eye gaze, the hand, and the color of each block, enabling the detection of their movements from the starting to the ending position. Results revealed that CVI subjects showed longer latencies between eye and hand movements for both tasks (free: 5.87 sec ± 1.79 SD and sequence: 7.44 sec ± 2.20 SD) compared to controls (free: 4.88 sec ± 1.01 SD and sequence: 5.39 sec ± 1.05 SD).

A similar trend was observed with an increased number of saccades for both tasks in the CVI group (free: 55.40 ± 9.53 SD and sequence: 50.80 ± 13.81 SD) compared to controls (free: 40.33 ± 8.80 and sequence: 41.58 ± 10.11 SD). These preliminary results provide objective evidence consistent with eye-hand dyscoordination and, more generally, visuomotor impairments associated with CVI.

NIH/NEI (R01 EY03097) to LBM

43.416 CUE-GUIDED SEARCH FACILITATES ATTENTIONAL SELECTION: EVIDENCE FROM AN EEG STUDY

Sizhu Han[1] (sizhu.han@uni-marburg.de), Anna Schubö[1]; [1]Philipps-University Marburg

Understanding how features guide attention amidst distractions is crucial in deciphering attentional templates. Early research has found that presenting target-relevant information (i.e., positive cue) ahead of the search task facilitates attentional guidance; more recent studies suggest that information about distractions (i.e., negative cue) can also be used to direct attention away from distractors. Yet, the underlying neural mechanisms are not well understood. The current study addressed this issue by adopting a modified cue-guided search paradigm, combined with EEG recordings, to examine how electrophysiological markers of attentional selection (N2pc and PD) were modulated by positive, negative and neutral cues. Behavioral results showed that positive cues led to faster response times (RTs) compared to negative and neutral cues. In line with this finding, ERP results time-locked to the search display showed an earlier onset of the target N2pc for positive cues compared to negative and neutral cues. Interestingly, results for negative cues were more heterogeneous across participants: 16 demonstrated improved RTs, while 14 showed increased RTs compared to neutral cues. This variability was also reflected in ERP results. With a negative cue, lateralized distractors in the search display were followed by a PD component, predominantly in participants who benefited from negative cues in RTs. Moreover, positive cues elicited larger lateralized alpha power than negative cues before the search display onset, suggesting a difference in anticipatory alpha activity. These findings suggested that both positive and negative cues can be used to facilitate attentional selection, but the strategy of using negative cues is more participant-dependent.

This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation – SFB/TRR 135, project number 222641018, TP B3).

43.417 DIFFERENTIAL EEG MARKERS OF SELECTIVE ATTENTION AND FEATURE BINDING IN VISUAL SEARCH

Jessica Ip[1] (jessica_ip@brown.edu), William C. Heindel[1], Elena K. Festa[1]; [1]Brown University

Previous studies have identified reliable electrophysiological markers of attentional processing within visual search tasks. However, none have identified markers that specifically reflect processes supporting feature binding within objects. This study aims to identify distinct EEG markers that are differentially sensitive to the feature binding and selective attention demands within a visual search paradigm. EEG was recorded while two conjunctive visual search tasks identical in their selective attention demands but differing in the demands placed on cross-cortical interactions were administered to 29 young adults (Mean Age=20.17, 22F). Participants were required to integrate a target's motion with either its luminance contrast (black or white) or its isoluminant color (red or green) to identify its location (left or right) among 0, 2, or 4 distractor items. Given that luminance and motion information are processed within the same dorsal cortical stream whereas color information is processed within the ventral cortical stream, the motion/color integration task places greater demands on cross-cortical interactions than does the motion/luminance integration task. Thus, these two search tasks differ only in terms of their feature binding demands. Both perceptual (P2, N2) and attentional markers (N2pc, P3b) were compared across conditions in posterior-occipital regions. Response times increased with set size but did not differ across conditions. Similarly, the magnitude of the P2 and N2 (complex pattern and motion-specific markers) and the N2pc (target selection marker) were modulated with set size but did not differ across tasks. In contrast, the contralateral P3b (a marker of top-down attentional enhancement) was unaffected by set size, but modulated across tasks, such that the amplitude was greater in the motion/color condition than in the motion/luminance condition. We believe that this modulation reflects increased reentrant attentional demands needed to maintain feature binding within the target in the across-stream binding condition.

43.418 REDUCED DISTRACTOR FILTERING WITH AGE: EVIDENCE FROM THE DISTRACTOR POSITIVITY ERP

Rosa E. Torres[1] (rt18dk@brocku.ca), Christine Salahub[2], Karen L. Campbell[1], Stephen M. Emrich[1]; [1]Brock University, [2]University Health Network

In our everyday lives, there are many instances during which we must guide our attention towards a goal while ignoring irrelevant information. In these situations, we rely on attentional control to engage cognitive resources necessary to ignore salient irrelevant distractors. Although this process can be facilitated by providing cues (e.g., a positive cue that indicates attend to blue), negative cues (indicating what to ignore) may initially bias attention towards distractors (Zhang et al., 2020). This may be especially the case in individuals with less efficient inhibitory control, such as anxious individuals (Salahub & Emrich, 2021) and older adults (Torres et al., 2023; Weeks et al., 2020) . To test the efficacy of target and distractor processing in a sample with lower inhibitory abilities, older adults' filtering performance was compared to that of younger adults during a search task while EEG was recorded. Participants were provided with either positive or negative pre-cues to indicate the feature of the target or distractor location, as well as a neutral control condition. The results indicate older adults only benefit from positive cues, as demonstrated by a higher mean amplitude of the N2pc component, as well as shorter reaction times, in response to lateral targets. However, in contrast to young adults, when presented with negative cues, older adults showed no Pd component to lateral distractors in any condition, suggesting that older adults did not inhibit the distracting information. These results suggest that older adults (with impaired inhibitory abilities) have particular difficulty suppressing distractors when a negative cue is provided, presumably because they have difficult disengaging attention from negatively cued items once it is directed there.

43.419 NEURAL RESPONSES TO NATURAL VERSUS AI-GENERATED AFFECTIVE IMAGES

Yujun Chen[1] (yujunchen@ufl.edu), Faith Gilbert[1], Ethan Smith[1], Ruogu Fang[1], Andreas Keil[1], Mingzhou Ding[1]; [1]UNIVERSITY OF FLORIDA

The International Affective Picture System (IAPS) contains 1,182 well-characterized photographs depicting natural scenes varying in affective content. These pictures are used extensively in affective neuroscience to investigate the neural correlates of emotional processing. Recently, in an effort to augment this dataset, we have begun to generate synthetic emotional images by combining IAPS pictures and diffusion-based AI models. The goal of this study is to compare the neural responses to IAPS pictures and matching AI-generated images. The stimulus set consisted of 60 IAPS pictures (20 pleasant, 20 neutral, 20 unpleasant) and 60 matching AI-generated images (20 pleasant, 20 neutral, 20 unpleasant). In a recording session, a total of 30 IAPS pictures and 30 matching AI-generated images were presented in random order, where each image was displayed for 3 seconds with neighboring images being separated by an interval of 2.8 to 3.5 seconds. Each experiment consisted of 10 recording sessions. The fMRI data was recorded on a 3T Siemens Prisma scanner. Pupil responses to image presentation were monitored using an MRI-compatible eyetracker. Our preliminary analysis of the fMRI data (N=3) showed that IAPS pictures and matching AI-generated images evoked similar neural responses in the visual cortex. In particular, MVPA (Multivariate Pattern Analysis) classifiers built to decode emotional categories from neural responses to IAPS pictures can be used to decode emotional categories from neural responses to AI-generated images and vice versa. Efforts to confirm these findings are underway by recruiting additional participants. Analysis is also being expanded to include the comparison of such measures as functional connectivity and pupillometry.

43.420 ACTIVATION PROFILES ACROSS NEURAL FEATURE DIMENSION MAPS GUIDE ATTENTION DURING VISUAL SEARCH

Daniel D. Thayer[1] (danielthayer@ucsb.edu), Thomas C. Sprague[1]; [1]University of California, Santa Barbara

Image-salient distractors and goal-relevant targets simultaneously compete for our attention. To resolve competition among items in the visual field, maps that preferentially respond to specific feature dimensions (e.g., color or motion) index salient and relevant locations based on their preferred feature, which is subsequently integrated into a feature-agnostic priority map, where the most important location guides attention. Even though feature dimension maps play a critical role in guiding attention, it is unclear how activation profiles in these neural maps resolve competition between task-relevant and irrelevant, but salient, items. Here, we used a visual search task to evaluate how relevant and salient items compete in neural feature dimension maps. On each trial, participants were cued to search for a target defined by a specific color or motion direction in a subsequent search array containing 8 colorful moving dot stimuli. All items in the array had homogenous features, except for the target item, which differed solely on the cued feature dimension. Occasionally, one of the non-target items was a salient distractor presented with either a different color or motion direction from the other items. We used an inverted encoding model to reconstruct spatial maps from activation patterns in feature-selective retinotopic regions (motion [TO1/TO2] and color [hV4/VO1/VO2] maps). Both targets and distractors were represented via spatially localized heightened activation in reconstructed spatial maps. Furthermore, we compared trials where the salient distractor likely captured attention (slow search RTs) to those where it was likely ignored (fast RTs). Activation within neural motion dimension maps reflected greater activation of the relevant target relative to the salient distractor when search was fast, and the inverse result when search was slow. These results indicate that neural feature dimension maps are crucial for computing attentional priority and that activation profiles across these maps guide behavior.

Funding: Alfred P Sloan Research Fellowship, National Eye Institute R01-EY035300

43.421 DECODING THE PROCESS OF MATCHING ATTENDED ITEMS TO TARGET TEMPLATES DURING VISUAL SEARCH

Xinger Yu[1] (xinger.yu@vanderbilt.edu), Geoff Woodman[1]; [1]Vanderbilt University

All major theories of visual attention posit a fundamental process wherein newly attended stimuli are systematically compared with target representations stored in memory to identify them as targets or reject them as distractors. In this study, we sought to test this proposal by decoding the electrophysiological brain activity as participants were cued to search for specific target objects in arrays. By strategically placing the target cues and the search elements at unique spatial locations, we were able to track when participants deployed external attention to select potential target items within the search array versus when they used internal attention to reactivate the representation of the sought-for target. We found evidence of the selection of the potential target object, followed by the reactivation of the target template held in memory. Further analyses demonstrated a correlation between the timing of target template reactivation and participants' behavioral reaction times, indicating that this reactivation process is closely linked to performing a detailed object comparison. Consequently, our study demonstrates how the electrophysiological brain activity provides a means to measure spotlights of attention as they shift between internal representations and the external environment, shedding light on the intricate mechanisms involved in the search for target objects.

Grants were provided by the National Science Foundation (BCS-2147064), and NEI (P30-EY08126 and T32-EY007135).

43.422 THE EFFECT OF BACKGROUND COMPLEXITY AND VARIABILITY ON VISUAL SEARCH IN VIDEO CONFERENCING DISPLAYS

Yelda Semizer[1] (yelda.semizer@njit.edu), Ruth Rosenholtz[2]; [1]New Jersey Institute of Technology, [2]Massachusetts Institute of Technology

The use of video conferencing tools increased significantly during the COVID-19 pandemic; however, there is little research on the visual factors that affect the video conferencing experience. For example, while searching for the current speaker, if a salient color is used as a highlight, it might "pop-out", increasing the efficiency of search by drawing attention. However, the level of saliency depends on the characteristics of the display, which in turn may impair the ability to predict the efficiency of search and to design a salient highlight, especially when the user has the freedom to control their own backgrounds during video conferencing. In this study, we investigated the effect of background complexity and variability on visual search performance using video conferencing displays. Observers searched either for the current speaker whose image, as in typical video conferencing tools, was highlighted (Experiment 1) or for a raised-hand symbol embedded on one of the backgrounds (Experiment 2) while we recorded their search time and tracked their eye movements. The number of virtual attendees in the display, the complexity and variability of attendees' backgrounds, and the presence of a target were manipulated. Results showed a classical set size effect and a significant effect of background complexity and variability, suggesting that search performance declined as the set size and visual complexity and variability increased. Our manipulation of background complexity was confirmed by a mathematical model of visual clutter. Further analysis revealed significant image-specific effects, suggesting that some backgrounds are sub-optimally designed from the point of perceptual processes.

Monday, May 20, 8:30 am – 12:30 pm, Pavilion

Spatial Vision: Crowding, eccentricity

43.423 MOTION DISCRIMINATION AROUND THE VISUAL FIELD DIFFERS BETWEEN ADULT HUMANS AND MACAQUES

Ekin Tünçok[1] (tuncokekin@gmail.com), Lynne Kiorpes[1], Marisa Carrasco[1]; [1]New York University

[Background and goal]. Visual discriminability varies with polar angle at isoeccentric locations. In human adults, performance is typically higher for stimuli along the horizontal than the vertical meridian (horizontal-vertical anisotropy, HVA), and along the lower than the upper vertical meridian (vertical meridian asymmetry, VMA). The macaque visual system is a prevalent animal model for understanding human vision. However, the natural-scene statistics may differ between these two species due to differences in bipedal locomotion and interactions with peri- and extra- personal space. Here we investigate whether polar angle asymmetries differ between adult human and macaque observers. [Method]. Adult macaque (n = 6) and human observers (n = 20) completed a 2AFC motion direction discrimination task at isoeccentric locations (7°) around the visual field. On every trial, a vertically-oriented Gabor target was presented (500 ms) among 7 equidistant distractors (Gabor patches with varying motion directions). Observers reported the motion direction of the target (left or right) whose location randomly varied among the 8 isoeccentric locations. [Results]. Performance was not homogenous as a function of polar angle for either observer group. Interestingly, the two groups of observers showed a striking difference in sensitivity (d'): For human observers, there was a pronounced VMA, a weaker HVA, and lower sensitivity at intercardinal than cardinal locations. For macaque observers, sensitivity was more homogenous as a function of polar angle, but was the poorest at the lower vertical meridian, showing an inverted VMA. [Conclusion]. The macaque visual system is used as an animal model to understand human vision, yet these two observer groups showed a pronounced difference in sensitivity around the visual field. The inverted VMA observed in macaques may reflect adaptive behavior by increasing discriminability at locations with greater relevance for visuomotor integration.

Funding: NIH R01-EY027401 to M.C.

43.424 PERFORMANCE DIFFERENCES AROUND POLAR ANGLE VARY SYSTEMATICALLY ACROSS EXPERIMENTAL CONDITIONS

David Tu[1] (dvid.tu@gmail.com), Shutian Xue[1], Marisa Carrasco[1]; [1]New York University

[Background] Visual performance varies at isoeccentric location around polar angles–it is better along the horizontal than vertical meridian (horizontal-vertical anisotropy, HVA) and at the lower than upper vertical meridian (vertical meridian asymmetry, VMA). Here, we examined whether the magnitude of these two asymmetries (1) differ from and (2) correlate with each other across multiple published studies. [Method] We calculated the magnitude of HVA and VMA across multiple studies using an orientation discrimination task, which varied in stimulus properties (e.g., eccentricity, stimulus size and degree of tilt) and reported measurements (e.g., contrast sensitivity and acuity). First, we assessed whether they alter the magnitude of HVA and VMA. Second, we assessed the correlation between the magnitude of HVA and VMA while controlling for stimulus properties and measurement. [Results] (1) HVA is stronger than VMA across stimulus properties (eccentricity and stimulus size), especially for contrast sensitivity-related measurements. (2) The extent of the HVA did not correlate with that of the VMA while controlling for stimulus properties and measurements. [Conclusion] Despite the variability in how performance for fundamental visual dimensions vary around polar angle across studies, an analysis across multiple studies shows that (1) HVA is more robust than VMA for some measures (e.g., contrast sensitivity). (2) HVA did not correlate with VMA. Results suggest that these two visual field asymmetries may derive from distinct neural computations

Funding: NIH R01-EY027401 to M.C.

43.425 NO POOLING, NO AVERAGING: HOW VARYING THE NUMBER OF IDENTICAL GABORS MODULATES ORIENTATION DISCRIMINATION IN THE PERIPHERY

Miao Li[1] (miao.li@univ-lille.fr), Sabrina Hansmann-Roth[2], Wolf M. Harmening[3], Bilge Sayim[1]; [1]CNRS, University of Lille, [2]University of Iceland, [3]University of Bonn

In crowding, flankers interfere with target perception. Pooling models suggest that features of the target and the flankers are averaged. A typical example is the (compulsory) averaging of orientation signals of peripherally presented Gabors. Here, we show that averaging accounts fail when the Gabors have identical orientations. We measured orientation discrimination thresholds for peripherally presented single and identical Gabors arranged in 'snake-like' and 'ladder-like' arrays (1, 2, 3, 5 or 7 Gabors). Gabor arrays were arranged radially or tangentially, typically yielding strong and weak crowding, respectively. In Experiment 1, participants were not informed about the identical orientations; in Experiment 2 they were informed prior to the experiment. In Experiment 3, participants reported the orientation of the innermost Gabor, and indicated whether all Gabors had the same orientation. In the radial conditions, where strong crowding was expected, we found a strong increase of thresholds with increasing numbers of Gabors in the radial snake conditions in both Experiments 1 and 2. Thresholds for radial ladders did not vary with the number of Gabors in Experiment 1, but increased slightly for larger numbers of Gabors in Experiment 2. In the tangential condition, thresholds did not vary significantly with the number of Gabors. In Experiment 3, observers frequently reported varying orientations of the presented (identical) Gabors. Reported orientation variation was higher in the radial snake (~50% of the trials) than ladder condition (~20%). A control experiment where participants adjusted the perceived orientation of each Gabor confirmed larger variance in the snake than in the ladder condition. Our results show that neither the number nor the orientation of Gabor arrays alone predicts performance when the Gabors have identical orientations. Observers perceived varying orientations in uniform displays. Averaging accounts fail when targets have identical orientations.

43.426 PERIPHERAL CROWDING MAGNITUDE IS SIMILAR UNDER PHOTOPIC AND SCOTOPIC LUMINANCE CONDITIONS

Dilce Tanriverdi[1] (d.tanriverdi@umcg.nl), Nomdo M. Jansonius[1], Frans W. Cornelissen[1]; [1]University Medical Center Groningen, Department of Ophthalmology, Laboratory of Experimental Ophthalmology

Visual crowding is the inability to distinguish a stimulus (target) in the presence of surrounding objects (flankers) and impacts various visual functions. Some of these functions are influenced by the luminance level of the environment, such as visual acuity, and contrast sensitivity. However, the potential interactions between crowding and luminance remain unclear. In this study, we explored the magnitude of crowding under five different luminance conditions, ranging from photopic to scotopic levels. Eight participants were asked to report the orientation of the gap in a white Landolt-C (2°diameter) on a gray background positioned at 10° eccentricity. The target Landolt-C was either presented alone or flanked by four Landolt-C's (2°diameter) placed at one of five center-to-center distances (2°,2.52°,3.22°,4.18°,5.5°). Participants responded by adjusting the orientation of a reference Landolt-C gap (5°diameter) placed at 0° eccentricity. Goggles with neutral density filters were used to manipulate stimulus luminance (0.02 to 200 cd/m²). Landolt-C gap sizes were individually adjusted to equalize acuity performance at isolated levels across luminance conditions. Participants' perceptual error (PE) was calculated based on the difference between their response and the actual orientation of the target gap. Crowding magnitude was defined as PE flanked - PE isolated for each luminance and distance conditions. Results showed that crowding magnitude was similar under all luminance conditions (BF10=0.23), while a monotonous decrease in crowding magnitude was revealed as the target-flanker distance increased in all luminance conditions. Moreover, participants' PEs were similar in all isolated conditions, confirming similar acuity performances (BF10=0.18). Our results support the idea that the same neural mechanisms operate under scotopic and photopic conditions for peripheral crowding. Additionally, our results align with prior research on contour interaction in scotopic conditions, suggesting a shared underlying process between contour interaction and crowding.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 955590

43.427 THE CHROMATIC AND ACHROMATIC CONTRAST SENSITIVITY FUNCTION IN THE FAR PERIPHERY

Norick Bowers[1] (norick.bowers@psychol.uni-giessen.de), Karl Gegenfurtner[1], Alex Goettker[1]; [1]Justus-Liebig-Universität Gießen

The contrast sensitivity function (CSF) defines one of the most fundamental aspects of vision: the boundary between the visible and the invisible. The CSF has been studied extensively since the 1950s, however most studies to date focused on the central region of vision and few have measured the CSF in the periphery. Furthermore, less investigation has been done examining the CSF with chromatic, as opposed to luminance, stimuli. The current study aims to fill this gap by examining the CSF across the entire horizontal visual field with achromatic and chromatic stimuli. Three curved wide-field displays were arranged in a semi-circle around the subjects. These monitors have a 1m circular curvature, so each point of the screen is equidistant from the subject, and combining all three allows for an effective screen size of 205°. A left/right 2AFC detection task was used to measure the CSF in 3 different experiments (Achromatic, Red-Green, and Yellow-Violet) via QUEST. Participants binocularly viewed the stimuli, which were 10deg Gabor patches of varying spatial frequency (0.1, 0.3, 0.5, 1, 3, 5, and 10 cpd). The stimuli were presented at 6 different locations along the horizontal visual field (5°, 15°, 45°, 60°, 75°, and 90° eccentricity). The measured CSFs in the more central regions of the visual field agreed with measurements from previous literature. However, achromatic sensitivities in the far periphery (>45°) were higher than models that account for peripheral contrast sensitivity would suggest. CSF models to date do not typically account for chromaticity and eccentricity together, making these data a valuable contribution to extend these models. These data will fill in an important gap in our understanding of human visual sensitivity that is vital for applications in full-field visual displays such as VR and AR, as well as cutting-edge technologies such as foveated rendering.

This work was support by the Alexander von Humboldt-Stiftung

43.428 EXPLORING THE RELATIONSHIP BETWEEN CONE DENSITY AND VISUAL CROWDING IN THE CENTRAL FOVEA

Krish Prahalad[1] (pkrishn5@ur.rochester.edu), Ashley M. Clark[1], Benjamin Moon[1], Austin Roorda[2], Pavan Tiruveedhula[2], Wolf Harmening[3], Aleksandr Gutnikov[3], Samantha K. Jenks[1], Sanjana Kapisthalam[1], Michele Rucci[1], Jannick P. Rolland[1], Martina Poletti[1]; [1]University of Rochester, [2]University of California, Berkeley, [3]University of Bonn

Visual crowding, the interference in target recognition caused by surrounding similar objects, occurs not only in the periphery but also at the very center of gaze, where visual resolution is highest. It is unclear what factors contribute to crowding at this scale and to what extent cone density defines crowding thresholds. Here we investigated this using an Adaptive Optics Scanning Laser Ophthalmoscope

(AOSLO) to stimulate the retina while bypassing optical limitations and maintaining the stimulus at a fixed retinal location despite the presence of fixational eye movements. Subjects (N=6) participated in a 4AFC digit identification task using Pelli's font, specifically designed to study foveal crowding. The target, flanked by other digits along the horizontal meridian, was presented for 500 ms and maintained at the preferred retinal locus (PRL). Flankers were set at different distances using a method of constant stimuli. Stimuli sizes were set to three times the thresholds obtained using a QUEST procedure with unflanked (isolated) stimuli under retinal stabilization. Target width covered approximately 4 cones in each subject. Our findings revealed the typical crowding effect, with critical spacings (maximum interference due to flankers) ranging from 0.52 to 1.93 arcminutes, about 10-100 times smaller than what observed extrafoveally. We then identified the cones covered by the stimulus array and assessed the average cone size. The ratio between critical spacing measures to cone diameter was on average 1.78±1.07. Behavioral performance, when spacing matched half cone diameter, dropped by an average of 50% compared to isolated target presentation. Mislocalization errors, typical in extrafoveal crowding, where subjects report either flanker on incorrect trials, did not exceed chance level. These results suggest information pooling for foveal crowding involves at most two cones. Hence, critical spacing could potentially be inferred from the average cone diameter at the PRL.

Research supported by R01 EY029788, R01 EY018363, EY001319, R01 EY023591, Ha5323/6-1 and Ha5323/8-1

43.429 VISUAL SENSITIVITY IN THE FOVEOLA IS SPATIALLY NON-UNIFORM

Lynn Schmittwilken[1,2] (l.schmittwilken@tu-berlin.de), Michele Rucci[1]; [1]University of Rochester, [2]Technische Universität Berlin

The foveola is the retinal region free from capillaries where cones are most densely packed. Although it only covers a tiny portion of the visual field (~1 degrees), foveal representation is disproportionally large throughout the visual hierarchy. Yet little is known about foveal sensitivity, primarily because of the technical challenges inherent in studying vision at such a small scale. The continual presence of eye movements (often at the resolution limit of standard eye-trackers) and the difficulty in accurate gaze localization, have traditionally prevented testing of adjacent retinal locations in isolation and contributed to the assumption that foveal sensitivity is approximately uniform. Building upon recent advances in eye-tracking, gaze-contingent display control, and localization of the line of sight, here we mapped visual sensitivity at 13 locations within the foveola of healthy observers. As in standard perimetry tests, subjects were asked to detect a small probe (5 arcminutes) briefly displayed (50 ms) over a uniform background (10 cd/m2). Unlike standard perimetry, however, probes were confined to the central 1 degree of the visual field. Probe locations were updated in real-time to compensate for eye movements, which were continually monitored by means of a digital Dual Purkinje image eye-tracker. Subjects reported their responses via button presses and received auditory feedback. By systematically varying the contrast of the probes, we measured contrast sensitivity functions at each location. Our results show that sensitivity is not uniform within the foveola. While idiosyncratic differences occur, peak sensitivity is consistently shifted towards the temporal visual field and favors the horizontal meridian relative to the vertical one. Furthermore, sensitivity progressively deteriorates with increasing eccentricity from the locus of peak sensitivity.

Supported by NIH grant EY018363 and P30 EY001319

43.430 REDUCED FOVEAL CROWDING CONTRIBUTES TO IMPROVEMENT IN ACUITY FOLLOWING FLICKER ADAPTATION

Selassie Tagoh[1] (stag209@aucklanduni.ac.nz), Samuel Schwarzkopf[1,2], Steven Dakin[1,2]; [1]The University of Auckland, [2]University College London

Flicker adaptation improves acuity for briefly presented (masked) stimuli (Arnold et al, PNAS, v113 p12556-61, 2016). We investigated the role that fixation-stability, pupil-size and susceptibility to foveal crowding might play in this effect. We measured visual acuity by having observers (n = 20) report the orientation of a white tumbling-“T” target. Observers were either unadapted or had adapted to a 60Hz flickering noise pattern (30s first trial, 4s trial-by-trial top-up). The target was presented for 110ms (immediately followed by a mask) and could appear in isolation or flanked by a pair of randomly oriented “T”s. A second cohort of observers (n = 35) performed the acuity task (flanked target only), while undergoing eye-tracking. This stimulus arrangement led to high levels of foveal crowding with recognition acuity for flanked letters (0.049 LogMAR) being around 0.16 LogMAR worse than for unflanked letters. Adaptation to flicker moderately improved visual acuity for flanked (-0.038 LogMAR) but not for unflanked (-0.008 LogMAR) optotypes. The magnitude of acuity-improvement was correlated with individuals' (unadapted) susceptibility to crowding; those more susceptible to crowding showed greater improvement (r = -0.55, p = 0.006). In conclusion, we confirm that flicker adaptation improves acuity for briefly presented crowded letters (Arnold et al., 2016). While we agree with suggestions that these improvements arise from a reduction in sensitivity to transient low spatial frequency image structure, our results demonstrate that this may improve acuity at least in part by reducing foveal crowding.

Acknowledgement: Supported by the Marsden fund council from government funding, managed by Royal Society, Te Apārangi.

43.431 CROWDING AND VISUAL APPEARANCE IN AMBLYOPIA

Ângela Gomes Tomaz[1], Sunwoo Kwon[1], Dennis M. Levi[1], Wolf M. Harmening[2], Bilge Sayim[3]; [1]Herbert Wertheim School of Optometry & Vision Science, University of California, Berkeley, Berkeley, United States of America, [2]Adaptive Optics and Vision Laboratory, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany, [3]Cognitive and Affective Sciences Lab (SCALab), UMR CNRS 9193, University of Lille, Lille, France

Persons with amblyopia, a neurodevelopmental disorder of spatial vision characterized by reduced visual acuity (VA), perceive stimuli distorted (e.g., Barrett et al., 2002, Pugh, 1958) and have stronger foveal crowding (i.e., worse performance with flanked targets) than neurotypical controls. Here, we explored the appearance of both crowded (with black bars on all sides) and isolated high contrast letters in observers with amblyopia and controls. Stimuli were presented in the fovea for 500ms with stimulus size set at 1.0 and 1.5 x the observer's VA threshold (measured with isolated targets). Observers viewed the target monocularly with their dominant/fellow (DE/FE) or non-dominant/amblyopic (NDE/AE) eye and replicated its appearance on a 9x9 square-grid interface with binocular viewing. Each square in the interface could be turned on and off with mouse clicks and their grey levels adjusted with the mouse wheel. Results showed that response-target structural similarity (RTSS) was lower for 1.0 than 1.5 x threshold sized stimuli, for flanked than isolated targets and for observers with amblyopia than controls. No differences were found between DE/FE and NDE/AE viewing conditions in amblyopic observers. However, their RTSS in both eye conditions were lower than controls' NDE condition. Surprisingly, the controls' RTSS was lower with DE than with NDE viewing. There was no interaction between groups (Amblyopic, Controls) and crowding conditions (flanked, isolated). When compared to the target letter, responses revealed truncation and extension of elements, fusion of separate elements (e.g., connecting the flankers) and shape distortions (e.g., depicting straight lines as curves) for both groups in all conditions. Taken together, our results reveal characteristics of target appearance in both groups showing stronger visual distortions in persons with amblyopia than controls, but no crowding differences between groups. These results highlight important appearance differences and similarities between normal and amblyopic visual perception when controlling for VA differences.

Monday, May 20, 8:30 am – 12:30 pm, Pavilion

Scene Perception: Categorization

43.432 A COMPARISON OF TASKS FOR CONSTRUCTING THE CATEGORY SPACE OF NATURAL SCENES

Pei-Ling Yang[1] (plyang2@illinois.edu), Diane M. Beck[1]; [1]University of Illinois Urbana-Champaign

Similarity is a core construct underlying various prominent category theories. This study compares the outcomes of two classic similarity judgment tasks—a same-different judgment task (implicit similarity task) and a spatial arrangement task (explicit similarity task)—for natural scene categories. We ask whether or not the two methods produce the same category space, that is, the similarity structure within and between categories. The same natural scene set (four categories: beach, city, highway, mountain; from Torralbo et al., 2013) was used for the two tasks. The same different judgement task (N = 218) required subjects to respond whether two side-by-side presented natural scenes belong to the same category or not. The similarity of the scene pair is implied by the response times of the judgement. The spatial arrangement task (N = 37) asked subjects to sort 24 scenes such that more similar scenes are closer scenes to each other. Two different sorting lists were included: within category (i.e., 24 images were all from the same category) and between category (i.e., six images per category). The multidimensional scaling (MDS) results were computed for both tasks, and the image distances were derived based on the MDS coordinates. We correlated the arrangement of each trial (i.e. 24 scenes) with the average MDS from the same-different task. Results showed that two tasks were significantly albeit weakly correlated for both within and between category distances (within category: r = .10; between category: r = .18). This suggests that while these implicit and explicit tasks might construct approximately similar spaces, there are still considerable differences, raising the question of whether we can use the tasks interchangeably to learn about the structure of natural scene category space.

43.433 A TURN OF EVENTS: UPSIDE-DOWN FIGURE JUDGMENTS SUPPORT A VERTICAL ATTENTION BIAS TOWARD AFFORDANCE LOCATION

Matthew Langley[1] (mdlangle@asu.edu), Madelaine Vu[1], Michael McBeath[1,2]; [1]Arizona State University, [2]Max Planck Institute for Empirical Aesthetics

Introduction: A robust Vertical Attention Bias (VAB) that directs attention toward tops of objects and bottoms of scenes was recently proposed and confirmed in both adults and 4-7 year-old children. This bias is consistent with Affordance Theory in that observers favor attending to the more informative locations of intentionality and functionally in the environment (locations relevant to our limbs and effectivities), such that we generally adopt a downward gaze. In the current analysis, we examine the effects of generic directionality by manipulating the vertical orientation of presented stimuli and test for an overall pattern of VAB for both object and scene images. Method: Participants observed picture triptychs and made similarity judgments between a central target object or scene and flanking comparison images that contained the same top- vs same bottom-half as the target image. Experiment 1 presented picture triptychs in an upright orientation, and Experiment 2 the same triptychs in an inverted upside-down orientation. This manipulated the affordance location, such that object tops were now positioned in the lower half of the image, and scene bottoms in the upper half of the image. Results: Results replicated past findings confirming a VAB for object tops and scene bottoms that varies as a function of informative aspects of visually attended stimuli. Here Experiment 1 followed the previous pattern while the results of Experiment 2 were inverted and somewhat weaker. Conclusion: The findings support our hypothesis that the VAB is driven by the regularity of affordance location. The overall VAB pattern for upside-down stimuli was maintained, though statistically weaker than for an upright orientation, particularly for object images. Taken together, the findings support that vertical information imbalance drives a generic downward vantage tendency that focuses attention on personal action space and body-level affordances.

43.434 ARE ABSTRACT RELATIONAL ROLES ENCODED VISUALLY? EVIDENCE FROM PRIMING EFFECTS

Alexander A Petrov[1] (petrov.11@osu.edu), Yuhui Du[1]; [1]Ohio State University

Whereas it is well established that the visual system encodes comparative relations such as Longer and spatial relations such as Above, it remains controversial whether it encodes abstract relational roles such as Agent and Patient in visual events. The present experiment tested whether the latency of a color localization task is affected by the abstract role bindings of a preceding event. Method: Each trial included a static target image preceded either by a brief silent video of a priming event (on Session 1) or by an audio-visual presentation of an English sentence describing the same event (on Session 2). Example sentence: 'The red goat on the left knocked down the blue goat on the right." There were 64 videos counterbalancing 4 event types (launching, deforming, breaking, and a relationally ambiguous control), 4 animal species, 2 role bindings (Agent is red vs blue), and 2 locations (red on the left vs right). The set of static targets were the final frames of the same videos. The role bindings were either repeated, switched, or ambiguous across the target and prime, counterbalanced across trials. The mask-enforced presentation duration of the target was calibrated individually to ensure 90% accuracy of the color localization task. Dependent variable is latency on correct trials. 15 participants x 2 sessions x 768 trials/session. Results: Whereas the role bindings of the linguistic primes had no statistically significant effect on the latency of the visual task, the role bindings of the video primes did have an effect: The latency on unambiguous trials (on which both target and prime had well-defined Agents and Patients) was significantly greater than that on ambiguous trials (on which at least one component lacked clear relational roles). This is evidence that the visual system is sensitive to (the ambiguity of) the role bindings of abstract relations.

43.435 CHARACTERIZING IDIOSYNCRASIES IN PERCEPTION AND NEURAL REPRESENTATION OF REAL-WORLD SCENES

Gongting Wang[1,2] (generalwgt@gmail.com), Matthew Foxwell[4], Lixiang Chen[1], David Pitcher[4], Radoslaw Martin Cichy[1], Daniel Kaiser[2,3]; [1]Department of Education and Psychology, Freie Universität Berlin, [2]Department of Mathematics and Computer Science, Physics, Geography, Justus Liebig University Gießen, [3]Center for Mind, Brain and Behavior, Justus Liebig University Gießen and Philipps University Marburg, [4]Department of Psychology, University of York

The efficiency of visual perception is not solely determined by the structure of the visual input. It also depends on our expectations, derived from internal models of the world. Given individual differences in visual experience and brain architecture, it is likely that such internal models differ systematically across the population. Yet, we have no clear understanding of how such differences shape the individual nature of perception. Here, we present a novel approach that uses drawing to directly access the contents of internal models in individual participants. Participants were first asked to draw typical versions of different scene categories (e.g., a kitchen or a living room), taken as descriptors of their internal models. These drawings were converted into standardized 3d renders to control for differences in drawing ability and style. During the subsequent experiments, participants viewed renders that were either based on their drawings (and thus similar to their internal model), based on other people's drawings, or based on arbitrary scenes they were asked to copy (thereby controlling for memory effects). In a series of behavioral experiments, we show that participants more accurately categorize briefly presented scene renders when they are more similar to their personal internal models. This suggests the efficiency of scene categorization is determined by how well the inputs resemble individual participants' internal scene models. Using multivariate decoding on EEG data, we further demonstrate that similarity to internal models enhances the cortical representation of scenes, starting from perceptual processing at around 200ms. A deep neural network modelling analysis on the EEG data suggests that scenes that are more similar to participants' internal models are processed in more idiosyncratic ways, rendering representations less faithful to visual features. Together, our results demonstrate that differences in internal models determine the personal nature of perception and neural representation.

European Research Council (ERC) starting grant (ERC-2022-STG 101076057); "The Adaptive Mind", funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art; China Scholarship Council (CSC)

43.436 COME HERE OFTEN? HOW FAMILIARITY AFFECTS SCENE PROCESSING

Elissa Aminoff[1] (eaminoff@fordham.edu), Bridget Meighan[1]; [1]Fordham University

Throughout our day, we understand the surrounding visual environment in a process that occurs quickly and automatically. However, if someone frequently encounters an environment, do they process it differently from an environment they rarely experience? In the current study, we examined whether familiarity with a context would affect one's scene perception. Scene perception was queried using a paradigm where on each trial two scenes were presented briefly, one after the other, and the task was to judge whether the scenes were the same or different. The two scenes were either exactly the same, two different exemplars of the same context, or two completely different scenes. Across two experiments, we examined how familiarity with the context affected performance. In experiment 1 (N = 145) we examined in separate models how one's experience with nature, streets, long-distance travel (e.g., airplanes), and COVID-19 pandemic closures (i.e., mandated lock-downs in 2020-2021), affected their performance on the respective categories. Across all four models we found an effect of experience. The more frequently a context was experienced, the better one was at identifying whether a scene was the same. In contrast, performance on frequently encountered scenes in the different conditions were worse, assessed either through accuracy or longer reaction times. Thus, the less familiar the scene, the better one was at identifying differences between scenes. In experiment 2 (N = 49), we tested this by examining the scene perception of students with their home university campus compared to the perception of an unknown university campus. Results of this study replicated the results from experiment 1. Together, we suggest that familiarity of a scene results in fast processing of a scene likely via a gist representation, which may result in the glossing over of details that are processed in an unfamiliar scene.

43.437 CROSSING CATEGORY BOUNDARIES: PERCEPTUAL HYSTERESIS FOR SCENES EVEN WITH ENDPOINT PREVIEW

Huiqin Chen[1] (huiqin.chen@mail.utoronto.ca), Mei Yang[1], Gaeun Son[1], Dirk Bernhardt-Walther[1]; [1]University of Toronto

In dynamically changing environments, how does our visual system make quick perceptual decisions based on information that is actively changing? The hysteresis effect suggests that when resolving perceptual ambiguity, people tend to stick with their current interpretation of sensory information and find it difficult to change their perception until there is a noticeable and significant change in input. We here demonstrate this effect for dynamically changing scenes. We further explored how prior exposure to target scene would affect the hysteresis effect. We used indoor scene images generated with a Generative Adversarial Network to create smooth yet realistic transitions between scene categories. Participants were asked to report when they perceived a shift in category during these transitions. Before each trial, participants were provided with information about the

scene category at the end of the transition, either in the form of words, images, or both. Each transition was repeated in both directions (A to B and B to A), and the differences in responses between the two opposite directions were analyzed. Our findings indicated that exposure to words or images had no effect on perceptual hysteresis. Even when participants had knowledge of the target category, their perception of the image categories was still biased towards the initial category. This result suggests that neither semantic knowledge nor visual representation of the future can influence the hysteresis effect. The top-down knowledge of the future direction does not impact or eliminate the conservatism of the visual system. This has important implications, indicating that hysteresis is an innate characteristic of the visual system and is not easily influenced by higher-level control. Thus, perceptual conservatism is likely instrumental for the apparent stability of visual perception in most real-world settings.

This work was supported by an NSERC Discovery Grant (RGPIN-2020-04097) to DBW.

43.439 EXAMINING THE EFFECTS OF REAL-WORLD EXPERIENCE ON LAB-BASED SCENE MEMORY

Maria S. Orlando[1] (morlando@yorku.ca), Alberto Umiltà[2], Federico Fornaciari[2], Elisa Ciaramelli[2], R. Shayna Rosenbaum[1]; [1]York University, [2]University of Bologna

Boundary extension (BE) is as an error in scene memory, such that participants retrieve details beyond the given boundaries of a scene image. Boundary contraction (BC) is the opposite effect, whereby participants retrieve less context within the boundaries of a given scene image. In the BE literature, there is variability in the types of stimuli that are used, how BE is tested, and the proposed mechanisms underlying the phenomenon. Some research supports the view that BE reflects (re)construction of the scene from an internal representation that was formed, whereas other research supports the view that BE (and BC) emerge from image-based properties, including the number of central objects and whether an object is pictured in close range or from a wider angle. Assessing the effects of prior knowledge and experience of a scene on this bias can help disentangle the role of visual perception and scene construction. The current study tested the influence of familiarity on scene recognition through the comparison of lab-based encoding of images of pre-experimentally familiar (real-world) places with images of unfamiliar places. Participants used a continuous rating scale to indicate how they perceived the boundaries of a test image relative to a previously studied image. There was a tendency for BC across both image conditions, with evidence of maintained, and an instance of greater, BC for familiar than unfamiliar scene images. Importantly, the lack of evidence for increased BE with greater familiarity favours an image-based theoretical account of BE and BC.

Thank you to Vision: Science to Applications (VISTA) at York University for funding this project

43.440 MEMORY DISTORTION OF PITCH ANGLE IN REAL-WORLD SCENES

Yichen Wu[1,2] (yichensamwu@outlook.com), Sheng Li[1,2,3,4,5]; [1]School of Psychological and Cognitive Sciences, Peking University, Beijing, China, [2]PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China, [3]Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China, [4]Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China, [5]National Key Laboratory of General Artificial Intelligence, Peking University, Beijing, China

Scene memory is prone to systematic distortions. A well-known distortion along the near-far axis of the three-dimensional space, called boundary transformation, demonstrates the observer's erroneous recall of the viewing distance of scenes. The investigators argued for the role of the normalization process to the high-probability prototypical viewing distance in boundary transformation. Here, we hypothesized that the prototypical viewpoint also includes the dimension of vertical angle of view (vAOV) and may cause memory distortion along the scenes' vertical axis. In two behavioral experiments, we found a systematic memory distortion in vAOV in both force choice ($n = 79$) and free adjustment ($n = 30$) tasks. The force choice task used the classic RSVP paradigm of boundary transformation. Subject were told to detect the vAOV change of two identical scenes. The free adjustment task adapted the RSVP paradigm and allowed subjects to adjust the scene's pitch, yaw, and depth to relive their memory. Computational modeling suggests that the vAOV biases could be predicted by the asymmetry of the complexity information in the vertical direction of the scenes and the subjective vAOV ratings from a group of online participants. These behavioral findings are in line with the normalization process. We further conducted an fMRI and an MEG experiment with the RSVP paradigm to explore the neural mechanism of vAOV bias. The fMRI experiment ($n = 24$) revealed the involvement of V4, PPA, and OPA in the normalization process. The results of MEG ($n = 20$) experiment provided temporal evidence that memory distortion appeared approximately 140 ms after scene onset and was accompanied by the estimation of low-level visual complexity. These findings suggest that the memory distortion is embedded in the feedforward sweep of scene processing. Estimation of the complexity information could serve as the initial step of vAOV bias.

43.441 PERCEIVING NATURAL IMAGES MAY CONSUME LESS COGNITIVE RESOURCES: EVIDENCE FROM IMAGE MEMORABILITY, EDGE MAGNITUDES, AND SPECTRAL CONTENT

Nakwon Rim[1], Omid Kardan[2], Sanjay Krishnan[1], Wilma A. Bainbridge[1], Marc G. Berman[1]; [1]University of Chicago, [2]University of Michigan

Past work has suggested that perceiving natural scenes requires less cognitive resources compared to perceiving urban scenes, which may explain the cognitive benefits attained after interacting with natural environments. While studies have shown that natural environments have restorative benefits, the hypothesized mechanisms have not been rigorously tested. Here, we investigated whether perceiving natural scenes may consume less cognitive resources. First, we conducted a continuous recognition task to probe the memorability of images and found that natural images are less remembered, which suggests that fewer cognitive resources are used to process natural stimulation. Next, using a Canny edge detection algorithm, we analyzed the number and significance of edges in the images and found that the proportion of "strong" edges (edges with higher gradient magnitude) is smaller in natural images. This association suggests that the number of edges that are essential to capture the information of a scene is smaller for natural images, which aligns with theories that perceiving natural images consumes less cognitive resources. Finally, we analyzed the spectral properties of the images by applying a discrete cosine transform to 8 x 8 pixel tiles (similar to JPEG compression). We found that natural scenes have a larger proportion of their spectral energy in high-frequency coefficients compared to man-made scenes. Since the human visual system may be less sensitive to high-frequency information, this result implies that natural images have less information that will be processed by humans, which is consistent with the idea that perceiving natural scenes will be less taxing. In conclusion, we found that natural scenes are less memorable, have a lower proportion of strong edges, and contain more high-frequency information compared to man-made scene images. These findings are consistent with theories that posit that natural images are less taxing to process.

This research was supported by National Science Foundation Smart and Connected Communities grant (NSF S&CC) #1952050 to MGB.

43.442 SPONTANEOUS ASSOCIATIVE THOUGHT FACILITATES SCENE-GIST MEMORY

Shira Baror[1] (baror.shira@gmail.com), Elissa Aminoff[2], Yoed Kenett[3]; [1]Edmond and Lily Safra Center for Brain Sciences, Hebrew University, Jerusalem, Israel, [2]Department of Psychology, Fordham University, Bronx, NY, USA, [3]Faculty of Data and Decision Sciences, Technion — Israel Institute of Technology, Haifa, Israel

Spontaneous associative processes (e.g., mind wandering, spontaneous memory recollection) are prevalent in everyday life, yet their influence on perceptual memory is under debate. Given that scene perception involves extraction of contextual associations, we hypothesized that associative thought would enhance scene memory by promoting encoding of contextual associations. In an online experiment (n=75), participants viewed scenes, and after each scene presentation participants either generated chained-free associations (associative processing), or, as control, listed words that begin with a specific letter (phonological processing). Scene memory was tested after an intermediate creativity task, which is also shown to rely on associative processes. Results revealed that associative thought, regardless of the conceptual distances involved, enhanced scene-gist memory, but hampered memory of scene details, implying that associative thought facilitates contextual encoding. At the same time, conceptual distances in the semantic spaces in scene-based associations were positively correlated with creativity, extending the creativity-associativity link to the scene-based domain.

43.443 THE IMPACT OF SEMANTIC DESCRIPTIONS ON LEARNING OBJECT-TO-OBJECT RELATIONSHIPS IN A SCENE

Victoria Nicholls[1] (nicholls@psych.uni-frankfurt.de), Kim Kira Philippi[1], Mohamed Abdelrahman[1], Jo Ackerman[1], Franka Schulze[1], Melissa Le-Hoa Vo[1]; [1]Goethe University Frankfurt

Our knowledge of scenes is thought to have a hierarchical structure, at the lowest level are local objects, often smaller objects such as soap. Followed by anchor objects, often larger objects, e.g. sinks. Local and anchor objects together e.g. soap on a sink, form a phrase. Phrases can have multiple local objects (co-locals). Multiple phrases combined form a scene. What is not clear is how we learn this structure, can this be learned with visual associations alone or is semantic object information required? To examine this, we performed two experiments. In the learning phase of the first experiment participants were presented with objects in isolation with audio descriptions of the object functions, or with non-descriptive audio. This was followed by two recall phases, the first where participants were presented with two objects which they rated on a scale from 1-9 how likely the objects would be grouped together based on the descriptions they received in the learning phase. In the second recall phase participants were shown a scene image containing all the objects and participants grouped the objects into phrases based on the object descriptions received. In the learning phase of the second experiment participants viewed videos of phrases in scenes, where each object was highlighted, along with descriptive or non-descriptive audio. This was followed by the same rating recall phase as in the first experiment. In the video conditions we found that participants learned the anchor-local relationships even with non-descriptive audio, while descriptive audio boosted learning the local- to-local relationships. This suggests that hierarchical scene knowledge can be learned through visual associations but the detail of the knowledge can be improved with the inclusion of semantic information such as descriptions of functions the objects perform together.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG-FOR German Research Foundation;Arena)

43.444 UNDERSTANDING NOVEL, REAL WORLD SCENES: GIST STRENGTH AND THE RICHES BEYOND

Adison Roman[1] (adisonmaryroman@usf.edu), Reilly Orman[1], Lewis Evans[1], Chad Hense[1], Aaron Sharrett[1], Skyler Torres[1], Gene Kim[1], Thomas Sanocki[1]; [1]University of South Florida

What happens when a full-color scene is shown to observers for 2 sec, followed by a written report describing what was perceived? Such full-report methods reveal that there is a shared perception of gist, as well as deeper perception that extends far beyond gist. We applied the method to a new set of 8 naturalistic scenes. The responses were accurate descriptions of stimulus content, validated against the scenes. Gist strength was measured by the frequency of the most common concepts in the scene-descriptions, and exceeded 80% for most scenes, but not all. The new findings are what is beyond gist: elaborations of the scenes, which constitute breadth of interpretation. Breadth was measured by counting distinct “aspects of interpretation” in the descriptions of each scene — aspects having have noticeably different and correct meanings. The number of aspects varied widely across scenes, ranging from 10 aspects for one scene (6 surfers riding a wave), to over 40 aspects for two otherwise ordinary scenes (a bench in a room, and an old wall and door). The aspects differed in multiple ways, including main objects and features, size scale, contextualization and attitude, and labels. The differing aspects suggest that perceptual experience of the world is partly idiosyncratic, driven by qualities of the perceiver. In other words, people see (somewhat) differently.

Attention: Tracking, shifting

43.445 ATTENTIONAL EFFECT IN MOTION-INDUCED POSITION SHIFT

Mohammad Shams[1,2] (m.shamsahmar@gmail.com), Peter Kohler[2,3], Patrick Cavanagh[1,2]; [1]Department of Psychology, Glendon College, York University, Toronto, ON, Canada, [2]Centre for Vision Research, York University, Toronto, ON, Canada, [3]Department of Psychology, York University, Toronto, ON, Canada

A moving bar shifts the perceived position of a nearby flash. This effect is strongly asymmetrical: a flash at or ahead of the bar is pushed further ahead in the direction of motion but a flash behind the bar shows little or no shift in the opposite direction (Shams-Ahmar et al., ECVP 2022). Recently, it has been proposed that attentional repulsion (Suzuki & Cavanagh, JEP:HPP 1997) may cause the shift (Shams et al., ECVP 2023) as attention leads a moving bar (Szinte et al., J Neurophysiol. 2014). Here, we manipulate the expected motion direction to see if an attentional manipulation influences the motion-induced position shift. A shape moved downward to the center of the screen, where the dot flashed on top of it and the shape then moved away either rightward or leftward. In three sessions run on separate days, the probability of the second motion direction was left and right equally often, left more likely, or right more likely. We found that the illusory shift of the flash was consistently in the direction of the motion that followed it, and that the likelihood of the direction significantly affected the induced position shift (Friedman test: p=0.001; medians: unlikely=0.07 dva; 'equally likely'=0.11 dva; likely=0.23 dva). Further, within each session, despite participants being aware of the more likely direction, the bias towards the more likely direction increased systematically across trials consistent with the notion of a gradual drift of attentional resources towards the expected direction of motion. No trial-to-trial effects were found, ruling out any contribution of serial dependence. We attribute the effects of direction frequency to a greater allocation of attention to the more frequent direction, increasing its effect on the shift. Attentional repulsion remains a viable explanation for the motion-induced shifts seen in static flashed tests.

43.446 ATTENTIONAL SWITCHING IN INFANTS AS A FUNCTION OF BIRTH EXPERIENCE: AN EYE MOVEMENT STUDY

Shir Kay[1] (shirkay4@yorku.ca), Scott A. Adler[1,2]; [1]York University, [2]Centre for Vision Research

Studies have shown that visuospatial attentional performance in a Posner-like attentional cueing task by 3-month-old and 6-month-old infants, and adults, is influenced by birth experience; that is, whether they were delivered via C-section or vaginally. In particular, these findings suggested that C-section birth produces aberrant bottom-up attentional mechanisms. What is not known is whether this birth experience effect produces deficits in selective attention that are specific to the mechanisms engaged by peripheral cueing or translates to other attentional phenomena. To this end, in this study, 6-month-old infants born either vaginally or via C-section participated in an expectation cueing task that used central cueing and their anticipatory and reactive saccades were measured. In the first condition, infants saw centrally presented cues, one of the cues predicted a target to the left and the other predicted a target to the right. Halfway through, the cue that predicted a target to the left side, for example, switched and now predicted a target on the right side and vice versa. In the second condition, which served as a control, the cues did not predict target location during the first half of trials. In the second half, the cues now reliably predicted the target location. Results have indicated that 6-month-old infants, irrespective of birth experience, form expectations for the cue-target location relation and exhibit anticipatory saccades when that relation is predictable. When the cue-target location relation is switched, infants' anticipatory saccades decreased irrespective of birth experience. Latencies of reactive saccade saccades, however, increased post-switch for infants born by C-section but not for those born vaginally. These findings suggest that C-section birth impacts selective attention mechanisms irrespective of the task. More importantly, that the effect of C-section birth is reflected in reactive saccades and not anticipatory saccades is consistent an aberrant bottom-up attentional mechanism.

43.447 EXAMINING THE INFLUENCE OF COGNITIVE PROCESSING ORIENTATION ON TRACKING PERFORMANCE USING A MODIFIED MULTIPLE OBJECT TRACKING TASK

Mengzhu Fu[1], Michael D Dodd[2]; [1]Shippensburg University of Pennsylvania, [2]University of Nebraska - Lincoln

Extensive research on multiple-object tracking (MOT) has shown that task performance may be influenced by both object-related and task-irrelevant factors (Meyerhoff, Papenmeier, & Huff, 2017). While some research suggests that certain expertise might be associated with more efficient tracking (Allen, McGeorge, Pearson, & Milne, 2004; Green & Bavelier, 2006), the question of how to consistently improve performance across individuals remains uncertain. The present study explores whether maintaining a wider attentional scope could enhance tracking abilities using a novel dual task paradigm. Participants tracked a number of moving targets (ranging 2-4) while performing a simultaneous secondary probe-detection task. While objects are in motion, a number of probes may appear on any of the moving objects (i.e., targets & distractors). In the global task condition, participants were instructed to respond only when probes appeared simultaneously on all tracking targets. This aims to prompt global processing by requiring participants to attend to the overall structure of the display to detect all probes at once. In the local task condition, participants were instructed to respond whenever a probe appeared on any tracking target. This aims to prompt local processing by directing participants to attend to individual objects within the display. It was observed that while tracking accuracy decreased as number of tracking targets increased, adopting a global processing orientation is associated with more efficient tracking performance than local processing orientation, especially with larger tracking sets. These results suggest that maintaining a broader attentional focus might efficiently improve tracking efficiency in dynamic scenes.

43.448 MACAQUE MONKEYS FOLLOW GAZE CUES OF HUMAN AVATARS

Ehsan Aboutorabi[1] (eaboutor@uwo.ca), Borna Mahmoudian[1], Julio Martinez-Trujillo[1]; [1]Robarts Research Institute, Western University, London, ON, Canada

Joint attention is a fundamental ability of humans and other social primates. Gaze direction could be informative of the behavioral relevance of objects in the environment. Paradigms to explore joint attention usually involve measurements of gaze in at least two subjects making it difficult to conduct with non-human primates. Here we show a novel paradigm using avatars presented in a virtual environment while exploring gaze behavior and joystick responses of monkeys in an experimental setup. We trained two rhesus monkeys to respond to a human avatar's attention by moving a joystick towards the gazed-at object. We designed our social cues by applying natural eye and head movements on a highly realistic human avatar (released by Reallusion). Each trial commenced with the avatar gazing at the animal, while four identical objects were presented at the screen's corners. After 500 ms, the avatar randomly shifted gaze towards one object, cueing the animal to move the joystick toward that object to obtain a juice reward. In 10% of the trials, the avatar gaze moved in between two objects. Eye positions were measured using EyeLink (SR Research). The animals followed the avatar's gaze, achieving 90% correct trials. We trained a classifier to identify the animals' choice from their eye positions during the cue period (350 ms), achieving an 80% accuracy. The classifier's accuracy decreased to 53% in catch trials when the avatar directed gaze to intermediate positions between objects. Here, the animals chose one of the two objects closest to the avatar's gaze final position, further indicating the animals followed the avatar gaze cues. Our results demonstrate the use of human avatars in experimental setups to explore joint attention in macaque monkeys. It also demonstrates a degree of cognitive flexibility and extrapolation of human gaze cues in macaques.

43.449 MULTIPLE OBJECT TRACKING AS A MEASURE OF SUSTAINED ATTENTION AND RELATION WITH FLUID REASONING

Taryn Perelmiter[1], Domenico Tullo[2], Jocelyn Faubert[3], Armando Bertone[1]; [1]McGill University, [2]Northeastern University College of Science, [3]Université de Montréal

The Multiple Object Tracking (MOT) paradigm has been a staple psychometric tool in cognitive sciences for over three decades, primarily designed to assess the ability to track multiple items simultaneously and explore the limitations of select subcomponents of attention. The present study seeks to extend the paradigm's application to real-world settings by investigating the paradigm's utility as a measure of sustained attention, while also considering individual differences in higher-order cognition, such as fluid reasoning intelligence. In this study, 61 typically developing adults completed the MOT task at varying demands of sustained attention (i.e., manipulations in trial durations consisting of 5, 8, 11, and 15 seconds). Participants also completed a measure of cognitive functioning via the Wechsler Abbreviated Scale of Intelligence – 2nd edition. The results demonstrated that fluid reasoning intelligence was a significant and robust predictor of MOT task performance across all trial lengths. Here, individuals with higher fluid reasoning scores outperformed those with lower scores, regardless of whether they also had high verbal intelligence, extending previous work examining the relationship between distributed attention and intelligence. Furthermore, performance on 8-second trials emerged as the best significant predictor of fluid reasoning intelligence, suggesting an optimal trial condition to characterize attention resource capacity for sustained attention. These findings contribute to the growing body of evidence affirming the strong link between fluid reasoning intelligence and sustained attention capabilities as measured by the MOT. Moreover, they support the paradigm's effectiveness in differentiating attention resource capacities among individuals. This study also uncovers the clinical significance of these insights, highlighting the optimal conditions for evaluating attentional capabilities in individuals with and without attention deficits. Taken together these results advocate for the MOT's broader application in both research and clinical settings, emphasizing its value in isolating specific subcomponents of attention.

Fonds de la Recherche du Québec - Santé (FRQ-S) Junior 1 Research Scholar Award

43.450 MULTIPLE-OBJECT-TRACKING SUPPORTED BY PARVOCELLULAR, MAGNOCELLULAR, AND KONIOCELLULAR PATHWAYS

Linan Shi[1] (shiii@ibp.ac.cn), Yunjie Song[2], Sheng He[3]; [1]Institute of Biophysics

Our study explores the differential contributions of the parvocellular, magnocellular, and koniocellular pathways to the task of multiple-object-tracking (MOT). Observers viewed moving stimuli designed to isolate these pathways. For the parvocellular pathway, red on green isoluminant stimuli were used; yellow on blue isoluminant stimuli were employed for the koniocellular pathway; and fast flickering discs on a red-green checkerboard background were adopted to isolate the magnocellular pathway. We also included a condition that allowed all visual pathways to participate for comparison. Observers' performance was measured while we manipulated two critical variables: the minimal inter-item distance and the motion speed of the moving objects. These manipulations allowed us to assess how spatial and temporal processing demands impact tracking accuracy across different visual pathways. Results reveal distinct pathway sensitivities in supporting attentional tracking: Koniocellular pathway exhibited marked sensitivity to both speed and spatial proximity. As object speed increased and inter-item distance decreased, a significant drop in tracking accuracy was observed. This suggests a pronounced limitation in the koniocellular pathway's ability to process rapidly moving and closely spaced stimuli. Magnocellular pathway demonstrated a higher sensitivity to spatial proximity. Reduction in inter-item distance substantially impaired tracking accuracy, underscoring this pathway's limitation in spatial in spatial resolution during motion perception. Parvocellular pathway showed a pattern similar to that observed when all visual pathways were engaged. Here, rapid object motion combined with reduced inter-item distance led to only a slight decrease in tracking accuracy, suggesting a more robust processing capability under high temporal and spatial demands. This research contributes to our understanding of how different visual pathways support the complex and attentional demanding task of MOT, with results highlighting the unique processing characteristics and their relations to attention of each pathway.

43.451 NON-IMAGE FORMING VISION AS MEASURED THROUGH IPRGC-MEDIATED PUPIL CONSTRICTION IS NOT MODULATED BY COVERT VISUAL ATTENTION

Ana Vilotijević[1] (a.vilotijevic@rug.nl), Sebastiaan Mathôt[1]; [1]University of Groningen

When exposed to light, the pupil constricts, whereas in darkness, the pupil dilates: this is the pupillary light response (PLR), which, for a long time, had been considered to be a reflex. The PLR is driven by all photoreceptors—rods, cones, and intrinsically photosensitive retinal ganglion cells (ipRGCs)—where rods and cones cause the pupil to immediately constrict in response to light, whereas ipRGCs cause the pupil to remain constricted for as long as light is on. Recent studies have shown that the initial PLR is modulated by covert attention; however, it remains unclear whether the same holds for the sustained PLR that is driven by ipRGCs. In our study, we investigated the effect of covert attention on the sustained PLR. To do so, we leveraged the fact that ipRGCs are predominantly responsive to blue light, causing the most prominent sustained constriction in response to blue light. We found that the pupil constricted more when covertly attending to bright as compared to dim stimuli (with the same color), an effect that emerged rapidly after stimulus onset, thus replicating the effect of covert attention on the initial PLR. However, we did not find any difference in pupil size when covertly attending to blue as compared to red stimuli (with the same luminosity), whereas we did observe this difference when participants directly looked at the same blue or red stimuli. This suggests that the sustained PLR is not modulated by covert attention. This finding implies that non-image forming vision, as measured through ipRGC-mediated pupil constriction, is not modulated by covert visual attention.

This research was supported by the Innovational Research Incentives Scheme VIDI (VI.Vidi.191.045) from the Dutch Research Council (NWO) to Sebastiaan Mathôt.

43.452 PUPILLOMETRIC IMAGING REVEALS THE SPATIOTEMPORAL DYNAMICS OF COVERT ATTENTION

Marnix Naber[1] (marnixnaber@gmail.com), Marinos Savva[2], Stefan Van der Stigchel[3], Samson Chota[4]; [1]Experimental Psychology, Helmholtz Institute, Utrecht University, The Netherlands

The visual system uses attention to process relevant aspects in the environment. Particularly covert attention plays an important role as it allows us to inspect information presented in the visual periphery before or without an eye movement. However, due to its latent and dynamic nature, it has been a challenge to characterize the spatial and temporal properties of covert attentional shifts. To date little is known about how the focus of attention moves across space and time. We developed a novel pupillometric imaging paradigm to directly probe and visualize spatiotemporal shifts of attention in observers that performed a classic Posner's cueing task. The distribution of attentional resources was measured by proxy of the amplitude of pupil orienting responses to salient probes that sampled various positions and timepoints around cue and target onsets. The resulting attention maps confirm that the analogy of attention as a local spotlight holds when stationary. The analysis of its temporal dynamics indicate that the attentional spotlight, when shifting between peripheral locations, gradually fades out at start positions and fades in at end positions across time. When shifting from foveal to peripheral locations, the degree of attention only decreases at its start position (i.e., fixation), resulting in relatively more attention at start and end positions before and after a shift, respectively. As the first two-dimensional imaging effort of covert attention shifts across peripheral locations, this study lays the foundation to characterize attentional properties at an unprecedentedly high spatiotemporal resolution.

43.453 SOMETIMES MORE (OVERLAP) IS BETTER! ACTION PLAN OVERLAP IMPACTS THE INTERFERENCE BETWEEN VISUALLY-GUIDED TOUCH AND MULTIPLE-OBJECT TRACKING (MOT)

Mallory E. Terry[1] (terry@uoguelph.ca), Vanessa Amelio[1], Lana M. Trick[1]; [1]University of Guelph

When two tasks are performed simultaneously their action plans can overlap with one another. Past findings suggest that the overlap can either improve or degrade performance, depending on the relatedness of the required actions (e.g., Fournier et al. 2015). In this study we assessed the impact of overlapping action plans in a multiple-object tracking (MOT) task. Participants tracked 1-4 MOT targets while also touching moving items in MOT that changed colour. To determine the effects of action plan overlap between the MOT and touch task, we manipulated the way that participants reported the identity of the targets at the end of the trial (untimed). In the touch task participants always used the index finger of their dominant hand. To report the targets participants either typed in the letters corresponding to the targets with their non-dominant hand (minimal overlap) or touched MOT targets with the index finger of their dominant hand (maximal overlap). Target report method had no effect on single-task MOT performance. However, when participants had to touch moving items that changed colour during tracking (dual-task), MOT performance was significantly worse when overlap was minimized. It also took participants longer to touch moving items that changed colour - even though target report occurred 7-8 seconds later. Nonetheless, MOT performance was always better and touch latencies lower when the touched items were targets as compared to distractors in MOT; report technique had no effect. This shows a dissociation between the effects of attentional selection in MOT and overlapping action plans.

NSERC

43.454 THE NATURE AND COMPUTATION OF ATTENTIONAL EFFORT: A PEAK/END RULE INTEGRATING OVER MOMENT-BY-MOMENT EFFORT DURING MULTIPLE-OBJECT TRACKING

Mario Belledonne[1] (mario.belledonne@yale.edu), Ilker Yildirim[1], Brian Scholl[1]; [1]Yale University

So much of perception is effortless, but a hallmark property of sustained visual attention is a vivid sense of effort. Nowhere is this more evident than during multiple object tracking (MOT), where keeping track of a group of moving targets amidst identical moving distractors involves a distinct sense of exertion. But where does this sense of effort come from? The answer is not immediately obvious, in part because of the dynamic nature of MOT: each MOT trial is almost an experiment unto itself, with a moment-by-moment ebb and flow of effort, as the proximities of targets and distractors constantly change. Accordingly, we asked a straightforward question (with a surprising answer): how does the feeling of retrospective effort (at the end of a trial) relate to the moment-by-moment experience of effort during a trial? To find out, we augmented MOT in two ways. First, during tracking, subjects reported their moment-by-moment sense of effort using a continuous dial (with the continuously varying pitch of a tone providing feedback that did not interfere with tracking). Second, immediately after each trial, subjects used a slider to report how effortful tracking was overall. Retrospective effort was not simply the average of moment-by-moment effort, but rather was best explained by certain brief moments — especially the *peak* effort, and the effort near the *end* of each tracking interval. These moments provided maximal predictive power: adding the other moment-by-moment effort ratings did not improve prediction of retrospective effort, and this was not true for any other temporal windows. This peak/end pattern is characteristic of retrospective reports of many other properties at longer time-scales — from the joy of a vacation, to the pain of a surgery. These results thus demonstrate a striking convergence between a hallmark effect of cognition and the moment-by-moment dynamics of visual phenomenology.

43.455 THE RESET EFFECT OF ATTENTION DEPENDS ON THE PHASE OF ONGOING ATTENTION OSCILLATION

DIKANG YIN[1] (2073279495@qq.com); [1]Institute of Biophysics, Chinese Academy of Sciences

A peripheral cue is believed to reset the phase of ongoing attention oscillation to the cued location. In the current study, we tested whether the reset effect of attention depends on the phase of ongoing attention oscillation. We devised a two-cue-1-probe behavioral paradigm to measure this effect. The first cue was used to trigger rhythmic attention sampling between two locations, the second cue was presented at two opposing phases of the rhythm, when attention was sampling either the same (in-phase) or opposing (out-of-phase) side of the first cue. The probe was presented at variable SOAs at 25-ms interval to measure the time course of reaction time. The two phases for presenting the second cue were determined by a pilot experiment in which only the first cue and the probe was presented. The results from a sample of 20 university students revealed that the amplitude of oscillation was enhanced when the second cue was presented in-phase with the attention sampling. Interestingly, when the second cue was presented out-of-phase or opposite to the currently sampled location, it reduced the amplitude of ongoing attention oscillation rather than reset the sampling of attention to the cued location. Taken together, these results suggest that the effect of reset event on attention sampling depends on the phase of ongoing attention oscillation.

43.456 THE SPEED OF ATTENTIONAL SHIFT IS SIMILAR FOR EXOGENOUS AND ENDOGENOUS CUES WITHIN ANTI-CUE TASKS

Yara Mohiar[1] (yara.mohiar@umontreal.ca), Rémy Allard[1]; [1]Université de Montréal

When navigating within complex environments such as a crowded street, our focus of attention often rapidly shifts between objects. Given that attention generally shifts faster when it is triggered involuntarily (e.g., with an exogenous cue near the target) than voluntarily (e.g., with an endogenous cue pointing to the relevant target), we could expect that ignoring an exogenous cue would be more difficult than ignoring an endogenous cue. The current study compared the speed at which attention can be shifted when using exogenous and endogenous cues when having to shift their attention to a location opposite to the cued location (anti-cue). On each trial, eight gabors were simultaneously presented for 600 msec uniformly distributed along an annulus centered on the fixation point and randomly rotated clockwise or counter-clockwise by 30 degrees after 300 msec. Participants were asked to report the motion direction of the cued gabor (pro-cue) or of a target on the opposite location of the cued gabor (anti-cue). The cue was either exogenous (a circle around one of the gabors) or endogenous (an arrow at fixation pointing to a gabor). The time required to shift attention to the target location was measured using a staircase procedure controlling the duration between the presentation of the cue and the occurrence of the motion (i.e., cue onset asynchrony, COA). In the pro-cue conditions, the COA was significantly lower with an exogenous cue than an endogenous cue, as expected. In the anti-cue conditions, the COA was globally significantly higher than pro-cue conditions, but no significant difference between the type of cue (exogenous vs endogenous) was observed. These results suggest that it is similarly demanding to attend to a target located at the opposite side of the cued location when the cue is exogenous or endogenous.

This research was supported by a grant from the Centre de recherche interdisciplinaire en réadaptation du Montréal métropolitain (CRIR) to RA.

MONDAY, MAY 20, 8:30 AM – 12:30 PM, PAVILION

Attention: Temporal selection

43.457 TEMPORAL FEATURE BINDING: ATTENTION MECHANISMS AND INDIVIDUAL DIFFERENCES

Alon Zivony[1] (a.zivony@sheffield.ac.uk), Martin Eimer[2]; [1]University of Sheffield, [2]Birkbeck College, University of London

Previous studies have shown that illusory conjunction can emerge for both spatially and temporally proximal objects. However, the mechanisms involved in binding in the temporal domain are not yet fully understood. In the current study, we investigated the role of attentional processes in correct and incorrect temporal binding, and specifically how feature binding is affected by the speed of attentional engagement. In two experiments (N=40), participants searched for a target in a rapid serial visual presentation stream and reported its colour and alphanumeric identity. Temporal binding errors were frequent. Critically, when participants reported the identity of a distractor instead of a target, they were also more likely to report the colour of this distractor. This finding suggest that attentional engagement facilitates the binding of temporally co-occurring features, and that variability in the speed of attentional engagement plays an important role in determining what we see. Moreover, in both experiments, strong correlations were observed between identity and colour reports across participants: individuals who were more likely to erroneously report the identities of post-target distractors were also more likely to report the colours of post-target distractors. This finding points to substantial and reliable individual differences in temporal selectivity, opening the door to future research on this issue.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 896192 to A. Zivony and from ESRC Grant No. ES/V002708/1 to M. Eimer.

43.458 KEEPING AN EYE ON ON-LINE LEARNING: ATTENTION MATTERS

Lester Loschky[1] (loschky@ksu.edu), Prasanth Chandran[1], Brian Howatt[1], Yifeng Huang[2], Jeremy Munsell[3], Brayden Wallace[1], Lindsey Wilson[1], Ally Dinkel[1], Sidney D'Mello, Minh Hoai[2], Sanjay Rebello[3]; [1]Kansas State University, [2]SUNY Stonybrook, [3]Purdue University

Computer-Assisted Instruction (CAI) is here to stay. However, given it's unsupervised nature, a key problem for CAI is estimating students' attention. Nevertheless, CAI is an ideal context for investigating these issues, because it allows comprehensive measurements of learners' behavior during naturalistic learning activities. Here, we report on the initial results of a study of 100 students' attentional states while studying a multimodal Physics module and their learning outcomes. Our study uses a 2x2 matrix to characterize online learners' attentional and cognitive states (D'Mello, 2016). First, when learners are looking at the screen, we distinguish between when they are thinking about the content versus when they are not (i.e., on-task vs. mind-wandering). Second, when learners are looking elsewhere (i.e., off-screen), we similarly distinguish between when they are thinking about the content versus not. Method: To measure students' attentional states, we combined information from a webcam, eye tracker, egocentric camera (showing what students looked at), and probes to measure mind-wandering. After learners had finished the module, we synchronized the multi-modal data and conducted a retrospective recall to ask students about their attentional and cognitive states whenever they had gazed away from the materials. To determine learning effects of the above attentional states, we used a 26-item pre-test, post-test, and 1-week retention test. Initial Results: Learners with less prior knowledge spent more time looking on-screen while thinking about the content, which was positively correlated with learning. However, learners with less prior knowledge also spent more time looking on-screen, but mind-wandering, which was negatively correlated with learning. Learners with greater prior knowledge spent more time looking off-screen, but thinking about the content, which was uncorrelated with learning. However, learners with greater prior knowledge also spent more time looking off-screen, but not thinking about the content, which was negatively correlated with learning.

Funding: National Science Foundation Grant 2100071 to LL, MH, SR

43.459 RHYTHMICITY IN PROACTIVE AND RETROACTIVE ATTENTION

Phillip Cheng[1] (pxcheng@princeton.edu), Sabine Kastner[1,2]; [1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, USA, [2]Department of Psychology, Princeton University, Princeton, NJ 08540, USA

Attention plays a key role in sampling relevant perceptual representations of the environment. The sampling seems rhythmic, manifested in fluctuations of behavioral performance. In previous research however, this rhythmic sampling has been limited to proactive attention (attending to upcoming events). The temporal dynamics of retroactive attention (attending to past events) remains unknown. Here, we investigate whether attention enhances perceptual processing in a rhythmic manner, both proactively and retroactively. We used an exogenous attention paradigm, where a cue drew attention to one of two possible target locations, either before or after the target onset. The cue was uninformative, resulting in half congruent trials and half incongruent trials, to minimize the attentional effect on post-perceptual decisions. The contrast of the target, a Gabor patch, was individually thresholded so that it was visible in approximately 50% of trials. Participants discriminated orientation and rated their awareness of the target. We found that the attentional enhancement of perception waxed and waned at a theta frequency depending on the cue-target interval, consistent with previous findings. Importantly, rhythmic behavior was not only observed in trials where the cue preceded the target but also when the cue followed the target, suggesting rhythmic properties of retroactive attention. Further, the behavioral rhythms at the cued and uncued locations were in antiphase. These space-based temporal dynamics were present regardless of target visibility. Note that all the above findings were significant at an individual level. Together, these findings suggest that attention is intrinsically rhythmic regardless of pro- or retroactive allocation.

43.460 TEMPORAL ATTENTION AND EXPECTATION INTERACT REGARDLESS OF EXPECTATION'S TRIAL SEQUENCE

Shiyang Ren[1] (syhren@nyu.edu), Aysun Duyar[1], Marisa Carrasco[1]; [1]New York University

[Background] Endogenous temporal attention is the voluntary deployment of attention to a specific moment, whereas temporal expectation is based on the temporal probability of event onsets. Endogenous temporal attention interacts with temporal expectation. Benefits of temporal attention are highest when stimuli appear early and decrease as onset is delayed. It is unknown whether attentional benefits depend on the timing of the prior trials–i.e. whether stimulus onsets affect observers' expectations in the subsequent trials and result in differential allocation of attention. [Goal] To investigate potential sequential effects of stimulus onset on the interaction between temporal expectation and attention. [Methods] Two oriented Gabors were presented sequentially at the fovea. Observers performed a 2AFC orientation discrimination task. Participants were instructed to attend to either the first (T1) or second (T2) stimulus (valid), or to both stimuli (neutral). T1 could appear within a temporal window of 1200-1600 ms (expected moment: 1400 ms), and T2 appeared 250 ms after T1 (expected moment: 1650 ms). A response cue indicated the target observers had to discriminate. [Results] There was a significant interaction between temporal attention and expectation. Temporal attention improved performance, and the improvements decreased as the stimulus onset was delayed. The timing of the preceding trial –early, expected, or late onset– did not affect the magnitude of the interaction between expectation and attention in consecutive trials. [Conclusions] Temporal attention benefits decreased as a function of stimulus onset delay. Once observers attended to an early moment, they could not reallocate attention to a later moment, regardless of the timing of the stimulus onset in the preceding trial. Stimulus timing within a trial did not affect

temporal expectations in the subsequent trial. Thus, the interaction between temporal attention and expectation was consistent regardless of the trials' sequential order, suggesting that observers optimize performance in each trial.

NIH R01-EY019693 to M.C.

43.461 EVIDENCE OF RHYTHMIC ENVIRONMENTAL SAMPLING IN A CUED TEMPORAL ATTENTION PARADIGM

Travis Powell[1] (tgppowell@ucdavis.edu), Megan Wong[1], José Samano Catalan[1], Yanelli Guzmán[3], Mingzhou Ding[2], George R. Mangun[1]; [1]University of California, Davis, [2]University of Florida, [3]California State University, Fullerton

An accumulated body of research supports a rhythmic model of environmental attentional sampling, oscillating in the theta-band (3—8 Hz). This phenomenon has largely been investigated in visuospatial and object-oriented attention domains. Previous studies have focused on the relationship between exploratory saccadic activity and spatial attention sampling, predicated on a common theta rhythm. It has been proposed that perceptual sensitivity and likelihood to saccade are in antiphase to each other, mediated by an overarching control network fluctuating at a theta rate. This mechanism is hypothesized to resolve the conflict between sensory and motor processes, allowing for shifts of attention in between periods of heightened sensitivity, thereby facilitating cognitive flexibility. Behavioral evidence of this attentional rhythm in domains beyond visuospatial attention would support a more generalized model, providing even stronger evidence of periodic capture of the environment. We investigate rhythmicity in attention sampling by using a modified Posner cueing paradigm employing high-density sampling across a range of cue-target intervals. We observed periodic fluctuations in mean reaction time (RT) that oscillated in the 3—8 Hz range. The oscillation was attenuated in the attended condition, replicating results found in previous rhythmic sampling research, but extending the evidence to temporal attention.

43.462 ANTICIPATORY ORIENTING OF COVERT ATTENTION WITH DYNAMIC GAZE CUEING

Srijita Karmakar[1] (srijita@ucsb.edu), Miguel P. Eckstein[1]; [1]University of California, Santa Barbara

Introduction: Cueing by eyes and head direction strongly orients an observer's overt and covert attention (Bayliss and Tipper, 2007), even in an anticipatory manner (Joyce et al., 2016, Han and Eckstein, 2023). However, the temporal dynamics by which gaze cueing influences the processing of peripheral information for perceptual decisions are not well understood. To investigate how observers use a foveally-viewed dynamic gazer to orient attention to peripheral information, we conducted a temporal reverse correlation study. Methods: Five observers covertly detected the presence of a bright spatial Gaussian target (SD = 0.93°) and localized it among dimmer distractors in one of two locations to the left and right (10.5° eccentricity) of a central gazer. Target and distractor luminances were perturbed with independent Gaussian noise every 25 ms. In 20%, cue-neutral, trials the central gazer's head did not cue either location, in 64%, normal-cue, trials the head turned for 375 ms to cue the left or right location. In the remaining 16%, reverse-cue, trials the head turned (for 225 ms) toward one location, then reversed its direction (for 400 ms) to cue the opposite location. The cue validity was 80%. Results: An inverse cueing effect was found for three observers in the reverse-cue condition (hit valid = 0.67; hit invalid = 0.73). For those three observers, localization in false-alarm trials was more frequent at cued than at uncued locations in the normal-cue condition (p=0.02) but opposite in the reverse-cue condition. Temporal reverse correlation of noise values for false-alarm trials revealed a time window for peripheral information integration that peaked 100-300 ms after the gazer's head movement initiation, which may explain the observers' inverse cue effect in the reverse-cue condition. Conclusion: Together, the findings suggest that observers anticipate the gazer's head direction to orient covert attention, which can sometimes result in task-suboptimal behavior.

Monday, May 20, 8:30 am – 12:30 pm, Pavilion

Attention: Divided, resource competition

43.463 EYE TRACKING IMPACTS OF AN IRRELEVANT SELF-VIEW SINGLETON

Gabriela Flores-Cruz[1] (ga692434@ucf.edu), Nelson Roque[2]; [1]University of Central Florida, [2]Pennsylvania State University

The following study was conducted to determine whether an irrelevant singleton (i.e., viewing oneself) has an impact on (1) on behavioral performance, i.e., target response accuracy and speed and (2) on eye tracking metrics (total dwell time within the self-view singleton area of interest) on a centered attentional blink (AB) task, which included 18 trials per lag (1, 3, 5, & 8). This experiment employed a within-subjects design with self-view (with two levels: on and off) displayed in a counterbalanced fashion (total of 144 trials). The self-view singleton included a live video of oneself when it was on and an avatar (like Zoom) when it was off. Results revealed a classic AB pattern, indicating that the manipulation of the task worked. In terms of self-view impacts on behavioral performance, we found that self-view on or off did not significantly influence target response accuracy or response times. However, eye tracking data revealed that participants had a higher total dwell time (on the self-view singleton when it was on compared to when it was off. Higher total dwell time when self-view was on did not lead to performance interference. Post-hoc analyses will be conducted on the percentage of time participants spent on the self-view singleton and on the AB task at a trial level. Future research should explore adjustments to task difficulty and self-view placement, incorporate more ecologically valid tasks (e.g., Zoom lectures), and include singleton of someone else's face. This experiment suggested that self-view can attract attention, even when it is unrelated to the primary task, but it did not affect performance with an AB task. However, it is possible that its impacts may differ in other tasks. Overall, the experiment shed light on how often people may look at themselves in remote meetings (classroom or work) with their cameras on.

43.464 AN IN-DEPTH EXAMINATION OF INTERRUPTIONS' EFFECT ON EMAIL CLASSIFICATION BEHAVIOR

Elisabeth J.D. Slifkin[1] (elisabeth.slifkin@ucf.edu), Mark B. Neider[1]; [1]University of Central Florida

Identifying and avoiding phishing emails is a critical task email users engage in, and performance directly impacts the degree of risk to which users are exposed. Though the impact of email, task, and participant characteristics on email classification performance have been studied extensively, the effect of environmental factors, such as interruptions, have been understudied. We recently demonstrated that limited task interruptions improved email classification ability, but increased time on task (Slifkin & Neider, 2023). The present studies build upon this work by evaluating the effect of interruptions on email classification and analyzing eye movements to provide an in-depth characterization of behavior. In three experiments, we assessed the effects of interruption length (Experiment 1), interruption difficulty (Experiment 2), and task-interruption similarity (Experiment 3) on email classification performance. Participants evaluated the legitimacy of 100 real emails (50% phishing), indicating their response by button press. 20% of trials included an interrupting math task 3s after trial onset. Within each experiment, participants experienced both interruption types, and uninterrupted trials. Preliminary results show across all three experiments, email classification accuracy was unaffected by interruption, regardless of interruption type (all ps>.087), but response time reliably increased (all ps<.003). To further characterize interruptions' impact, we analyzed the number of fixations post-interruption to locations fixated pre-interruption. Higher refixation rates indicate a disruptive effect of interruptions on previously encoded email information; lower refixation rates indicate better preservation of previously encoded information. Results were mixed across experiments; Experiments 1&3 showed an increase in refixations (all ps<.039), while Experiment 2 did not (p=.075). Overall, participants maintained accuracy when interrupted, but took longer to classify emails. In Experiments 1&3 interruptions appear to have disrupted representation of email information accrued prior to interruption; Experiment 2 revealed a better-preserved representation. Combined, these findings suggest that certain interruption types may disproportionately impair email classification performance.

43.465 EVIDENCE OF PARALLEL PROCESSING DURING DIVIDED ATTENTION TO FACES

Samantha C. Lee[1] (samanthalee@nevada.unr.edu), Tyler A. Touchstone[1], Lars Strother[1]; [1]University of Nevada, Reno

Visual processing of single faces is associated with reduced attentional demands relative to non-face objects, but evidence that this applies to multiple faces is limited. Although face ensemble studies show that multi-face information can be extracted in parallel as a summary statistic, whether we can divide attention across multiple faces to support single-face tasks is not known. We used a divided attention paradigm in which participants made three different kinds of face judgments: (1) gender (male/female), (2) orientation (upright/inverted), and (3) color (red/gray). Judgments were made for either one (single-task) or two (dual-task) faces in a simultaneously-viewed face pair. Our paradigm enabled us to measure dual-task deficits for each type of judgment and to assess the magnitude of deficits with respect to three benchmark models. The unlimited capacity parallel model predicts little if any deficit. The all-or-none serial model predicts a deficit magnitude consistent with an ability to process only one face at a time. Finally, the fixed-capacity model predicts an intermediate result. In addition to measuring dual-task deficits, we also examined the relationship between the two responses in the dual-task condition. Specifically, we used conditional accuracy to test for evidence of response trade-offs, which we again assessed with respect to the three benchmark models. Consistent with previously published results, the unlimited capacity parallel model was the best predictor of the color judgment results. Dual-task deficits for both the gender and orientation judgments were surprisingly small and consistent with the predictions of a fixed-capacity model. The fixed-capacity model also best predicted the trade-off results for gender judgments, but the unlimited-capacity model best predicted the orientation judgment results. Our overall findings show no evidence of all-or-none serial processing of face pairs, unlike results reported in similar divided attention studies of semantic categorization for non-face objects.

P30 GM145646

43.466 SINGLE OBJECT ATTENTION NARROWS THE PERCEPTUAL TEMPLATE, IMPROVES EXTERNAL NOISE EXCLUSION AND ENHANCES THE STIMULUS RELATIVE TO DUAL OBJECT ATTENTION

Barbara Dosher[1] (bdosher@uci.edu), Songmei Han[1], Zhong-Lin Lu[2]; [1]University of California Irvine, [2]NYU-Shanghai & New York University

Object attention refers to preferred processing of multiple aspects of one object rather than across objects. We examine the mechanisms of object attention by manipulating the contrast of external noise added to the stimulus. We applied the perceptual template model (PTM) to evaluate external noise exclusion and stimulus enhancement (internal additive noise reduction). We compared subject's performance in reporting the orientation and phase of one peripheral Gabor with their performance in reporting the orientation of one Gabor and the phase of another (single versus dual object report). Dual object report deficits are the classic demonstration of object attention. Two Gabors were presented in each trial (±7.3 deg eccentricity on the horizontal meridian; vertical +/-10 deg orientation; sine or cosine phase; Gabor window sigma = 0.45 deg). Single and dual object report were blocked, and the location of first (or sole) object report was indicated by a precue. External noise contrasts were [0, 0.02, 0.04, 0.08, 0.17, 0.33], and each external noise was tested at 7 stimulus contrasts, yielding 24 psychometric functions (6 external noises x 2 judgments x 2 attention conditions) in two-alternative judgments (n= 5 observers). The data were interpolated at 65%, 75%, and 85% correct to generate threshold versus external noise contrast (TvC) curves, which were fitted (average $r2=.9812$) with an extended PTM (ePTM) to account for discrimination between overlapping (non-orthogonal) target templates. The single-object report conditions showed mixtures of external noise filtering and stimulus enhancement (reduced external noise and internal additive noise with Af= 0.76 and Aa= 0.62 for orientation judgments; Af= 0.87 and Aa=0.88 for phase judgments) and more narrowly tuned perceptual templates relative to dual-object report. Distributing attention across two objects for different judgments (orientation and phase) is mediated by less precise perceptual templates.

Supported by NEI (# EY–17491).

Monday, May 20, 8:30 am – 12:30 pm, Pavilion

Plasticity and Learning: Disorders, atypical vision

43.467 VISUAL MODE SWITCHING FOR GLASSES: OBSERVERS CAN LEARN TO IMMEDIATELY CORRECT SPATIAL DISTORTIONS PRODUCED BY PRESCRIPTION LENSES

Sean Liu[1] (liu02542@umn.edu), Stephen Engel[1]; [1]University of Minnesota

Wearing prescription glasses often produces geometric distortions while improving focus. For instance, cylindrical lenses, prescribed for astigmatism, create skew by scaling the image along an oriented axis, causing rectangles to appear like slightly tilted parallelograms. Most people adapt to these distortions, and report the world returning to normal appearance after a few days. They additionally note normal appearance immediately when they put on the glasses back on, and when they take them off. However, there have been few laboratory tests of learned immediate adjustment to compensate for geometric distortions; here we examined whether observers can learn to rapidly switch to "glasses mode”. Five observers wore +1.25D, 45-deg axis, cylindrical lenses binocularly, which skewed and blurred the images reaching the eyes. They wore the lenses in glasses for four hours during each of five consecutive days. Perception of skew was measured using a cancellation method. Observers adjusted the physical skew (along the 45 degree axis) of a parallelogram presented against a grayscale naturalistic background. Observers set the physical skew so that it canceled the effects of the glasses, and the parallelogram appeared to contain 90 degree angles. When first donning the glasses on day 1, observers settings were around 2% on average (scaled by 1.02 along the -45 degree axis); this scaling counteracted the skew produced by the glasses, producing a perceived rectangle. By day 5, however, this initial setting reduced to an average of 1%, indicating that the world appeared much less skewed when putting the glasses on. This trend across days was visible in all individual observers and reliable at the group level ($t(4) = -4.9859$, $p < 0.01$). These results suggest that observers can learn to rapidly adjust their perception to counteract spectacles' distortions. Failure of such learning may underlie dissatisfaction with glasses, especially for more complex prescriptions.

43.468 ACTION VIDEO GAMES TRAINING IN CHILDREN WITH DEVELOPMENTAL DYSLEXIA: A META-ANALYSIS

Simone Gori[1] (simone.gori@unibg.it), Giovanna Puccio[2], Giulia Gazzi[2], Sara Bertoni[1], Sandro Franceschini[2], Martina Mancarella[3], Lucia Ronconi[2], Andrea Facoetti[2]; [1]University of Bergamo, [2]University of Padua, [3]Unit, KU Leuven

Longitudinal studies and meta-analyses have shown a causal link between attentional dysfunctions and developmental dyslexia (DD). We carried out a meta-analysis to test the effectiveness of action video games (AVGs) training on visual attention in children with DD. PubMed, Cochrane, Science Report, EBSCO Database, Scopus, ProQuest Dissertation and Theses, and IEEE Explore were consulted. Only quantitative studies with measures of pre- and post-treatment reading skills, written in English, and with an active control group were considered. The risk of bias was evaluated according to RoB2 and ROBINS-I assessment tools. Out of 2073 records, nine experiments using AVGs in 238 children (aged 5–15) with DD were selected. The Hedge’s g results indicate that AVGs training affects visual attention as well as reading-related functions. Studies with a larger sample including follow-up assessments and neurobiological studies are needed to verify AVGs long-lasting effects on DD.

43.469 EXAMINING THE ROLE OF THE ARCUATE FASCICULUS ON READING DEVELOPMENT BY STUDYING REPETITIVE HEAD IMPACT

Nii-Ayi Aryeetey[1], Kelly Hiersche[1], Jeff Pan[1], James Onate[1], Ginger Yang[2], Sean Rose[2], Jaclyn Caccese[1], Zeynep Saygin[1]; [1]Ohio State University, [2]Nationwide Children’s Hospital

A myriad of changes occur as a child learns to read including in gray matter (e.g. word-selectivity in the visual word form area (VWFA)) and white matter (e.g. fractional anisotropy (FA) of the left arcuate fasciculus (AF)). Further, pre-readers at risk for developing dyslexia show lower FA in left AF, revealing potential white matter mechanisms that drive reading development. While we cannot know whether arcuate development is causally responsible for much of early reading ability without disrupting the system, here we explore this question by studying head impacts in otherwise typically-developing 8-12 year-olds. Youth football leagues usually start tackle football around age 8, when most children acquire literacy. We scanned eleven children before their initial season of tackle football and followed them longitudinally after the season, and compared their neurodevelopment to a matched control cohort. Preliminary findings show that at preseason, there are no significant between-group reading differences in VWFA selectivity, and FA in AF; however, we observed FA differences in the left AF postseason. Children who played their first season of tackle football showed no developmental changes in the left AF while control children showed increasing FA longitudinally (4-5 months). These white matter differences might precede observable deficits in reading, as we found no significant between-group differences in reading scores postseason (albeit trending, with modest reading score improvements in children in the control but not football group). Significant reading ability changes in typical development are expected at follow-up. Overall, these results suggest that exposure to repetitive head impacts may lead to atypical white matter development, and that the growth of the left AF may be especially important in driving reading development. Ongoing longitudinal investigations will explore whether children catch up in AF growth with increasing time post-season, and dose-response relationship of head impacts on reading and other cognitive outcomes.

43.470 IN BLINDNESS, PPA-LIKE RESPONSES TO SCENE SOUNDS ARE DRIVEN BY STRONGER DEACTIVATION FOR HUMAN VOCALIZATIONS, NOT ACTIVATION FOR SCENES

Elizabeth Saccone[1] (esaccon2@jhu.edu), Akshi LNU[1], Marina Bedny[1]; [1]Johns Hopkins University

A prominent view is that the parahippocampal place area (PPA) preserves a role in scene perception, regardless of visual experience. This claim is based on findings that in congenitally blind people the medial vOTC (i.e., PPA) responds more to scene sounds (e.g., waves on a beach) than to human vocalizations (e.g., laughter, crying) - so called 'face sounds' (e.g., van den Hurk et al., 2017, PNAS). We tested whether in blindness the PPA: 1) responds to place sounds or deactivates to vocalizations, 2) whether the same vertices respond to scene sounds and tactile scenes and 3) whether there is specialization across left/right and anterior/posterior PPA. Fifteen English-speaking, congenitally blind adults performed a 1-back auditory task during fMRI. Conditions were scene sounds (e.g., waves crashing on a beach, city traffic; van den Hurk et al., 2017), non-verbal vocal sounds (e.g., laughter), language (English sentences), and foreign speech (Russian, Korean, Mandarin). The same participants also performed a tactile 1-back task, touching 3D-printed models of scenes and faces (Ratan Murty et al., 2020, PNAS). We replicate the previous finding of greater response to scene sounds than non-verbal vocalizations in medial vOTC. However, this effect is driven by greater deactivation to vocalizations, while responses to scene sounds are below (left PPA) or not different (right PPA) from rest. There is no evidence for anterior/posterior specialization. Responses to tactile scenes in the PPA were either below (left) or barely above rest (right) and no different than in lateral vOTC, where the fusiform face area is typically located. Rather, medial vOTC is distinguished from lateral vOTC by its deactivation to vocalizations. Results are not consistent with a role of medial vOTC for perception of places in blindness. We find no evidence that this region preserves its function for scene perception in the absence of visual experience.

43.471 DISTINCT TACTILE AND VISUAL BRAIN RESPONSES TO ALPHABETIC LETTERS CONVERGE TO SHARED REPRESENTATIONS IN BLIND AND SIGHTED READERS

Santani Teng[1,3], Radoslaw Cichy[2], Dimitrios Pantazis[3], Aude Oliva[3]; [1]Smith-Kettlewell Eye Research Institute, [2]Free University of Berlin, [3]CSAIL, Massachusetts Institute of Technology

Functional networks are reorganized in blind individuals to engage typically visual cortical sensory areas in nonvisual tasks. Braille has served as a major model to investigate such crossmodal plasticity, but the underlying brain representations and spatiotemporal dynamics of these computations remain poorly understood. Here, we presented individual visual and tactile (braille) alphabetic letters to sighted and early-blind participants, respectively, while recording brain activity with magnetoencephalography (MEG). Both groups of participants read letters passively, responding via button press to occasional vigilance targets. For each group, we used multivariate pattern analysis to compare brain responses to alphabetical letters across different anatomical locations and time points in the trial epoch. Next, we modeled low-level stimulus representations in tactile and visual brain responses, as well as higher-level distributional statistics of letter occurrences in a text corpus. Using these representational similarity models and a shared-variance approach, we examined the spatiotemporal dynamics of the representational cascade, identifying a convergence from distinct low-level to shared higher-level representations of individual letter stimuli between early sensory and left fusiform regions. Taken together, the results reveal spatiotemporally dissociable representations of individual letter processing, common and distinct computations in blind and sighted individuals, and a possible role for early "visual" cortex in the reorganized functional brain networks of blind readers.

43.472 TOWARD COMPARING SCOTOMAS: USING MICROPERIMETRY PAIRED WITH CORTICAL MAGNIFICATION FACTOR TO QUANTIFY RETINAL FUNCTIONAL HEALTH IN PATIENTS WITH CENTRAL VISION LOSS

Elam Cutts[1] (ecutts99@uab.edu), Marcello Maniglia[2], Matthew Defenderfer[1], Pinar Demirayak[1], Dawn DeCarlo[1], Kristina Visscher[1]; [1]University of Alabama at Birmingham, [2]University of California, Riverside

In macular degeneration, the primary cause of vision loss among older adults, photoreceptor death in the retina results in diverse patterns of vision impairment. Complex visual tasks such as reading or navigation require both basic visual sensation (here, 'low level vision') as well as neural processes beyond basic sensation (here, 'high-level vision'). Patients with similar retinal damage can differ widely in their performance on complex visual tasks, suggesting that compensation for this impairment varies between patients. Quantification of this compensation is a necessary step to understanding the neural mechanisms underlying compensatory visual strategies. Traditional tests like visual acuity and contrast sensitivity gauge only limited aspects of vision and may not capture the broader visual field, leading to discrepancies between test outcomes and real-world function. To bridge this gap, we developed a method using outcomes from the Macular Integrity Assessment (MAIA), a microperimetry method that evaluates sensitivity across the retina. Quantitatively comparing MAIA results across patients has been a challenge, especially given that lesions in central vision lead to worse impairment than peripheral vision. Here we introduce a method to quantitatively account for that difference, using the concept of the "Cortical Magnification Factor" (CMF). Different parts of visual cortex correspond to distinct regions of vision (retinotopic maps), and the CMF describes how much more cortex is devoted to each portion of the visual field. By weighting MAIA scores with CMF, we derived a measure called retinal functional health (RFH). RFH reliably reflects the clinical impression of the severity of a scotoma. RFH was significantly correlated to contrast sensitivity, as well as acuity. Further, models incorporating RFH were better predictors of high-level visual processing (aggregate performance on a range of complex visual tasks). These results validate our measure of RFH to compare scotoma severity across participants.

43.473 STABLE INDIVIDUAL PLASTICITY PATTERNS IN BLINDNESS: A LONGITUDINAL STUDY

Lenia Amaral[1], Peyton Thomas[1], Amir Amedi[2,3], Ella Striem-Amit[1]; [1]Department of Neuroscience, Georgetown University Medical Center, Washington, DC, USA, [2]Ivcher School of Psychology, The Institute for Brain, Mind and Technology, Reichman University, Herzliya, Israel, [3]The Ruth & Meir Rosenthal Brain Imaging Center, Reichman University, Herzliya, Israel

Rehabilitation for individuals with blindness remains a challenge, primarily due to the considerable variability in functional outcomes, despite supposedly successful restoration of the peripheral light sensitivity. Brain reorganization may be a limiting factor in allowing a person who has been blind for a substantial duration to regain sight. To identify potential candidates who may benefit most from rehabilitation, it is crucial to search for neuromarkers that can predict success. However, the visual cortex in blindness was suggested to partake in multiple networks over time, flexibly shifting its function based on task demands. This would make it impossible to reliably use its connectivity or responses as neuromarkers. Here we specifically tested whether this is the case, by examining the stability of individual connectivity patterns from the visual cortex in a group of eight congenitally blind individuals over a period of two years. Interestingly, our findings reveal that functional connectivity from the primary visual cortex in this small sample of repeatedly sampled congenitally blind individuals is both distinctive and remarkably stable over time. Using multivoxel pattern analysis, we show that the unique patterns of reorganization within the visual cortex allow to decode participant identity very accurately across the three scan intervals. These results align with recent evidence highlighting significant individual differences in visual cortex connectivity, indicating a consistent and distinct role for the visual cortex in blindness that may vary across individuals. In conclusion, our study sheds light on the stability of functional connectivity within the visual cortex of congenitally blind individuals, adding a potential to identify stable neuromarkers associated with sight rehabilitation success. Such neuromarkers could facilitate targeted and personalized treatment strategies, enhancing the overall efficacy of visual restoration efforts.

TUESDAY MORNING POSTERS IN BANYAN BREEZEWAY

TUESDAY, MAY 21, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Color, Light and Materials: Lightness, brightness

53.301 CRITERION CHANGES MIGHT INFLUENCE THE COMPARISON PROCEDURE IN MAXIMUM LIKELIHOOD DIFFERENCE SCALING (MLDS)

Yangyi Shi[1] (shi.yang@northeastern.edu), Rhea T. Eskew, Jr.[1]; [1]Northeastern University

Maximum Likelihood Difference Scaling (MLDS) is an efficient perceptual scaling method used in many psychophysical studies. In the 'Quad' version of MLDS, two stimulus pairs varying in the same physical property (for example, luminance contrast) are shown, and observers pick either the more similar pair or the more different pair according to instruction. A maximum likelihood procedure then assigns perceptual scale values to each physical level. In our previous study, we used MLDS to estimate the asymmetry of perceived luminance contrast between achromatic increments (A+) and decrements (A-) square patches. We showed that the MLDS results are consistent with pedestal discrimination results. We proposed a model, in which the perceptual scale estimated by MLDS is used to predict the forced-choice discrimination thresholds. Logically, the instruction (pick the more similar pair or the more different pair) should not influence the observers' choices, because MLDS is a two-alternative forced choice procedure, switching the instruction is just picking the other pair. However, in our study, we found that this was not always true. While some observers did not show a criterion difference, there were some observers whose estimated perceptual scales varied significantly when the instruction given to them changed. We verified this by using their MLDS results under two criterions to predict their forced-choice discrimination results, based on the model we proposed in the previous study. Our results showed that those observers' "more similar" results can predict the discrimination thresholds better than their "more different" results do. Also, the difference between criterions is larger in A- than in A+. We suggest that the criterion differences likely result from attending to different features in the stimuli, even though they are simple square patches.

NSF BCS-2239356

53.302 HETEROCHROMATIC BRIGHTNESS CHANGES BETWEEN MONITORS

Yuan Zhang[1] (zhanglhm@163.com), Karl Gegenfurtner[2], Jing Chen[1]; [1]Shanghai University of Sport, China, [2]Justus-Liebig-Universität Gießen, Germany

The RGB primaries vary substantially between different monitors. It is known that color perception adapts to the environment. For instance, there is an adaptive change in unique yellow settings across seasons (Welbourne et al., 2015). However, it is unknown whether color perception would change between monitors. Here, we simulated 3 different RGB primary ratios on one monitor, with their luminance ratio at 1:2.8:0.4, 1:3.4:0.4, and 1:3.8:0.4. These 3 simulated monitors differed only in the green primary but not in red and blue. Observers (N = 20) ranked the brightness of seven heterochromatic patches (7 principal colors: R, G, B, RG, RB, GB, RGB). We employed a non-linear max-weighted RGB model to establish the relationship between brightness perception and the weights of R, G, and B luminance values. The model predicted 88.9% of the observers' rankings correctly. The weight of the green primary decreased across the three simulated monitors (Ps < 0.01), as the intensity of the green primary increased. The weights of the other two primaries showed no significant differences (Ps > 0.3). Thus, when a monitor has a higher intensity in one primary (i.e., green primary in the current setting), this primary's contribution to brightness perception is weighted less by observers, in adaptation to the monitor color statistics. This adaptation was only partial, approximately 61.1% relative to the intensity change. When looking at the time course of adaptation, we found that the effect requires a few dozens of trials to build up. In conclusion, the present study suggests that our perception of heterochromatic brightness adapts to different monitors.

European Research Council Advanced Grant Color3.0, 884116

53.303 THE CONTRAST DEPENDENCY OF THE MUNKER-WHITE ILLUSION

Conrad Schulte[1] (conrad.schulte@student.maastrichtuniversity.nl), Chien-Chung Chen[2]; [1]Maastricht University, Netherlands, [2]National Taiwan University, Taipei, Taiwan

In the Munker-White illusion, a grey bar placed on the black stripes of a black-and-white grating appears brighter than the one place on the white stripes despite having the same luminance. We investigated the mechanisms underlying this illusion with an asymmetric matching paradigm. Participants viewed a screen displaying a point at the center and a matching bar on one side and a test pattern with a bar embedded either in a black or a white stripe of a black-and-white square wave grating on the side of the display. Participants were instructed to compare the brightness of the test with the matching bar and indicated which bar seemed brighter. In each run, the luminance contrast between the test bar and the background was either a positive or a negative constant (brighter or darker than the background respectively) while the contrast of the test bar varied from trial to trial with the participants' response and a one-up-one-down staircase algorithm. The test contrast, polarity and location all had a significant effect on the strength of the illusion measured as the difference between the test and the matching contrast). The illusion reduced quickly as the test bar contrast increased when the test was embedded in a stripe of the opposite contrast polarity. However, the illusion was enhanced first before decreasing with the test contrast when the test was embedded in a stripe of the same contrast polarity. The test contrast effect is inconsistent with the high-level theories of Munker-White illusion, such as the belongingness theory, which would suggest no test contrast effect as the stimulus configuration remained the same regardless test contrast. This effect is qualitatively consistent with the contrast theory, which predicted an increased illusion with higher test bar contrast. However, the polarity effect suggested an involvement of a nonlinear process.

53.304 EVALUATING THE ROLE OF EDGE-TO-SURFACE RECONSTRUCTION AS A UNIFYING EXPLANATION FOR SIMPLE TO COMPLEX LIGHTNESS ILLUSIONS

Srijani Saha[1,2] (srijanisaha@g.harvard.edu), George Alvarez[1,2]; [1]Harvard University, [2]Kempner Institute for the Study of Natural and Artificial Intelligence

Lightness illusions demonstrate that how bright an object appears depends on an elaborate constructive process, to the point that the same surface can be perceived as either black or white depending on the context. Why does the biological visual system work this way? Traditionally, distinct mechanisms have been proposed to account for simple lightness illusions (e.g. the Craik-O'Brien-Cornsweet Illusion) and for more complex illusions (e.g., the moon illusion: discs in different hazy backgrounds, Anderson & Winawer 2005). The Craik-O'Brien-Cornsweet illusion seems to depend on local cues — a dark/light difference at a singular edge— whereas the moon illusion seems to require more than just local cues by parsing the input into distinct scene properties. Our work examines the degree to which an edge-to-surface reconstruction mechanism can provide a unified explanation for a range of lightness illusions. First, across two behavioral tasks, 59 participants matched the brightness of a disc presented in a dark or light hazy background to illustrate the effect of the context on lightness perception. Second, we trained a reconstructive U-Net model to output a filled-in image from edge-only inputs, a computational goal that is analogous to filling in surfaces from edge-selective neurons in the biological visual system. Surprisingly, we found that, when reconstructing the discs, the U-Net model made systematic errors consistent with lightness illusions measured in people, suggesting that an edge-to-surface reconstruction is a plausible mechanism underlying this complex illusion. Finally, we applied the reconstructive model to a suite of additional lightness illusions — Adelson Haze Illusion, Snake Illusion, Koffka Illusions, 3D Cornsweet Illusion (Purves et al. 1999) and Kanizsa Square Illusion — finding that an edge-to-surface reconstruction mechanism successfully recapitulated illusions when edge information is present, and the polarity of the edge is consistent around the boundary of the illusory surface.

This work was supported by NSF PAC COMP-COG 1946308 to GAA.

53.305 TESTING PARABOLIC BRIGHTNESS MATCHING FUNCTIONS ACROSS FOUR CONTRAST POLARITY CONDITIONS

Osman B. Kavcar[1] (okavcar@nevada.unr.edu), Michael E. Rudd[1], Christabel Arthur[1], Mohana Kuppuswamy Parthasarathy[1], Michael A. Crognale[1]; [1]University of Nevada, Reno

It is well established that brightness (perceived luminance) or lightness (perceived reflectance) of a patch is modulated by its surroundings. Such influence can be measured using the disk/annulus paradigm (Wallach, 1948) where the luminance of a match disk is adjusted to match the appearance of a target disk that is surrounded by an annulus. Parabolic Brightness Matching Functions (BMFs) are obtained when a 2nd-order polynomial model is fitted to the average perceptual matches made at varying annulus luminance levels on a log-log scale. Rudd (2010) proposed a model that explains the parabolic BMFs, on which a contrast gain control operates between the outer and inner edges of the annulus to influence the target brightness. The model posits that, when the first-order coefficient (k1) of BMFs is plotted against the second-order coefficient (k2), the resulting plot should be a linear function whose slope depends on the contrast polarity of the target disk to its annulus and the luminance of the target disk. Rudd, Kavcar, and Crognale (2023) showed that this linear relationship does not depend on the annulus luminance or the annulus size but instead depends on the contrast polarity—all in-line with the model predictions. Here, we further tested the model by varying the target disk luminance to four additional levels. The model predicts that the slope of the k1 vs k2 plot should itself vary as a linear function of the target disk luminance, and rate of change will be -1.0 for conditions where match disk is decrement to its background and -0.5 where the disk is increment. Our results confirmed the first prediction (strong, negative relationship) across four contrast polarity conditions, but the specific slope predictions were disconfirmed.

53.306 THE RELATIONSHIP OF LIGHTNESS ILLUSIONS UNCOVERED BY INDIVIDUAL DIFFERENCES AND ITS ADVANTAGE IN MODEL EVALUATION

Yuki Kobayashi[1,2] (y-koba@gst.ritsumei.ac.jp), Arthur Shapiro[1]; [1]American University, [2]Ritsumeikan University

Numerous computational models have been proposed to account for brightness/lightness illusions. To compare these models, researchers often start with a set of test illusions and then count the number of illusions in the set that each model can correctly predict. Models are considered successful if they can account for more illusions than other models; however, since lightness illusions are not independent of each other, one class of model may seem better than another class of model, but the ranking of may be due to a stimulus set that overrepresents correlated illusions. Here, we collected the magnitudes of various lightness illusions through two online experiments and then examined the response with an exploratory factor analysis. We found that the illusions in typical test sets can be divided into three classes: assimilation illusions, contrast illusions, and Whites effect combination illusions. We also had one example of a contrast enhancement illusion

(Agostini's glare illusion) that seemed to be separate from the other three classes. We then examined three well-known computational models (ODOG, LODOG, and FLODOG) using the obtained information about relationships of the illusions. We show that the assumption of illusions' independence does not markedly distort the relative evaluation of the models, but performances of some models are substantially unbalanced across the three classes. ODOG and LODOG tend to be better at contrast effects, while FLODOG has better balance across three classes, but all three models were not very good at accounting for the contrast enhancement example. The results are consistent with the idea that there are separate processes for assimilation and contrast and a higher order stage that combines these separate processes. The study highlights the need to assess model performance based on their explanation of underlying processes rather than focusing solely on individual illusions.

The first author received financial support from the Japan Society for the Promotion of Science (Grant No. 22K13878).

53.307 DYNAMICAL NEURAL MODEL OF LIGHTNESS COMPUTATION AND PERCEPTUAL FADING OF RETINALLY STABILIZED IMAGES

Michael Rudd[1] (mrudd@unr.edu), Idris Shareef[1]; [1]University of Nevada, Reno

We recently proposed a neural model that accounts to within <6% error for lightness matches made to Staircase Gelb and simultaneous contrast displays comprising real illuminated surfaces. Here, we demonstrate how the model accounts for the perceptual fading that occurs when images are stabilized on the retina (Troxler, 1804, Riggs et al., 1953). In the model, cortical lightness computations are derived from transient ON and OFF cell responses in the early visual system that are generated in the course of fixational eye movements. The ON and OFF responses are sorted by eye movement direction in visual cortex to produce a set of spatiotopic maps of ON and OFF activations. Activations within these maps trigger spatial filling-in of lightness and darkness within independent ON and OFF networks, which are combined at the final modeling stage to compute perceived reflectance (lightness). We elaborate these mechanisms to produce a more detailed neurophysiological theory. We propose how early temporal responses of ON and OFF cells are read out (decoded) in visual cortex to trigger lightness and darkness induction signals, and we explicitly model cortical magnification, which further improves the fit to psychophysical data. Two key takeaways are: 1) the model accounts for multiple lightness phenomena, including fading of stabilized images, with high quantitative precision and in a biologically plausible way; 2) estimated rates of fixational eye movements known as microsaccades (Martinez-Conde et al., 2004) are too low to explain the dynamics of lightness phenomenology. We suggest that the higher rate eye movements known as tremor can better account for the perceptual data within the context of an otherwise identical neural framework. Correspondences between the model's processing stages and cortical neurophysiology will be discussed, and the computations performed at different model stages will be illustrated through a combination of still images and movies.

53.308 LIGHTNESS ILLUSIONS THROUGH AI EYES: ASSESSING CONVNET AND VIT CONCORDANCE WITH HUMAN PERCEPTION

Jaykishan Patel[1,2] (jay96@my.yorku.ca), Alban Flachot[1,2], Javier Vazquez-Corral[3,4], Konstantinos George Derpanis[1,5], Richard Murray[1,2]; [1]York University, [2]Center for Vision Research, [3]Universitat Autònoma de Barcelona (UAB), [4]Computer Vision Center, UAB, [5]Lassonde School of Engineering

Inferring surface reflectance from luminance images has proven to be a challenge for models of human vision, as many combinations of illumination, reflectance, and 3D shape can create the same luminance image. Traditional models struggle with this deep ambiguity. Recently, convolutional neural networks (CNNs) and vision transformers (ViTs) have been successful computer vision approaches to inferring surface colour. These architectures have the potential to be foundational models for lightness and color perception, if they process image information similarly to humans. We trained CNN and ViT backbones including ResNet18, VGG19, DPT, and custom designs to infer surface reflectance from luminance images using a custom dataset of luminance and reflectance images generated in Blender. We used these models to infer surface reflectance from several well-known images that generate strong lightness illusions, including the argyle, Koffka-Adelson, snake, simultaneous contrast, White's, and checkerboard assimilation illusions, as well as their control images. These illusions are often thought to result from the visual system's attempt to infer surface reflectance from ambiguous images using the statistics of natural images, and we hypothesized that networks trained on simple scenes rendered with shading and shadows would be susceptible to similar illusions. We found that all networks did in fact predict illusions in most test images, and predicted stronger illusions than in the control conditions. The exceptions were that the models typically failed to predict the argyle illusion, and to predict assimilation illusions. Model saliency analysis showed that the networks' outputs were strongly dependent on pixel-information in the shadowed regions of the image. These results support the hypothesis that some lightness phenomena arise from the visual system's use of natural scene statistics to infer reflectance from ambiguous images, and show the potential of CNNs and other deep learning architectures as starting points for models of human lightness and colour perception.

TUESDAY, MAY 21, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Scene Perception: Ensembles, natural image statistics

53.309 AVERAGE TEMPERATURE FROM VISUAL SCENE ENSEMBLES WITHOUT RELIANCE ON COLOR, CONTRAST OR LOW SPATIAL FREQUENCIES

Vignash Tharmaratnam[1], Dirk Bernhardt-Walther[2], Jonathan S. Cant[1]; [1]University of Toronto Scarborough, [2]University of Toronto

Summary statistics for groups (i.e., ensembles) of faces or objects can be rapidly extracted to optimize visual processing, without reliance on visual working memory (VWM). We have previously demonstrated that this ability extends to complex groups of scenes. Namely, participants were able to extract average scene content and spatial boundary from scene ensembles. In the present study we tested whether this ability extends to scene features that are not solely attributable to visual processing and are instead computed cross-modally. Specifically, we examined ensemble processing of scene temperature. Given that the apparent temperature (i.e., how hot or cold a scene would feel) of single scenes is accurately rated by observers (Jung & Walther, 2021), we predicted that average scene temperature could be extracted by observers, without reliance on VWM. Crucially, across 4 experiments, we tested if this ability would depend on low-level visual features. Participants rated the average temperature of scene ensembles, with either colored stimuli (Exp 1), gray-scaled stimuli (Exp. 2), gray-scaled stimuli with a 75% contrast reduction (Exp. 3), or gray-scaled high spatial frequency filtered stimuli (> 6 cycles/degree, Exp. 4). In all experiments, we varied set size by randomly presenting 1, 2, 4, or 6 scenes to participants on each trial, and measured VWM capacity using a 2-AFC task. Participants were able to accurately extract average temperature in all experiments, with all 6 scenes being integrated into their summary statistics. This occurred without relying on VWM, as fewer than 1.2 scenes were remembered on average. These results reveal that computing cross-modal summary statistics (i.e., average temperature) does not rely on lower-level visual features. Overall, these results reveal that with minimal low-level visual information available, abstract multisensory information can be rapidly retrieved and combined from long term memory to form statistical representations.

53.310 DO ENSEMBLE REPRESENTATIONS GUIDE VISUAL ATTENTION IN A VISUAL SEARCH TASK?

Kristina Knox[1,2] (kristina.knox@mail.utoronto.ca), Jay Pratt[1], Jonathan S. Cant[2]; [1]University of Toronto, [2]University of Toronto Scarborough

Ensemble processing plays an important role in our daily lives by condensing abundant visual information in our environment into statistical representations. Our study examined how these statistical representations are prioritized in the attentional system by asking whether ensemble representations, such as the average orientation of a set of items, can guide attention in a subsequent task. To explore this, we integrated an orientation-based ensemble-processing task with a visual search task. On each trial, participants were shown an initial display of eight bars of varying orientations. The subsequent task—either a search or an average task—was signalled by the colour of the fixation cross. When the cross turned orange (25% of trials), participants engaged in the search task. They had to locate and click on the shortest bar among six others displayed around the fixation point. Importantly, in half of these search displays, the target bar matched the average orientation of the initial eight-bar display. When the fixation cross turned blue (75% of trials), participants performed the average task. This task involved a display of two bars to the left and right of the fixation point, and participants had to determine which of these two bars corresponded to the average orientation of the initial eight-bar display. In both tasks, participants were instructed to respond as quickly and accurately as possible. The results revealed shorter response times (RTs) in the search task when the target bar matched the average orientation of the initial eight-bar display compared to when the orientation of the target bar did not match the average orientation of the initial display. On a local level, this finding indicates that ensemble representations guide attention in subsequent tasks. On a global level, this means that the representation of an item never explicitly perceived can guide attention and subsequent behaviour.

53.311 LEARNING TO ENSEMBLE? REPEATED EXPOSURE LEADS TO MORE EFFICIENT PROCESSING OF SUMMARY STATISTICS FOR WEIGHT AND REAL-WORLD SIZE

Greer Gillies[1] (greer.gillies@mail.utoronto.ca), Keisuke Fukuda[1,2], Jonathan S. Cant[1,3]; [1]University of Toronto, [2]University of Toronto, Mississauga, [3]University of Toronto, Scarborough

Ensemble coding is the brain's ability to rapidly extract summary statistics from groups of similar items (e.g., average colour of leaves on a tree). Ensemble coding has also been found for information that cannot be gleaned solely from retinal input (e.g., average animacy of objects). We extended this line of ensemble coding research by examining if observers were sensitive to two different features that require access to stored information in long-term memory. Specifically, participants made judgements about the average weight or real-world size of groups of objects. We found that participants were unable to integrate information from multiple items to produce accurate summary statistics for either judgement. Next, we examined how learning and memory may influence ensemble coding for weight and real-world size. Could associative learning (e.g., learning which items belong in an ensemble via their relationship to other items) aid in ensemble coding? Specifically, we examined if repeated exposures to the different items within an ensemble would enable people to form summary statistics more efficiently. To test this, participants made judgements about the average weight or average real-world size of groups of object stimuli, and we manipulated how frequently people were exposed to certain subsets and combinations of the stimuli in a given ensemble. We found that repeated presentation of the ensemble stimuli leads to improved performance on the ensemble tasks for both the weight and size judgements. In summary, while observers were unable to extract average weight and real-world size information from objects when stimuli were not repeatedly presented, they were able to produce accurate summary statistics with multiple stimulus exposures. We speculate that this may be due to mechanisms governing associative statistical learning and memory for repeatedly encountered visual information.

53.312 MODELLING THE REPRESENTATION OF VISUAL ENSEMBLES IN THE HUMAN BRAIN

Patxi Elosegi[1] (pelosegi@bcbl.eu), Ning Mei[2], Yaoda Xu[3], David Soto[1,4]; [1]Basque Center on Cognition, Brain and Language, Donostia, Spain, [2]Psychology Department, Shenzhen University, Shenzhen, China, [3]Psychology Department, Yale University, New Haven, CT 06511, United States, [4]Ikerbasque, Basque Foundation for Science, Bilbao, Spain

The human brain compensates for processing capacity limitations by compressing redundant features of the visual input in the form of ensemble representations. While psychophysical evidence demonstrates that ensemble representations can efficiently be extracted for low- mid- and high-level visual features, neuroimaging studies have provided mixed results regarding the neural underpinnings of this ability. Studies employing low-level stimuli indicate the involvement of ventro-occipital areas in ensemble perception, while those using high-level stimuli show activations primarily in parietal regions. However, to date, all fMRI studies on ensemble perception have predominantly focused on univariate activation changes, overlooking potential differences at the multi-voxel level. Here, we aim to characterise the representational geometry of high-level animacy ensemble perception, combining computer vision models and high-precision human fMRI data (5 sessions per participant; N=2).This exploration involves understanding the relationships among high-dimensional brain patterns, offering a more comprehensive perspective than traditional univariate analysis on how ensemble information is encoded in the BOLD signal. Behaviourally, we show that participants could successfully discriminate the predominant class well beyond chance in a 12 item ensemble, comprising living and non-living objects, even in the most difficult conditions. At the neural level, searchlight representational similarity analysis showed that DenseNet-169 penultimate layer representations, when averaged across the single-items comprising the ensemble, predicted activity across dorso-parietal and frontal substrates. This work provides a framework to characterise the representational geometry of ensembles in the brain and offers evidence supporting the involvement of the dorsal visual pathway in high-level ensemble perception.

P.E. acknowledges support from the Basque Government PREDOC grant

53.313 MULTIPLICATIVE EFFECT OF TASK-IRRELEVANT FEATURE VARIABILITY ON VARIABILITY JUDGMENTS

Suyeon Kim[1] (eoddysl@naver.com), Oakyoon Cha[1]; [1]Department of Psychology, Sungshin Women's University

Pooling within the neural population processing a specific feature, or population coding, has been proposed as a mechanism for ensemble judgments of simple features (e.g., mean orientation judgments, orientation variability judgments). According to the population coding model, task-irrelevant features have no influence on ensemble judgments, and if they do, the influence would be additive since they can only influence judgments after the target ensemble property has been estimated. In present study, we investigated whether task-irrelevant feature variability has additive or multiplicative influence on variability judgments. To this end, participants were asked to compare the orientation variability between the two arrays of tilted oval shapes. Tilted oval shapes were randomly placed within a 4-by-5 grid on one of the two arrays (random array). In the other array, they were placed so that their orientations were sorted from counterclockwise to clockwise starting from the top left corner of the grid (sorted array). We expected participants to judge the sorted array as less variable if the orientation variability was similar between the two arrays. Furthermore, we manipulated the color variability of the two arrays. In half of the trials, oval shapes had a uniform color within each array, while in the other half, they had variable colors. If the influence of task-irrelevant color variability is additive, differences in the perceived variability between the random and sorted arrays would decrease. On the other hand, if the influence is multiplicative, differences in the perceived variability would increase. Participants consistently judged the sorted array as less variable when the two arrays had the same orientation variability, and the introduction of task-irrelevant color variability led to increased differences between the perceived variability of random and sorted arrays. These results suggest that task-irrelevant feature variability influences ensemble judgments prior to the estimation of target ensemble property.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2023-00211668).

53.314 FACE DETECTION SENSITIVITY FOLLOWS THE SPATIAL DISTRIBUTION OF EXPERIENCED FACE LOCATIONS

Michelle Greene[1] (mgreene@barnard.edu), Benjamin Balas[2], Quinn O'Connor[1], Kaiki Chiu[1], Emily Lo[1]; [1]Barnard College, Columbia University, [2]North Dakota State University

Visual perception is influenced by the statistical properties of our environment. Notably, cortical regions with visual selectivity for faces and body parts exhibit a topographic relationship (Weiner & Grill-Spector, 2011) that evokes the relative locations of heads to bodies. Here, we explore the hypothesis that detection sensitivity for faces and hands is biased according to their frequent spatial locations in the visual field. We used the Google MediaPipe face and hand detectors to analyze over 200 hours of the Visual Experience Dataset, comprising more than 23.1 million frames of egocentric video. Our findings revealed a distinct spatial distribution: the median center of faces was located 13 degrees above central fixation, whereas hands typically appeared 23 degrees below (Wilcoxon rank sum, $p<0.0001$). We reasoned that these position statistics would translate into enhanced face detection in the upper visual field and better hand detection in the lower field. To test this, we embedded faces and hands within 1/f noise at 50% transparency (with chroma values selected from a range of skin tones). Targets were presented in either the upper or lower visual field. In a challenging detection task, 22 participants identified the presence of a body part (face or hand) versus 1/f noise within an 80 ms exposure followed by a dynamic pattern mask. Our face detection results supported our hypothesis, with higher accuracy in the upper visual field (81%) compared to the lower (75%, $p<0.05$). However, hand detection did not exhibit a significant difference between the two fields (70% vs 72%). These outcomes will be discussed in relation to the spatial selectivity of face (FFA and OFA) and body (EBA) areas and the potential for complex relationships between visual field biases, eye movement patterns, and peripheral visual recognition.

NSF 1920896 to MRG and BJB.

53.315 INTEGRATIVE PROCESSING IN DEEP NEURAL NETWORKS AND HUMAN VISUAL CORTEX PREDICTS THE BEAUTY OF NATURAL SCENES

Sanjeev Nara[1] (sanjeev.nara@math.uni-giessen.de), Daniel Kaiser[1,2]; [1]Mathematical Institute, Department of Mathematics and Computer Science, Physics, Geography, Justus Liebig University Gießen, Germany., [2]Center for Mind, Brain and Behavior (CMBB), Philipps-University Marburg and Justus Liebig University Gießen, Germany

During daily life, we encounter a large variety of natural scenes. Some of them appear beautiful to us while others do not. Research in empirical aesthetics demonstrates that the beauty of natural images is already determined during perceptual analysis. Although research has characterized preferences for certain visual features over others, it remains unclear which overarching perceptual computations give rise to the perception of beauty. Here, we tested whether the perceived beauty of natural scenes can be predicted by the amount of spatial integration, a perceptual computation that reduces processing demands by aggregating image elements into more efficient representations of whole images. Theories of processing fluency suggest that the ease of visual analysis is a critical factor for experiencing beauty. Hence, we reasoned that increasing amounts of integration reduce processing demands in the visual system, thereby leading to an increase in perceived beauty. We quantified integration in a deep neural network (DNN) model trained on scene categorization. Specifically, we compared DNN activations averaged across complementary image halves to activations for the whole image. We quantified integration as the deviation between activations to the whole and the average of the parts. Critically, the degree of integration was indeed positively related to beauty ratings across four studies featuring different images and task demands. By manipulating the images supplied to the DNN in targeted ways, we further charted the contribution of a set of candidate visual properties to this prediction. We show that neither basic features like color or luminance, nor high-level configural properties exclusively drive predictions of beauty. Finally, complementary fMRI recordings from human participants revealed that integration in scene-selective visual cortex predicts perceived beauty in a similar way as integration in DNNs. Together, our results establish integration as a computational principle that eases perceptual analysis and thereby predisposes the perception of beauty.

53.316 LEVERAGING VISION AND LANGUAGE GENERATIVE MODELS TO UNDERSTAND THE VISUAL CORTEX

Andrew Luo[1] (luo_andrew@outlook.com), Margaret Henderson[1], Leila Wehbe[1], Michael Tarr[1]; [1]Carnegie Mellon University

Understanding the functional organization of the higher visual cortex is a fundamental goal in neuroscience. Traditional approaches have focused on mapping the visual and semantic selectivity of neural populations using hand-selected, non-naturalistic stimuli, which require a priori hypotheses about visual cortex selectivity. To address these limitations, we introduce two data-driven methods: Brain Diffusion for Visual Exploration ('BrainDiVE') and Semantic Captioning Using Brain Alignments ('BrainSCUBA'). BrainDiVE synthesizes images predicted to activate specific brain regions, having been trained on a dataset of natural images and paired fMRI recordings, thus bypassing the need for hand-crafted visual stimuli. This approach leverages large-scale diffusion models combined with brain-gradient guided image synthesis. We demonstrate the synthesis of preferred images with high semantic specificity for category-selective regions of interest (ROIs). This method further enables the characterization of differences and novel functional subdivisions within ROIs, which we validated with behavioral data. BrainSCUBA, on the other hand, generates natural language descriptions for images predicted to maximally activate individual voxels. Utilizing a contrastive vision-language model and a pre-trained large language model, BrainSCUBA generates interpretable captions, enabling text-conditioned image synthesis. This method shows that the generated images are semantically coherent and achieve high predicted activations. In exploratory studies on the distribution of 'person' representations in the brain, we observe fine-grained semantic selectivity in body-selective areas. Together, these two methods offer well-specified constraints for future hypothesis-driven examinations and demonstrate the potential of data-driven approaches in uncovering visual cortex organization.

This work used Bridges-2 at Pittsburgh Supercomputing Center through allocation SOC220017 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296.

53.317 SIMILAR VISUAL COMFORT RATINGS FOR NATURAL TEXTURES AND DISEASE IMAGERY BY TRYPOPHOBIC AND NON-TRYPOPHOBIC INDIVIDUALS

Christopher DiMattina[1] (cdimattina@fgcu.edu), R. Nathan Pipitone, Martin R. Renteria, Kriston J. Ryan; [1]Florida Gulf Coast University

Trypophobia is the discomfort felt by some individuals when viewing images containing clusters of bumps or holes. Several evolutionary hypotheses have been proposed to explain this phenomenon, with more current literature supporting the notion that visual discomfort functions as part of the behavioral immune system, helping organisms avoid skin disease and/or ectoparasites. Although both skin disease imagery and trypophobic imagery are visual textures, to date there has been no direct comparison of the visual discomfort elicited by these textures in the larger context of the discomfort elicited by a wider variety of natural textures. In this study, we measured participants level of trypophobia using the Trypophobia Questionnaire (TQ) and recorded the visual comfort ratings elicited by a large set of natural texture images from the Brodatz database, trypophobic textures, and skin disease textures. Results showed that while all observers find skin diseases uncomfortable to view, only those scoring high on the TQ rated trypophobic imagery equally uncomfortable. We further observe that the high-TQ participants rated skin disease images as being significantly more uncomfortable than low-TQ participants. Comparing the ratings for high-TQ and low-TQ participants to the standard textures, we find remarkably consistent rank-order preferences, with the most unpleasant textures (as rated by both groups) exhibiting qualitative similarities to trypophobic imagery. However, we find that low-level image statistics which have been previously shown to affect visual comfort are poor predictors of the visual comfort elicited by natural textures, and greatly under-predict the visual discomfort elicited by trypophobic or disease imagery. Our results suggest that a full understanding of the visual comfort elicited by natural textures, including those arising from skin disease, will ultimately depend upon a better understanding of the perception of surface and material properties, and how these percepts of surface and material properties evoke appropriate behavioral responses like disgust.

53.318 THE VISIBILITY OF EIDOLON DISTORTIONS IN THINGS AND STUFF

Swantje Mahncke[1] (swantje.mahncke@tu-darmstadt.de), Lina Eicke-Kanani[1], Thomas S.A. Wallis[1,2]; [1]Technical University of Darmstadt, Germany, [2]Centre for Mind, Brain and Behaviour (CMBB), Universities of Marburg, Giessen and Darmstadt, Germany

The differentiation between things and stuff, man-made and natural scenes or different scene complexities have been previously identified as key components of scene appearance. However, there remains uncertainty around how to categorize these aspects of scenes using image-based metrics. Groen et al. (Journal of Neuroscience, 2013) found that natural and man-made image content could be loosely characterized using two dimensions: spatial coherence (variation in edge density such that low variation means high coherence) and contrast energy (average local contrast). Here, we find that these statistics can similarly differentiate images in the THINGS and STUFF databases. While the two databases are not perfectly separated, images from THINGS tend to have high spatial coherence and high contrast energy (scene-like), where as images from STUFF tend to have low spatial coherence and low contrast energy (texture-like). To test whether variation in these statistics is correlated with differences in perceptual processing, we examined human sensitivity to Eidolon distortions in sets of images from each of these two quadrants, independent of their database membership. Participants discriminated between a natural and an Eidolon-distorted image in a 2IFC task, for different distortion intensities (reach) and spatial frequencies (grain). Images were presented 6.4 degrees to the right of fixation, and subtended 7.5 degrees in diameter. We found that Eidolon distortions were easier to detect (lower reach thresholds) at all grain values in more scene-like images compared to texture-like images. Together, these data indicate that the low-dimensional representation of spatial coherence and contrast energy can provide a placement of images onto a scale ranging from things to stuff, at least in terms of perceptual sensitivity to spatial distortions.

Co-funded by the European Union (ERC, SEGMENT, 101086774), and the cluster project "The Adaptive Mind" as part of the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art.

53.319 WHAT MAKES UP THE GIST OF ABNORMALITY IN MAMMOGRAMS?

Karla Evans[1] (karla.evans@york.ac.uk), Cameron Kyle-Davidson[1]; [1]University of York

Humans have rapid access to global structural and statistical regularities which allows them to extract the "gist" of an image, central to efficient assessment and orienting in complex environments. This ability is most likely based on our experience with the regularities of the natural world. Mammograms can be thought of as a specialized class of scenes and radiologists as experts who have tuned their visual system to regularities in these unusual scenes. Consequently, we have found that the gist of the abnormal in radiographs, viewed only for 500 milliseconds, allows radiologists to detect the presence of disease independent of the locus of any lesion and up to 3 years before the onset of cancer. Using three different approaches to image analysis we aimed to determine the textural components which contribute to the gist signal. We first computed the mammographic power spectrum via the Fourier transform of patches drawn from the parenchyma. Secondly, we employed a radiomics toolkit to compute 90 different textural features related to cancer. Thirdly, we used Portilla-Simoncelli analysis to identify differences across different textural scales. For frequency data, we compared spectra directly,and using machine learning techniques identified frequencies which appear most important for dividing normal from abnormal data. Within the radiomic feature spaces, we used the same process to identify the most important radiomic features. As the textural meaning of these features is obfuscated, we synthesized artificial mammograms which display a heightened degree of important radiomic markers, and pass them through a deep neural network trained to classify mammographic patches. This reveals the contribution of these radiomic features to abnormality, and their textural impact. Finally, through Portilla-Simoncelli modeling, we find significant differences across scales. Together, these metrics are indicative of where in the texture the experts' visual system is tuning to the gist of the abnormal.

This research was funded by Cancer Research UK and Engineering & Physical Sciences Research Council (EPSRC) Early Detection and Diagnosis Project award EDDCPJT/100027 to Karla K. Evans

53.320 'ATTENTIONAL TRANSPLANTS' CAUSE RECIPIENTS TO LIKE IMAGES SIMILARLY TO DONORS: EVIDENCE FOR INTER-OBSERVER COMMONALITIES IN HOW ATTENTION DRIVES PREFERENCES

Hong B Nguyen[1] (nguyh376@newschool.edu), Benjamin van Buren[1]; [1]The New School

When different people view a scene, they attend to different things, and these differences in attention influence how much they like the scene. Patterns of attention may be highly individually specific. However, the effects of different patterns of attention on preferences may not be. Here we demonstrate this, using a new method of 'attentional transplants'. We show that, if an observer likes an image, it is possible to transplant their viewing pattern into another observer—and that this causes the recipient to like the image better, compared with transplanting the viewing pattern of a donor who disliked the image. In Experiment 1, 50 observers viewed images of landscapes by using their cursor to move a small circular viewing window around each image for three seconds. After viewing each image, they rated how much they liked it. For each image, we identified two 'attentional donors'—the Liked-it-Best observer who rated the image highest (relative to their other image ratings) across observers, as well as the Liked-it-Least observer who rated the image lowest across observers. Next, we recruited 100 new observers to serve as 'attentional recipients'. These observers viewed each image, but now passively, through a moving window which reproduced the viewing pattern of either the previous observer who Liked it Best, or the observer who Liked it Least. Recipients gave substantially higher ratings to an image when they received the viewing pattern of the observer who Liked it Best, compared to when they received the viewing pattern of the observer who Liked it Least. In subsequent experiments, we replicated this effect, and found that Liked-it-Best attention patterns are more predictable. We conclude that individual differences in preferences for scenes are partly explained by differences in how we attend—but that attention drives preferences in similar ways across observers.

TUESDAY, MAY 21, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Visual Memory: Working memory and attention

53.321 VISUAL EXPLORATION GOES DOWN WITH HIGHER WORKING MEMORY LOAD, WHILE MORE SACCADES NEGATIVELY IMPACT RECALL

Amit Rawal[1,2] (amitrawal1643@gmail.com), Rosanne L. Rademaker[1]; [1]Ernst Strüngmann Institute for Neuroscience in Cooperation with the Max Planck Society, Frankfurt, Germany,

[2]Faculty of Behavioural and Human Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands

During active visual exploration and engagement with our daily environment, we can also entertain many thoughts. Given a limited processing capacity, a balance must be struck between how much we prioritise the information held in mind, and how much we explore the external world. With rising memory demands, might the tendency to sample new inputs go down? Conversely, might visual exploration impact how well a memorised item is recalled? To investigate how internal memory maintenance and external exploration interact, we varied working memory load while participants freely viewed images with varying semantic content. While we tracked their gaze, 20 participants remembered 1, 3, or 5 randomly oriented gratings for 5.5s until cued to report one item via method-of-adjustment. During the delay, participants could freely view an intact natural scene, or a locally phase-scrambled natural scene, for 5s. As expected, recall error rises monotonically with increasing set size. Semantic content also impacted recall, with scrambled images resulting in smaller errors than natural images. To explore if memory load impacts ocular behaviour, we looked at scanpath length (total distance travelled) and number of saccades. Load had no effect on the number of saccades, but higher load did lead to shorter scanpaths. Natural scenes induced more saccades and longer scanpaths compared to scrambled scenes. To explore if ocular behaviour is related to memory performance, we median-split trials by recall error, and found a larger number of saccades on trials with poorer recall, but no association between performance and scanpath length. This was true irrespective of load or semantic content, and complements our previously discovered positive association between recall error and saccade rate during fixation. In sum, we propose a bi-directional link between ocular behaviour and memory recall, where mental contents impact visual exploration, and the amount of ocular activity impacts recall.

53.322 DOES SENSORY RECRUITMENT SUPPORT GUIDANCE OF ATTENTION BY VISUAL WORKING MEMORY?

Zexuan Niu[1] (zexuan-niu@uiowa.edu), Andrew Hollingworth[1]; [1]UNIVERSITY OF IOWA

Visual working memory (VWM) maintains visual information to support ongoing cognitive processes. This core function requires that VWM resist perceptual interference from intervening sensory events, such as scene or gaze changes. On the other hand, many theories of attention hold that VWM representations interact dynamically with sensory processing to boost the salience of matching items. These two, potentially conflicting, demands could be reconciled if VWM content can be maintained across different brain regions and at different levels of abstraction: Higher-level representations in extra-sensory regions may maintain stable representations, whereas sensory-level representations interact with new perceptual processing. Here, we tested whether sensory recruitment plays a functional role in the guidance of attention. First, we tested whether sensory activation alone is sufficient to guide attention. Participants were exposed to a color before viewing a search array in which either the target (valid) or a distractor (invalid) matched that color. The color was not predictive of the target, and there was no demand to remember it. We observed no evidence of attention guidance from the exposed color (i.e., no validity effect), even at color-array SOAs in the range of iconic persistence, indicating that sensory activation alone is insufficient for attention guidance. We further tested whether VWM-based sensory recruitment is necessary for attention guidance. The basic task was modified so that participants were required to remember the color. In addition, a multi-color mask was or was not presented at variable times during the interval between the color and search array. Robust attention guidance was observed. Critically, the magnitude of attention guidance was unaffected by perceptual masking. The results do not support the hypothesis that VWM-based guidance effects have a specific sensory locus, which is potentially consistent with the claim that sensory recruitment plays a minimal role in functions attributed to VWM.

53.323 CHUNKING IN VISUAL WORKING MEMORY CHANGES THE GUIDANCE OF ATTENTION IN A VISUAL SEARCH

Logan Doyle[1] (logan.doyle@mail.utoronto.ca), Susanne Ferber[2]; [1]University of Toronto

Pairing two colors regularly across displays enables them to be accessed by visual working memory (VWM) more efficiently in a process often called chunking (Brady et al., 2009). Previous investigations have revealed that this VWM advantage may be due to the contribution of explicit LTM (Huang & Awh, 2018; Ngiam et al., 2019). Simultaneously, colors actively maintained in VWM guide the deployment of attention (van Moorselar, 2014) separately from representations maintained in LTM (Carlisle et al., 2011). We asked if the representational changes that underlie chunking also impact how those VWM representations guide attention. Using an incidental capture design, we presented participants with four reliable (high-probability) color pairs across 10 blocks of VWM tests or visual search trials. Attentional guidance of the maintained colors was calculated as the difference in reaction time as a function of the distance between the search target and the maintained color pair. Across three experiments, we found that participants with full explicit awareness of the color pairings were significantly more accurate than unaware participants in the VWM task ($F(1, 266) = 22.93$, $p < 0.001$), replicating Ngiam et al. (2019). Surprisingly, aware participants were significantly less guided toward those high-probability pairs in the search task when compared to unaware participants ($F(1, 36) = 4.77$, $p < 0.05$). This slowing is attributable to chunking, as it is limited to aware participants maintaining high-probability pairs. There is no difference between aware and unaware participants for low-probability pairs ($F(1, 32) = 0.187$, $p = 0.67$) or search displays with single colors ($F(1, 42) = 1.207$, $p = 0.28$). Overall, participants who leveraged chunked representations to improve in the VWM task show slowed attentional capture of maintained colors during the visual search task.

53.324 CONTEXTUAL CUEING IN CHANGE DETECTION TASKS

Courtney Turner[1] (courtney.l.turner@durham.ac.uk), Mikel Jimenez[1], Anna Grubert[1]; [1]Durham University

Visual scenes contain detailed features and objects at certain locations that form a global visuospatial context. In attention research, visual search was found to be faster and more accurate in constant as compared to random global contexts, i.e., when a target was repeatedly presented at a fixed location within a spatially constant context of distractors. Such contextual cueing effects demonstrate that global contextual information is learned implicitly to facilitate search. Here we ask whether contextual cueing can directly be measured at the level of working memory (which is used to guide selection in visual search). We measured working memory capacity K and the contralateral delay activity (CDA) of the event-related potential under low and high-load conditions of a visual change detection task. In different blocks, memory displays contained two or four coloured squares. Participants had to retain these colours at their spatial locations for 1000ms and then compare them to a colour set in a test display in which the colours were either identical (50% no-change trials), or one of them changed (50% change trials). The critical manipulation concerned the global context of the memory displays – in half of the trials, the memory displays showed the same colour squares at fixed locations (constant context), in the other half, the colours and locations of the squares were selected completely randomly (random context). K values and CDA amplitudes (measured during the retention period) in response to trials with constant as compared to random contexts were substantially increased, both in the low and high memory load conditions. These converging behavioural and electrophysiological findings suggest that implicitly learned global stimulus configurations directly affected visuospatial working memory capacity. In a wider framework, this observation proposes that previously observed contextual cueing effects in visual search may depend on learning effects in visuospatial working memory.

This work was funded by a research grant of the Leverhulme Trust (RPG-2020-319).

53.325 PREDICTIVE PROCESSES OF OBJECT-SCENE INTEGRATION SUPPORT ATTENTIONAL AND MEMORY MECHANISMS IN HEALTHY AND PATHOLOGICAL COGNITIVE AGEING

Elena Allegretti[1] (elena.allegretti@uniroma1.it), Moreno I. Coco[1,2]; [1]Sapienza, University of Rome, [2]I.R.C.C.S. Fondazione Santa Lucia, Rome

The consistency between the semantics of an object and its embedding scene (e.g., a torch vs. toothpaste in a bathroom) modulates its attentional prioritization (e.g., foveal vs. extrafoveal effects) and memory processes associated with it (e.g., the strength of encoding). Yet, little is known about the influence of healthy and pathological cognitive ageing on these mechanisms. In two independent eye-tracking studies: (a) younger (N = 26) and healthy older participants (N = 24), and (b) people with MCI (N = 30) and healthy age-matched controls (N = 32) were asked to detect whether an object, either consistent or inconsistent with the scene context, changed on its identity (became another object), location (moved to another position) or concurrently both these features. During successful recognitions, only younger participants displayed a significantly higher probability of fixating as first targets on inconsistent objects compared to consistent ones. When the object remained in the same spatial location (i.e., identity change) all groups displayed a shorter latency of first fixation for objects encoded as inconsistent. Moreover, in general, gaze duration was higher for inconsistent compared to consistent objects, again independent of age or pathology. Finally, changes happening on inconsistent objects were detected better across all groups but only when they moved location, which is the only condition where the object maintains the same consistency between encoding and recognition. Overall, our results extend our understanding of object-scene integration mechanisms, a timely topic in vision science, by revealing that object semantics guides early overt attention and influences short-term memory processes independently of healthy, and even pathological, cognitive ageing. The only subtle age-related difference we observed was possibly a decline in the useful field of view, which has reduced the accessibility of object semantics in extrafoveal vision.

Funded by Progetti di Ricerca Medi, RM122181673FAF68 (Sapienza, University of Rome) awarded to M. I. Coco.

53.326 FAST-TRACKING IMPROVEMENTS OF METACOGNITIVE ASSESSMENTS OF VISUAL WORKING MEMORY

Hana Yabuki[1] (hana.yabuki@mail.utoronto.ca), Caitlin Tozios[1], Susanne Ferber[1], Keisuke Fukuda[1,2]; [1]University of Toronto, [2]University of Toronto Mississauga

Not only is visual working memory (VWM) limited in capacity, some VWM representations may be maintained inaccurately even though we feel confident about them. Given that such confident errors can cause severe costs (e.g., traffic accidents), we tested a new approach to improve observers' insights into the accuracy of their representations. Previously, we successfully reduced confident errors through a 1.5-hour-long VWM training during which participants received performance feedback based on the accuracy of metacognitive assessment of their VWM representations. Participants remembered a briefly presented array of six colored squares, then reported each item with their confidence in the accuracy of their report. Critically, they received 10, 5, or 0 points for an accurate VWM report coupled with high, low, or no confidence, respectively. Conversely, they lost 10, 5, or 0 points for inaccurate responses coupled with high, low, or no confidence. This training reduced the occurrence of confident errors (i.e., errors coupled with high confidence). Interestingly, training benefits emerged in the first several minutes of training. Thus, in the current study, we tested whether shorter training (10 mins) was sufficient to produce a training benefit that also generalizes across different stimulus types. Specifically, participants (n = 78) performed two VWM tasks (10 mins each) where they remembered an array of six colored squares or oriented bars and reported each item with their confidence (high, low, or no confidence). After measuring baseline performance, participants repeated the color or orientation VWM task with feedback. Participants then repeated the two VWM tasks without feedback to assess the training benefit and its stimulus generalizability. Here, despite a shorter training session, participants reduced confident errors during and after training relative to baseline. This benefit also generalized to the untrained stimulus. Taken together, our results demonstrate rapid and stimulus-general improvement of metacognitive assessment of VWM representations.

Visual Memory: Working memory and encoding, retrieval

53.327 COMPARING NEURAL CORRELATES OF FOVEALLY AND PERIPHERALLY ENCODED MEMORIES

Güven Kandemir[1] (g.kandemir@vu.nl), Chris Olivers[2]; [1]Vrije Universiteit Amsterdam, [2]Vrije Universiteit Amsterdam

Current consensus ascribes visual working memory to the maintenance of the incoming sensory signal through top-down feedback mechanisms, when the stimulus is no longer available. Crucially, we know from perception research that both feedforward and feedback mechanisms favor central over peripheral vision. Yet, it is currently unknown whether and how these biases impact visual working memory processes, such as encoding and maintenance. We, therefore, compared the neural correlates of foveally and peripherally encoded memories of oriented Gabor stimuli presented either at the centre of fixation or left/right-lateralized at 15° eccentricity. Behavioral, EEG and eye data were recorded from 30 participants who completed the task. Multivariate decoding analyses of the EEG data revealed clear dissociations across presentation location (foveal, peripheral) and time window (encoding, early maintenance, late maintenance period). During stimulus encoding, decoding was strong for foveal items, yet untraceable for peripheral items. Orientation information of peripheral items emerged only after stimulus offset, was location-specific, but of comparable strength as for foveal items, suggesting an equivalent involvement of feedback-based maintenance mechanisms for both central and peripheral items. Later during maintenance, the memory signal disappeared from the voltage data, and orientation information gradually emerged in alpha power instead. This transformation also included a conversion from spatially specific neural codes to a spatially generalized format. Moreover, transformed representations were accessible through impulse-driven perturbations, further unveiling the underlying memory state. We conclude that feedback mechanisms do not necessarily suffer from eccentricity. Furthermore, the eccentricity-driven dissociation between disparate sensory and common maintenance representations indicates that storage activity patterns as measured by EEG must reflect signals beyond the primary visual cortex.

53.328 SELF-INITIATING STIMULUS IN SHORT TERM MEMORY: DIFFERENTIAL EFFECTS OF TEMPORAL PREDICTION AND MOTOR CONTROL

Rocio Loyola Navarro[1,2,3] (rocio.loyola@umce.cl), Ricardo Mendoza[3], Pedro Maldonado[3,4]; [1]Universidad de Chile, Centro de Investigación Avanzada en Educación, Instituto de Estudios Avanzados en Educación, [2]Universidad Metropolitana de Ciencias de la Educación, Departamento de Educación Diferencial, Facultad de Filosofía y Educación, [3]Universidad de Chile, Departamento de Neurociencia, Facultad de Medicina, [4]Centro de Inteligencia Artificial (CENIA), Santiago Chile

Classic studies of Short Term Memory (STM) required passively waiting for the stimulus. In contrast, sensory activation is typically the result of self-initiated movements in natural behavior. It has previously shown that self-initiation improves accuracy in an STM task (Loyola et al., 2022). However, it is still unknown whether this effect is related to the temporal predictability or to the activation of motor systems produced during self-initiation (motor control). We hypothesize that both mechanisms have independent effects on STM accuracy. Undergraduate students underwent a modified Stenberg task, which was designed to manipulate the degree of motor control and temporal predictability of the stimulus onset. We created five conditions: Active Predictable (AP), Active Unpredictable (AU), Passive Predictable (PP), Passive Unpredictable (PU) and Motor Only (MO). In the active conditions, the participants had full motor control over the stimulus onset, while in the passive conditions, they had no motor control. The stimulus was presented at a fixed time in the predictable conditions, while in the unpredictable conditions, the stimulus appeared randomly. MO condition consisted of button presses at given times, with no STM task related. Contrary to our previous results, our current results show that the AP condition performs worst (accuracy mean = 0.79). Two-way paired ANOVA yields a significant main effect of predictability on accuracy only ($p = 0.004$). Paired t-tests corrected by Holm method show a significant difference between AP and AU ($p = 0.034$), and between AP and PU ($p = 0.043$). These results suggest that STM is taxed more strongly when subjects control the onset of the stimulus and the time of presentation is predictable, but also indicate that predictability seems to be a major factor in this effect.

ANID FONDECYT postdoc 322004 and Support from ANID/PIA/Basal Funds for Centers of Excellence FB0003 are gratefully acknowledged. Funds from PROYECTO ACE 210007 is gratefully acknowledged

53.329 RECENT STIMULUS HISTORY IMPACTS THE SPEED OF RESPONSE FOR A NEW STIMULUS

Kuo-Wei Chen[1] (kchen172@asu.edu), Gi-Yeul Bae[1]; [1]Arizona State University

Recent stimulus history can automatically bias the perception of a new stimulus. The present study investigated whether this stimulus history effect reflects the biased representation itself or the biased decision during the report. In the experiment, participants performed a motion direction estimation task where they reproduced the perceived motion direction from random-dot-kinematograms (RDKs). We hypothesized that response time (RT) should be faster for the trials with large direction differences from the previous trial if the stimulus history effect is driven by the decision during the report as the decision should be easier when the current motion direction differs significantly from the prior motion direction. Results showed a typical repulsive serial bias effect. More importantly, RT was faster for the trials with large direction differences, supporting the hypothesis that the serial bias is decision-driven. If this RT effect truly reflect the decision related to the serial bias, then the RT effect should be absent when the serial bias is absent. To test this prediction, we had participants perform a 'dual-feature' version of the RDK task where RDK was presented with colored dots and participants reported the color, instead of the direction, on the random half of the trials. We found no serial bias in the direction reports following the color report. Critically, RT effect was absent after the color report, demonstrating that the RT effect was associated with the serial bias. In subsequent analysis for the data with a large sample size (N>70), we confirm that the RT effect was mostly driven by the expedited decision for large direction differences rather than the delayed decision for small direction differences. Together,

these results provide converging evidence that the prior stimulus impacts the speed of decision for a current stimulus, supporting the hypothesis that the stimulus history effect is driven by decisional processes.

53.330 IMPLICIT PRIORITIZATION OF INFORMATION IN VISUAL WORKING MEMORY?

Chris Jungerius[1,2,3] (d.c.jungerius@uva.nl), Freek Van Ede[2,3], Heleen Slagter[2,3]; [1]University of Amsterdam, [2]Vrije Universiteit Amsterdam, [3]Institute Brain and Behaviour Amsterdam

It is well-established that the brain can exploit statistical regularities to selectively guide attention towards task-relevant information in the external world, even without explicit awareness of these regularities. This study examined whether such 'statistical learning' can also guide the selection of previously encoded information held in working memory, using an online experiment. To this end, 69 participants performed a visual working-memory task. In each trial, two tilted colored bars were shown, followed by a sound (low or high in pitch). After a short delay, participants had to either reproduce the orientation of one of the bars (5/6 of trials) or report the pitch of the sound (1/6 of trials). Critically, unbeknownst to the participant, the sound's pitch functioned as an implicit retrocue, predicting which bar was likely to be the target in that trial. A Bayesian hierarchical model showed tentative evidence that participants exploited the predictive cue, as they were slightly faster in reporting validly cued targets over invalidly cued targets, while showing no explicit awareness of the sound-to-memory-item regularity. However, this effect failed to reach significance in an equivalent frequentist model. In a control experiment (n=31) where participants were informed of the contingency beforehand, validly cued targets were reported significantly faster and more precise than invalidly cued targets, and to a much greater extent. These findings suggest that guidance of internal attention by contextual cues may be possible to some extent in the absence of explicit awareness and may be sensitive to implicit statistical learning. We are currently conducting a follow-up experiment (n=65) using implicit temporal expectations instead of an implicit contextual cue to guide internal attention, to further characterize the boundary conditions of implicitly learning to prioritize information in working memory.

53.331 PHANTOM MEMORIES: ENHANCING MEMORY THROUGH COMPETITION WITH IMAGINARY ITEMS

Timothy Brady[1] (timothy.brady@gmail.com), Maria Robinson[1]; [1]University of California, San Diego

Visual working memories are known to be imprecise because perceptual information is corrupted by internal noise during the memory delay. Can memory noise create strong but entirely false memories if no perceptual information is actually present? If so, how do these memories compete with true memories? In a series of two experiments (total N=90), participants were exposed to 2 or 4 colors for a single frame (~16ms), followed by a dynamic noise mask for 200ms and an 800ms delay. Crucially, in half the cases, only 1 item was presented, but the masking and test positions implied the full set of items was present. In Experiment 1, participants were probed about a specific position, which might be one where a color was never presented, and then asked to report their confidence. In Experiment 2, participants were free to choose which position they want to report on half of trials. Replicating prior work, there was a robust confidence-accuracy relationship for items actually shown. However, participants often gave high confidence responses to locations that did not have items, and regularly chose these locations when reporting on their "strongest" memory, indicating a propensity for false memories. Critically, when only a single item was presented, participants' memory was more precise when they freely chose that item as their strongest than when they were forced to report it, even though in both cases they reported on the only actual item presented. Overall, these results imply that the phantom items are genuinely experienced. Furthermore, an actually shown item wins the "competition" against phantom items more often when this item is strongly represented. This provides evidence for the general view that participants are always reporting based on an underlying noisy sensory signal even when they are reporting an item they have no information about ('guessing').

NSF BCS- 2146988 to TFB

53.332 WORKING MEMORY ERRORS FROM DISTRACTION ARE NOT INFLUENCED BY SENSORY NOISE AT ENCODING

Holly Kular[1], John Serences[1]; [1]University of California San Diego

Prior research suggests that ongoing sensory input during the delay period of a visual working memory task interferes with the storage of remembered information. However, it is unclear whether sensory uncertainty during encoding increases susceptibility to interference or whether memories are relatively stable once information is encoded (even if encoding is highly inaccurate). Here, we tested whether distracting stimuli influenced memory accuracy while simultaneously manipulating sensory noise by changing the width of a bandpass filter applied to white noise stimuli to titrate the amount of orientation information in each remembered stimulus. In a behavioral experiment (N=24), participants viewed a phase-reversing orientation grating rendered at one of two levels of sensory noise (high and low) for 500 ms, followed by a 3,000 ms delay with one of three distractor conditions (no distractor, ignore oriented white noise distractor, or attend oriented white noise distractor). At the end of each trial subjects responded by rotating a white line to match the remembered orientation. We found strong evidence that increasing noise in the remembered oriented grating decreased working memory performance (BF10=3.49e12). Notably, the presence of oriented distractors interfered with working memory performance (BF10=72.2), with evidence of memory attraction toward the orientation of the noise distractor. However, there was no interaction between stimulus noise and distractor presence (BF10=0.13). These data provide evidence that while working memory representations are prone to adaptive distortions during delay periods, these distortions are not modulated by sensory noise during encoding.

NEI

53.333 INTER-INDIVIDUAL DEPENDENCY IN VISUAL COGNITION: OBSERVERS ADJUST PERCEPTUAL DECISIONS BASED ON OTHERS' PERCEPTION

Gi-Yeul Bae[1] (gbae2@asu.edu), Kuo-Wei Chen[1]; [1]Arizona State University

Observing how others perceive objects is a fundamental human behavior as it provides valuable information for adapting to new environments. In the present study, we demonstrate that human observers adjust their perceptual decisions based on others' perceptions, even when those perceptions are entirely irrelevant to the ongoing behavioral goal. In Experiment 1, pairs of participants performed a visual working memory task where they remembered the orientation of a simple visual stimulus and reproduced the orientation after a short delay. A response cue, presented shortly after the stimulus, determined which participant should report the orientation (i.e., stimulus ownership) for that trial. Therefore, participants could ignore the stimulus orientation if the cue indicated that their partner should report it. The orientation reports were randomly assigned to each participant, and the orientations between trials were entirely independent. Nevertheless, one participant's orientation reports were systematically biased by the orientation presented in the previous trial, even when their partner had reported the orientation, demonstrating that participants adjusted their perceptual decisions based on their partner's stimulus. This was replicated in Experiment 2 where the stimulus ownership was determined at the time of stimulus presentation so that the encoding of the stimulus into working memory was completely unnecessary for one participant. The same pattern of results was obtained in two follow-up experiments where the stimuli for each participant were spatially separated (i.e., lateral presentation) and where different stimuli (i.e., orientation and location) were used for each participant in the pair. In the last experiment, we found that the reports were systematically biased by the partner's decision even when the partner's stimulus was invisible. Together, these results demonstrate the inter-individual dependency in visual cognition: observers automatically adjust their perceptual decisions based on how others perceive objects even when such decisions are completely irrelevant to the task at hand.

53.334 RECALL REQUIREMENTS CAN DRASTICALLY MODULATE WORKING MEMORY REPRESENTATIONS IN HUMAN VISUAL CORTEX

Giuliana Martinatti Giorjiani[1,2] (giuliana.mg17@gmail.com), Rosanne L. Rademaker[1]; [1]Ernst Strüngmann Institute for Neuroscience in Cooperation with the Max Planck Society, [2]Department of Cognitive Neuroscience, Vrije Universiteit Amsterdam, The Netherlands

Visual working memory (VWM) allows for temporary storage of relevant information to support adaptive behavior. Prior fMRI studies investigating the mechanisms of passive storage have decoded mnemonic information from activity patterns in early visual cortex (EVC). Given that stored information is ultimately used to accomplish specific tasks, might EVC involvement be invoked by anticipated recall requirements? Here, we manipulated recall requirements during an orientation memory task to investigate how adaptive behavior impacts EVC representations. Specifically, we used four recall conditions based on a method-or-adjustment paradigm, as even a simple task of this sort involves a multitude of processes, and allows us to dissociate the roles of visual input, motor output, and online monitoring. To report a remembered orientation, participants (1) used "closed-loop" button presses to rotate a thin dial on the screen, (2) viewed a "matched replay" of a pre-recorded response, and indicated a clockwise or counter-clockwise offset of the final dial orientation, (3) viewed a "mismatched replay", i.e. a pre-recorded response to a random orientation, and pressed a button if the dial crossed the currently remembered orientation, and (4) used button presses to rotate an "invisible dial" with an initial orientation shown only briefly before response onset. We uncovered a drastic increase in memory decoding during recall when the remembered orientation matched the final dial orientation ("closed-loop" and "matched replay" conditions). This implies that visual input alone (a rotating dial with final orientation unrelated to the remembered orientation, as in "mismatch replay"), motor output alone (high during "closed-loop", low during "matched replay"), or online monitoring alone (not required during "exact replay" until the very end of the trial) cannot account for high memory decoding in EVC during recall. Instead, our findings suggest amplified information in EVC when mnemonic contents match sensory input during recall.

53.335 INFLUENCE OF INVISIBLE IMAGES ON REACTION SPEED IN A WORKING MEMORY TASK

Siyun Liu[1] (liusiyun@ibp.ac.cn), Sheng He; [1]Insititute of Biophysics, Chinese Academy of Sciences

Invisible stimuli have the potential to influence cognitive processes without entering conscious awareness. Previous studies suggested that invisible stimuli can even affect the contents of working memory, but the claim of stimulus invisibility requires more vigorous validation. Here we employed a 2-alternative-forced-choice paradigm to validate invisibility and explored how invisible stimuli affect visual working memory. In a working memory task, participants (n=21, 9 males) briefly viewed a sample figure for 0.5 seconds and memorized it. After a two-second blank screen, participants were presented with noise, visible, or invisible interfering figures for four seconds. Invisibility was achieved through continuous flashing suppression (CFS). The number of interfering figures varied from two to four, randomly. While these interference figures shared the sample's layout, they were never identical. Following another two-second blank screen, a probe figure was presented. The participants evaluated and responded as soon as possible whether the probe matched the sample. The probes were different from the interfering figures in all trials. Three validation blocks were conducted to check the invisibility of the suppressed interfering images, with participants making forced-choice judgments about their location. Accuracy remained at chance levels, confirming invisibility. Participants' normalized reaction times in different conditions showed that invisible stimuli increased reaction times compared to noise, particularly when the probe matched the sample ($t(1789) = 2.90$, $p = .003$). The number of stimuli figures did not yield significant differences in reaction times. Our study reveals that unconscious exposure to invisible stimuli sharing the layout as the content in the working memory can delay reaction times in subsequent match-to-sample task. When the invisible interfering pattern differed from the working memory content, participants exhibited slower responses in determining a match between the probe and the sample.

53.336 SERIAL DEPENDENCE AS A MECHANISM INVOLVED IN SLOW CHANGE BLINDNESS

Haley G. Frey[1], David Whitney[1]; [1]University of California, Berkeley

Change blindness is a robust phenomenon where attentive observers fail to notice large changes in a visual stimulus, a surprising limitation of visual processing. These failures range from missing a change that coincides with a "mudsplash" or that occurs between two images separated by a blank screen, to failing to notice that the identity of a stranger has changed mid conversation. Individuals will even miss changes in items they are looking right at when the change takes place gradually, a less well-researched phenomenon called slow change blindness. Though there is still uncertainty in what drives the phenomenon of slow change blindness, one theory for how the brain maintains a stable visual representation across time is serial dependence. This proposed mechanism capitalizes on the assumption of an unchanging world and allows us to retain information about our surroundings across different head, body, and object orientations. Serial dependence predicts that perception at any given moment not only reflects currently available visual information, but also pulls from past visual information to stabilize and smooth perception. Recent work shows that after viewing a slowly aging face, observers rate a face as several years younger than just observing the face without any preceding morph. This illusion of stability indicates that past information can bias perception of a current stimulus but does not characterize or measure the moment-to-moment perception that occurs during a continuously changing stimulus. In this study, we present observers with two slowly oscillating stimuli and ask them to make continuous judgements about their similarity over time. We find that observers perceive two phase-shifted slowly changing stimuli to be more similar than they actually are at any moment. This bias is consistent with serial dependence and may provide evidence for the role of such a stabilization mechanism in slow change blindness.

53.337 UNFOLDING SERIAL DEPENDENCE ACROSS PERCEPTION AND WORKING MEMORY PROCESSES

Hyung-Bum Park[1] (hbpark@uchicago.edu), Edward Awh[1]; [1]University of Chicago

The debate on serial dependence questions its origin in perceptual versus post-perceptual working memory (WM) components, with mixed findings on whether perceptual decisions exhibit attractive serial dependence toward previous target or repulsive bias linked to sensory adaptation. The present study aims to examine the dynamics of serial biases across perception and WM processes at consolidation and retrieval. Participant performed a two-part experiment, with the first part involved immediate continuous estimations of color on color-wheel within view (perceptual report). The second part involved WM recall with dual-response. A memory color followed by masks was reported after stimulus offset (consolidation report), and the same item was reported again after delay on another color-wheel (retrieval report). Mouse trajectories for all responses were recorded to index direction and magnitude of serial bias. Behavioral reports showed a gradual evolution of serial bias across opposite directions: repulsion in perceptual reports, moderate attraction in consolidation, and stronger attraction in retrieval. Critically, we found a unique mouse trajectory pattern in consolidation reports, with a repulsive curvature from the previous target appearing on the color-wheel, while the overall trajectory exhibited attraction bias. Further examination of moment-by-moment trajectories, coupled with a median-split of movement onset latency, captured a 'repulsion-to-attraction' transition in consolidation reports, with early repulsion evolving into attraction as movements progressed. Moreover, the median-split analysis revealed a magnification of bias from early-onset to late-onset trials within their directions (stronger repulsion vs. attraction in perceptual and WM reports, respectively). These suggest that changes in serial biases across processes are not solely time-dependent, but originate from distinct perceptual and WM mechanisms. The repulsion-to-attraction transition during the WM consolidation phase offers an intermediate window into the interplay between sensory adaptation and serial dependence, jointly shaping serial bias in behavioral reports. Together, our study provides evidence supporting the mnemonic origin of serial dependence.

53.338 PERCEPTUAL CONTEXT BOUNDARIES CREATE SEPARATE EVENTS IN VISUAL WORKING MEMORY

Yuanxiu Zhao[1] (yuanxiu.z@zju.edu.cn), Entong Jin[1], Chi Zhang[1], Mowei Shen[1], Zaifeng Gao[1]; [1]Zhejiang University

In daily life, individuals perceive a continuous stream of external information yet spontaneously segment the ongoing sensorimotor flow into describable, memorable, and discrete units. These event boundaries, occurring during context shifts (e.g., environmental changes), extensively influence subsequent cognitive activities. However, little is known about how events are formed and stored in VWM. This study investigates event-level organization in Visual Working Memory (VWM). Experiment 1 explores whether perceptual context shifts create independent events in VWM. To mimic that our vision captures the transient inputs from the ever-changing world, we asked participants to memorize two sets of items (colors and shapes) separated by a context shift (i.e., the change of both presentation mode and stimuli category), with fixed memory load for one event (e.g., 3 colors) and manipulated load for the other (e.g., 1/2/3 shapes). Results indicate independent storage of two events, that is, the memory load of one event did not affect the memory performance of the other event. Experiment 2, without event boundaries (i.e., color and shape appeared in a chaotic order), refutes independent color and shape storage. Experiment 3, which employs solely the background color change as the event boundary, confirms independent storage between shape events. Experiment 4 assesses the effect of perceptual context boundaries on temporal order memory. The temporal order memory test incorporated two shapes selected from either the same event or two adjacent events for recency discrimination. The results showed that participants had worse temporal memory for information spanning event boundaries than information from the same event, providing additional evidence for event structure presence. These findings align with the notion that across-event elements have weaker associations than within-event ones. In conclusion, context shifts during encoding shape subsequent VWM organization, with temporally close but perceptually divided events stored separately.

TUESDAY, MAY 21, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Visual Memory: Working memory and behavior, models

53.339 ASSESSING INDIVIDUAL DIFFERENCES IN CHUNKING STRATEGY IN VISUAL WORKING MEMORY

Yin-ting Lin[1] (lin.3913@osu.edu), Andrew B. Leber[1]; [1]The Ohio State University

Visual working memory is highly limited, and there are large individual differences in this limitation. Many studies suggest that individual differences in visual working memory are related to storage capacity and attentional control abilities. However, recent work has proposed that these individual differences may also reflect how efficiently individuals encode information into working memory (e.g., Nassar et al., 2018). Here we ask whether individuals with high visual working memory capacity are more likely to use chunking strategies to optimize encoding of information. In a continuous report task, participants viewed a display containing six colours and subsequently reported the colour of three items on a colour wheel. Critically, we showed participants the same sets of displays, and participants were free to choose which three items they report on each trial. This allows us to examine the degree to which participants selectively chunk and report similar items for each display. We also independently measured individuals' working memory capacity in a change localization task. Here, we computed within-cluster variance in order to quantify the colour similarity for each display (i.e., the extent that the colours in the display can be chunked). Results show that memory performance is better for displays with high colour similarity, replicating previous work (Nassar et al., 2018). Furthermore, we assessed the extent that individuals use chunking strategies by examining the colour similarity of items that individuals chose to report on each trial. However, we found no evidence that individuals with high capacity were more likely to report items sharing greater colour similarity. Taken together, these findings suggest that while chunking strategies improve performance, individual differences in working memory do not seem to predict the use of chunking strategies. Future work will examine the impact of working memory load on the use of chunking strategies across individuals.

This work is supported by NSF BCS-2021038 to ABL

53.340 EFFECTIVE DISTRIBUTION OF VWM RESOURCES DOES NOT DEPEND ON VWM CAPACITY.

Lyric R. Ransom[1] (ransom.120@osu.edu), Yin-ting Lin[1], Julie D. Golomb[1], Blaire Dube[1,2]; [1]The Ohio State University, [2]Memorial University of Newfoundland

Attention serves as a filter to capacity-limited visual working memory (VWM), ensuring that irrelevant information is not encoded. Dube et al. (2017) suggested that this attentional filter also regulates the distribution of VWM resources, ensuring the most relevant items are encoded with the greatest precision (the Filter and Distribute account). There are individual differences in VWM capacity, and high- and low-capacity individuals differ in their ability to filter distraction (Vogel et al., 2005). Here we examine whether high- and low-capacity individuals also differ in their ability to flexibly distribute VWM resources. We first used a change localization task to measure VWM capacity: participants viewed an array of colored squares and identified the item that changed color when the array reappeared. We separated participants into high- and low-capacity groups using a median split on capacity estimates. Next, participants viewed four colored shapes (two circles/two squares) before reporting the color of a probed shape in a continuous report task. We manipulated the likelihood that a square (or circle) would be probed (target shape counterbalanced across participants), such that the probed item was 60%, 70%, 80%, or 90% likely to be the pre-designated target shape (blocked conditions). We observed flexible resource distribution in both VWM groups: the precision of color report increased with increasing probe probability. Unlike the ability to filter out distraction, our results suggest that low-capacity VWM individuals do not show reduced ability to flexibly distribute resources in VWM. Thus, counter to the suggestion made by the Filter and Distribute account, the ability to filter information in/out of VWM and the ability to flexibly distribute resources among encoded information may be supported by distinct mechanisms.

This project is supported by NIH R01-EY025648 (JG), NSF 1848939 (JG), and NSERC PDF.

53.341 VISUAL FIELD MAP SIZE PREDICTS SPATIAL WORKING MEMORY PERFORMANCE

Nathan Tardiff[1] (nt2487@nyu.edu), Xingyu Ding[1], Xiao-Jing Wang[1], Clayton Curtis[1]; [1]New York University

Working memory (WM) performance demonstrates substantial individual variation. This variability is an important predictor of cognitive functioning and real-world outcomes, including intelligence and psychiatric disorders. Despite its importance, the neurobiological substrates of individual differences in WM are unknown. One possibility is that individual differences stem from basic properties of the neural populations supporting WM. In perception, for example, differences in psychophysical performance correlate with differences in the surface area of early visual cortex (e.g., Song et al., 2013; Himmelberg et al., 2022). In this study, we used functional neuroimaging (fMRI) to test the hypothesis that the size of visual field maps in frontal and parietal cortex predict individual differences in WM precision. To bridge behavioral and neural variability, we implemented neural network models to identify mechanisms by which neural population size could affect WM performance. Human subjects (male and female) underwent population receptive field mapping to identify retinotopically-organized regions of visual, parietal, and frontal cortex, using fMRI. Separately, we assayed subjects' WM using a memory-guided saccade task. We then correlated the size of subjects' visual field maps with their average WM error. Confirming our hypothesis, we found significant negative correlations between the size of visual field maps and WM error in regions along the precentral and intraparietal sulci. We used a neural network model of WM to explore how size improves WM precision. We asked whether larger size: 1) makes WM representations more resilient to noise; 2) allows greater averaging over noise during readout; 3) increases encoding precision via finer tuning of units across stimulus space. We found that both resilience to noise and improved readout contributed to size effects. In sum, our findings identify a mechanistic basis for individual differences in WM and demonstrate the power of combining individual differences with computational modeling for uncovering basic cognitive mechanisms.

NIH R01 EY-016407; NIH R01 EY-033925; NIH 5T32EY007136-30

53.342 ATTRACTIVE AND REPULSIVE BIASES IN A FRAMEWORK OF CONCEPT GENERALIZATION

Isabella DeStefano[1] (idestefa@ucsd.edu), Timothy Brady[1]; [1]University of California San Diego

When multiple stimuli are perceived or remembered, separated in space or time, reproductions of one item tend to be biased towards or away from the other item(s). Previous work has proposed that these biases could arise from assuming that items belong to the same or different objects (Chetverikov, 2023). Here we use a bayesian model of concept generalization (Tenenbaum & Griffiths, 2001) in conjunction with a signal detection framework (Schurgin et al., 2020) to demonstrate that attractive and repulsive biases arise as a natural consequence of assuming items belong to either the same or a different 'concept'. Under this model, concepts provide groupings over stimuli (i.e., for numbers, concepts might be "20-30"); generalization is governed by how likely stimuli are to arise from the same underlying concept; and similarity between stimuli arises from generalization. When concepts are assumed to reflect ranges of different sizes across stimuli space (as in Shepard, 1989), this model accurately predicts full error distributions. When a same/different assumption is incorporated into this model, unique patterns of responses arise: for repulsion, responses are skewed away from the other stimulus with the true stimulus as the median; for attraction, responses are symmetric with the median shifted away from the true stimulus. The magnitude of bias can be modulated via manipulating the maximum ranges of concepts considered. Additionally, by defining heterogeneous concepts (e.g., for color, greater ranges within yellows than across the yellow-orange boundary), stimulus-specific biases can emerge. This model suggests that attractive and repulsive biases are not unique to vision, but instead apply broadly; consistent with this, we provide preliminary empirical evidence of conceptual rule-based repulsion using symbolic number. Overall, understanding this putatively visual phenomena under a framework of concept generalization broadens our understanding of the cause and nature of such biases.

53.343 A PROCESS MODEL OF ENSEMBLE CODING WITHIN WORKING MEMORY

Jacob Zepp[1] (jacobzepp@usf.edu), Chad Dubé[1]; [1]University of South Florida

The cognitive science literature has recently noted a large body of effects attributed to ensemble codes, or the ability to integrate environmental information into summary representations. While the prevalence of summary effects has been well documented, a description of the process by which the ensemble code arises within the more general visual information processing stream is still missing. Within the current work, a description of such a process is given. The ensemble code forms over information that has been encoded into working memory, forming an approximation of the external environment from a fidelity-weighted integrated representation. The nature of the information that contributes to the summary is vulnerable to bottom-up perceptual effects as well as goal-driven top-down effects. The resulting model gives a description of how the wide range of newly noted summary effects arise within the wider systems of memory and perceptual processing.

53.344 ENCODING MODELLING FOR WORKING MEMORY RESEARCH: PATTERN SIMILARITY, REPRESENTATIONAL GEOMETRY, AND MODEL COMPARISON

Thomas B. Christophel[1,2] (tbchristophel@gmail.com), Andreea-Maria Gui[1,2], Carsten Allefeld[3], José M. Baldaque[1,4], Joana Pereira Seabra[1,2]; [1]Department of Psychology, Humboldt Universität zu Berlin, Rudower Chaussee 18, Berlin, 12489, Germany, [2]Bernstein Center for Computational Neuroscience and Berlin Center for Advanced Neuroimaging, Charité Universitätsmedizin, corporate member of Freie Universität Berlin, Humboldt Universität zu Berlin, and Berlin Institute of Health, Berlin, Philippstraße 13, Haus 6, 10115, Germany, [3]Department of Psychology, City, University of London, London EC1V 0HB, United Kingdom, [4]School of Psychology & Neuroscience, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom

Research into the neural basis of working memory relies heavily on the use of multivariate decoding techniques to ascertain the presence of neural representations of working memory content. Recent work attempts to understand the tuning properties underlying cortical representations using encoding modelling. Encoding modelling aims to explain the multivariate patterns underlying mnemonic function using stimulus dependent regressors called basis functions. These basis functions are capable of explicitly modelling the similarity relationships between different stimuli analogous to tuning functions. Here, we aim to evaluate the use of encoding modelling for the quantification of mnemonic representations using simulated and real data. We make use of a recently developed flexible simulation toolbox to simulate patterned neural activity generated using different underlying voxel tuning distributions across a large number of possible experimental designs (designSim). We quantify the information content of these neural signals as the variance explained by a given encoding model using cvCrossMANOVA. We demonstrate that realistically modelled neural representation can be more reliably identified using encoding models with realistic similarity assumptions as compared to a simplistic classifier-like models. We show that estimates of representational similarity between two conditions (e.g., cross-classification accuracy, correlation, and variance explained between conditions) are strongly biased by the signal-to-noise ratio of individual representations and provide an SNR-independent measure of pattern similarity by comparing variance explained within and between conditions. Finally, we ask whether encoding modelling can ascertain the precise representational code used for memorization. We show that, consistent with prior work, stimulus-driven representation of memorized contents can be fitted and explained using a large variety of differently shaped encoding models such. This means that a reliable fit alone gives little insight into the representational geometry ('feature fallacy'). Instead, we demonstrate that comparing the explained variance of two or more competing models allows to reliably identify the true model.

Tuesday, May 21, 8:30 am – 12:30 pm, Banyan Breezeway

Plasticity and Learning: Properties

53.345 EXPOSURE-BASED LEARNING IMPROVED ORIENTATION DISCRIMINATION UNDER VISUAL CROWDING

Yanru Chen[1] (2101110693@stu.pku.edu.cn), Kunyun Wang[2], Cong Yu[3], Junyun Zhang[4]; [1]Peking University

Visual crowding refers to the impairment of object recognition in the presence of adjacent objects. Perceptual learning reduces peripheral crowded orientation discrimination and learning shows specificity to trained locations. Here we manipulate attention to crowded stimuli, to separate the impacts of top-down attention and bottom-up exposure, on learning with crowded orientation and its transfer to other locations. Observers reported the target orientation (a circular sinusoidal grating centered at 8°-eccentricity, 36°/126°) with two grating flankers with randomized orientation in pre-test and post-test. The orientation discrimination threshold was adjusted by staircase. Four groups of observers underwent five sessions of training or exposure. Results: (1) Baseline-group (N=8): Training improved crowded orientation discrimination, and the reduction of crowding was specific to the trained location. (2) Active-exposure-group (N=8): Crowded contrast discrimination training enabled complete learning transfer to crowded orientation discrimination, and the transfer was specific to the exposed location. (3) Passive-exposure-group (N=8): Observers responded to a central RSVP task while passively exposed to peripheral crowded gratings. Crowded orientation discrimination was substantially improved as continued training produced no further gains, and the reduction of crowding was evident in unexposed locations. (4) Subliminal-exposure-group (N=8): A "continuous-flash-suppression" technique was used to suppress the exposure of crowded gratings into sub-consciousness, meanwhile observers were asked to do a fovea dot color task. Crowded orientation discrimination was mostly improved, and the improvements partially transferred to unexposed locations. (5) A control group (N=9) ruled out the possibility that the improvements were due to the test-retest effect. The results demonstrated the capacity of the visual system to learn to reduce crowding by repeated exposure to crowded stimuli, which provides a complementary of plasticity to practice, attention-based, learning. Releasing spatial attention to crowded stimuli might decrease the location specificity in crowding learning. These findings shed new light on the mechanisms of crowding and learning.

Supported by Natural Science Foundation of China Grants 32371097 (JYZ).

53.346 TRAINING IN A CROWDING TASK AND POSSIBLE TRANSFER EFFECTS TO ECCENTRIC VISUAL ACUITY, CONTRAST SENSITIVITY, WORD RECOGNITION AND READING

Elena von Perponcher[1] (elena.von-perponcher@ur.de), Niria Krüger[1], Anna Mayr[1], Katja Prölß[1], Mark W. Greenlee[1], Tina Plank[1]; [1]Institute for Experimental Psychology, University of Regensburg

Visual crowding, the difficulty in recognizing objects in the periphery amid distractors, can be reduced by perceptual learning. However, it is uncertain how much this improvement transfers to other visual tasks. We examined whether reducing the crowding effect in the peripheral visual field through training transfers to the performance in a word recognition (WR) task (lexical decision task), a rapid serial visual presentation (RSVP) task (reading words that form sentences) and on the FrACT test (four visual acuity and three contrast sensitivity tasks), in three normally sighted samples (N1=36, N2=36, N3=42, respectively). Every sample was randomly assigned to either a training or a control group. The training groups performed a Landolt-C gap detection training daily over 4 days at 10° eccentricity in all four quadrants. A 2-down, 1-up adaptive procedure adjusted the critical spatial distance of two ring-shaped flanker distractors positioned radially and tangentially to the Landolt-C with respect to central fixation. Pre- and post-training, both trained and control (untrained) groups performed the WR (sample 1), RSVP (sample 2) or FrACT tasks (sample 3) at the same positions at 10° eccentricity. Each training group showed a significant ($p < .05$) decrease in peripheral visual crowding following training. Preliminary results suggest significant ($p < .05$) improvements in WR and RSVP task performance for the training groups, while no such improvements were observed in the control groups, indicating successful transfer of crowding training. In the FrACT test, on average visual acuity and contrast sensitivity both showed improvements, but no significant differences emerged between training and control groups. The results suggest that training on a Landolt-C crowding task can enhance performance in tasks directly impacted by crowding effects, such as WR and RSVP tasks, but may not extend to more basic single visual features like visual acuity or contrast sensitivity.

This research was supported by the Deutsche Forschungsgemeinschaft (DFG: PL641-1/1) and the ZSER e.V. (Support Association for Children and Adults with Rare Diseases e.V. c/o University Hospital Regensburg) funding program.

53.347 WHEN DOES OBJECT FAMILIARITY LEAD TO A TASK ADVANTAGE?

JHILIK DAS[1] (jhilik.das.94@gmail.com), SP Arun[1]; [1]Indian Institute of Science

While it is well known that humans quickly become familiar with items that are repeatedly viewed, it is unclear whether this familiarity confers any task advantage. With extreme familiarity, such as in the case of reading, readers discriminate between letters of their own scripts better than those of unknown scripts. However, it is unclear at what time scales such effects develop, and whether these effects are dependent on the nature of experience with letters. Here, we investigated whether short-term familiarity with novel shapes can lead to improved shape discrimination. We performed five experiments. In each experiment, we trained two groups of naïve participants on letters or bigrams of two novel scripts. We also conducted familiarity tests with varying difficulty, at various time points, to assess their familiarity to the viewed items. All participants were highly accurate on the familiarity test (with even as few as 100 exposures), even when they had to identify a familiar letter compared to a slight modification of the letter, and even when tested two months after training. Thus, familiarity was quick, robust, and long-lasting. However, does this familiarization produce a task advantage? In Experiments 1 & 2, participants discriminated familiar letters faster only on hard trials of a visual search task. By contrast, in Experiments 3 & 4, when letters were dissimilar to each other, familiarizing them did not produce such a task advantage. In Experiment 5, when we created hard visual search trials by embedding familiarized letters against slightly modified distractors, participants were more accurate on familiar compared to unfamiliar letters both in a visual search task as well as a same-different task. Taken together, our results show that, while the effects of familiarity are fast and robust, this familiarity leads to a task advantage only for highly similar shapes.

This research was funded through a Senior Fellowship from the DBT-Wellcome India Alliance (Grant # IA/S/17/1/503081) and the DBT-IISc partnership programme (both to SPA).

53.348 COMPARING AUDITORY AND VISUAL CATEGORY LEARNING

Casey L. Roark[1] (casey.roark@unh.edu); [1]University of New Hampshire

Introduction: Categorization is a fundamental skill that spans the senses. Categories enable quick identification of visual objects in our surroundings and phonemes and words in spoken speech. While categories are ubiquitous across modalities, the amodal and modality-specific mechanisms of perceptual category learning are not well understood. I investigated learning of artificial auditory and visual categories that shared a higher-level unidimensional rule structure. If learners build amodal category representations, they should benefit from simultaneous learning of categories from different modalities that share a higher-level structure. If learners build representations separately across modalities, their learning should either be unaffected or impaired by simultaneously learning categories from different modalities. Methods: Learners were randomly assigned to learn two auditory and two visual categories either simultaneously (interleaved) or separately (blocked). The higher-level category structure was the same across modalities – learning required selective attention to one dimension (temporal modulation, spatial frequency) while ignoring a category-irrelevant dimension (spectral modulation, orientation). After 400 training trials (interleaved: auditory and visual together; blocked: auditory then visual or vice versa), participants completed two separate generalization test blocks for both modalities (counterbalanced order). Results: When learning categories separately, accuracies were no different across modalities, indicating that the categories were well-matched for difficulty. When learning categories simultaneously, learners were significantly more accurate for visual than auditory categories. Importantly, there were no significant differences in test performance across blocked and interleaved training conditions in either modality. Conclusion: These results indicate that learners build separate, modality-specific representations even when learning auditory and visual categories simultaneously. Further, learners do not exploit the shared amodal structure of categories across modalities to facilitate learning. These results have important implications for understanding learning of real-world categories, which are often multimodal, and highlights the importance of considering the role of modality in models of category learning.

53.349 LEARNING RELATIONAL CATEGORIES THROUGH GUIDED COMPARISONS

Andrew Jun Lee[1] (andrewlee0@g.ucla.edu), Hongjing Lu[1], Keith Holyoak[1]; [1]University of California, Los Angeles

Visual scenes are not perceived as simple constellations of objects, but rather as objects in relation to one another. Humans can efficiently learn visual categories based on relational knowledge from just a handful of examples; however, the learning mechanisms remain unclear. Here, we hypothesize that analogical comparisons can facilitate learning of visual categories defined by relations. Method: We examined learning using the Synthetic Visual Reasoning Test (SVRT), a collection of 23 relational category learning problems (Fleuret et al., 2011). Each problem consists of images involving artificial-island-shaped objects; positive image exemplars instantiated a rule based on spatial relations and negative exemplars did not. Participants categorized each successive test image into the correct set until an accuracy criterion was met. Feedback was provided on each trial. We conducted two experiments that varied the display format and coloring scheme for the SVRT images. In both experiments, images from previous trials were displayed on the screen as a visual record. In Experiment 1, these record images were either spatially segregated or intermixed by category membership. In Experiment 2, the record images were colored in a way that differentiated object entities based on relations. The colored display of the record images thus guided analogical comparisons between them. Results: Learning was more efficient when prior images in the display were spatially segregated by category membership, resulting in an average 53% reduction in proportion of SVRT problem failures. Furthermore, when objects were assigned with corresponding colors to facilitate the alignment of related objects across images, learning was more efficient relative to the uncolored condition (33% reduction in failure proportion). Conclusion: Human learning of visual relational categories depends on the ability to efficiently extract relational knowledge from visual inputs. Visual displays that facilitate relation extraction promote learning on the basis of analogical comparisons.

NSF Grant IIS-1956441

53.350 VISUAL ADAPTATION IS MORE PRONOUNCED AT THE HORIZONTAL THAN VERTICAL MERIDIAN

Hsing-Hao Lee[1] (hsinghaolee@nyu.edu), Marisa Carrasco[1]; [1]New York University

[BACKGROUND] Contrast sensitivity varies with polar angle (i.e., performance fields); it is better along the horizontal than vertical (horizontal-vertical anisotropy, HVA), and along the lower than upper vertical meridian (vertical meridian asymmetry, VMA). Adaptation, which suppresses contrast sensitivity, is more pronounced at peripheral than parafoveal locations. Here we investigate whether contrast adaptation similarly decreases sensitivity at cardinal locations, diminishes or exacerbates the performance fields. [METHODS] Observers performed a 2-AFC orientation discrimination task, with and without adaptation. The adaptor was a 5-cpd horizontally- or vertically-oriented Gabor patch (100% contrast) presented at the left or right horizontal meridian, and upper- or lower-vertical meridian at 8° eccentricity. Observers determined whether a 5-cpd Gabor patch (4° diameter) was tilted clockwise or counterclockwise compared to a horizontal (Exp.1) or vertical (Exp.2) reference. The contrast of the target was titrated for 75% accuracy with 4 randomized independent staircases at each location. [RESULTS] In the non-adapted condition, contrast threshold was higher at the vertical than horizontal meridian, and at the upper than lower meridian, consistent with the typical HVA and VMA. Critically, in both experiments, the adaptation effect [(adapted - non-adapted threshold) / (adapted + non-adapted threshold)] was more pronounced (contrast sensitivity was more suppressed) at the horizontal than vertical meridian, and there was no difference between the upper and lower meridian. [CONCLUSION] This study reveals that the extent of contrast adaptation varies around the visual field, with stronger adaptation effects at the horizontal than vertical meridian, regardless of the stimulus orientation. This effect may be due to the larger cortical area corresponding to the horizontal than vertical meridian. This difference in the adaptation effect diminishes performance asymmetries, rendering visibility to be more homogenous around the visual field.

Funding: NIH R01-EY027401 to M.C.

Binocular Vision: Clinical

53.351 DIFFERENCES IN BI-STABLE PERCEPTION ACROSS THREE PARADIGMS IN PEOPLE WITH SCHIZOPHRENIA

Kyle W. Killebrew[1] (kkillebr@umn.edu), Hannah R. Moser[1], Scott R. Sponheim[2,1], Michael-Paul Schallmo[1]; [1]Department of Psychiatry and Behavioral Sciences, University of Minnesota, [2]Veterans Affairs Medical Center, Minneapolis, MN

People with schizophrenia (SZ) experience psychotic symptoms that include hallucinations and distorted perception of real stimuli. Bi-stabile perception occurs when the same physical stimulus yields two alternating percepts that compete for dominance. Previous work from our group and others has demonstrated abnormal bi-stable switch rates in SZ compared to healthy controls. Surprisingly, differences in bi-stable switch rates between groups may vary across bi-stable paradigms. Previous reports have found slower switch rates during binocular rivalry (BR) in SZ, whereas we have recently found faster switch rates during a bi-stable structure-from-motion (SFM) paradigm. In an ongoing study, we examined bi-stable switch rates using three different paradigms: BR, SFM (i.e., the rotating cylinder), and biological motion (BM; i.e., ambiguous point-light-walker), in 23 SZ participants and 16 healthy controls. Preliminary results suggest a trend toward faster switch rates among SZ participants vs. healthy controls during the BM task. For SFM, switch rates were also numerically higher on average in SZ vs. controls, but this difference was not significant. However, during the BR task there was a trend for slower switch rates in SZ vs. healthy controls, consistent with results in the literature. Additionally, we did not see significant differences in bi-stable switch rates among people with bipolar disorder (19) vs. controls. Our results suggest that atypical perception of bi-stable stimuli in SZ may not be uniform across paradigms. Switch rates appeared slower in paradigms involving inter-ocular suppression (i.e., BR), whereas bi-stable motion paradigms tended to evoke faster switch rates in SZ, which may indicate multiple abnormalities in SZ at distinct levels of visual processing.

I01 CX001843A

53.352 HEAD-GAZE RELATED INFLUENCES ON HETEROPHORIAS REVEALED BY AN HMD-BASED COVER TEST

Federico Ferracini[1], Samuel Smithers[2], Francesca Peveri[1], Andrea Canessa[1], Agostino Gibaldi[3], Peter Bex[2], Silvio P. Sabatini[1]; [1]University of Genoa, [2]Northeastern University, [3]Magic Leap Inc.

Strabismus is a disorder in which the visual axes of the eyes are misaligned and affects 3-5 percent of children; it can be latent or visible and, if visible, it may be persistent or intermittent. Angle of deviation may change at different gaze positions for incomitant (non-comitant) strabismus or may not change for comitant (concomitant) strabismus, which accounts for most forms of childhood strabismus. Common clinical practice prioritizes the primary position of gaze and a fixed distance, since testing other gaze directions and variable distances would require a longer and more difficult procedure. Employing Head-Mounted Displays (HMD) with embedded eye-trackers in clinical practice would allow to generalize standard procedures, like the cover test, to eventually quantify strabismus angle in natural conditions (free head movement). To this end an HTC Vive Pro Eye was used to administer an alternating cover test: subjects begin by fixating at a 1.6° visual angle grey cross with a rotating grey ellipse at its center, placed 40 cm within the virtual HMD environment. The initial binocular fixation established a 2s baseline, followed by monocular covering, virtually occluding each eye view for 2.25s, alternating for 10 trials each lasting 8s. Specifically, in this work we investigate the extent of heterophorias under different head rotations (45° of yaw and 45° of pitch). The ocular deviation is quantified in ten subjects who have never been diagnosed with strabismus. We conducted an analysis to explore potential trends associated with head rotations. Results revealed patterns of deviation across different head rotations, independently of the specific direction of rotation. Individual differences suggest complex relationships between head movements and ocular alignment. Preliminary results show heterogeneous behaviors for different eye-head relative positions, which would pave the way to a systematic assessment of multiple and varying parameters captured under complex, ecological conditions.

53.353 COARSE STEREOPSIS AND EYE ALIGNMENT IN STRABISMUS

Akosua Kesewah Asare[1] (kesewah@student.ubc.ca), Kimberly Meier[2], Christopher Lyons[1], Conor Mulholland[1], Laurie M Wilcox[3], Deborah Giaschi[1]; [1]University of British Columbia, [2]University of Houston, [3]York University

Depth order judgements are supported by both fused (fine) and diplopic (coarse) binocular disparities. There is evidence that they are subserved by distinct mechanisms, for instance, coarse stereopsis matures earlier than fine and is spared in some types of amblyopia whereas fine stereopsis is disrupted. We hypothesize that coarse stereopsis may be used during development to achieve binocular fusion, align the eyes, and promote fine stereopsis. Here we assess this possibility in children receiving eye alignment surgery for strabismus. We tested 14 children pre-surgery, then 3 and 12 months post-surgery. Stereoscopic stimuli were viewed through liquid crystal shutter glasses. Depth-order discrimination was measured for greyscale cartoon characters, relative to a zero-disparity reference frame. We tested 5 disparities characterized as fine (0.17, 0.67 degrees) or coarse (2.0, 2.5, 3.0 degrees) based on previous work. Aged-matched norms from a previous study with the same stimuli and test paradigm were used to identify atypical performance (>2 SD) in strabismic children. Fusion and interocular suppression were measured using the Worth-4-Dot test and a dichoptic eye chart, respectively. Eye alignment (prism dioptres) was measured with a prism cover test. We found that better eye alignment post-surgery was associated with higher coarse disparity accuracy ($r = -0.7$, $p = .02$), but not fine ($r = 0.2$, $p = .6$), pre-surgery. Pre-surgery performance was atypical for fine disparities for all participants, but typical for coarse disparities for 86% of participants. Fine stereopsis remained atypical for all except 1 participant at 3 months post-surgery, but improved beyond test-retest variability in 73% of participants (all except 1 with typical coarse stereopsis) by 12 months post-surgery. Five participants (all exotropia) achieved typical fine stereopsis, and three of these exhibited typical sensory fusion and interocular suppression. Our results support the hypothesis that coarse stereopsis facilitates eye alignment and the subsequent development of high-resolution stereoscopic mechanisms.

53.354 TRAINING AT-HOME ON A DICHOPTIC READING APPLICATION IMPROVES VISUAL FUNCTION IN ADULTS WITH AMBLYOPIA

Nicole A. Dranitsaris[1,2] (nicoledranitsaris16@gmail.com), Ken Chong[2], Robert F. Hess[1,2], Alexandre Reynaud[1,2]; [1]McGill University, [2]Research Institute of the McGill University Health Centre

Recent research on treatments for amblyopia has shifted its focus from conventional patching, which is only applicable in childhood, to exploring dichoptic tasks. This study leveraged a new approach using an important daily task, reading, to improve amblyopic vision. Here, we assessed if training at-home on a dichoptic E-book application can be an alternative treatment for binocular vision in amblyopia. The dichoptic reading application (DEBRA) was uploaded onto tablets displaying E-books in red/green/black presentation, with each word of text being one of the three colors. By using anaglyph red/green glasses different text could be shown to each eye simultaneously, forcing the individual to combine the input from both eyes. At an initial visit, adult amblyopic participants were given an ophthalmic assessment, then reading speed and eye movements patterns while reading were recorded. Next, participants brought the technology home and trained for one hour per day across two weeks. At the outcome visit they were reassessed on the tests from the initial visit and completed a visual comfort questionnaire. Preliminary results demonstrated improved visual acuity, and contrast sensitivity in most participants. There was variability in task compliance, with some participants being able to easily read everyday for an hour per day and others having more difficulty following this training protocol. Based on the visual comfort questionnaire responses, majority of participants did not experience visual discomfort while completing the dichoptic training for two weeks. Overall, this preliminary study demonstrated that daily training on a dichoptic reading application at home for two weeks can improve amblyopic visual function. More data will clarify if eye movement patterns and other altered ophthalmic factors in amblyopia can be treated by completing the training. Future steps are aimed at collecting more data from amblyopes and ameliorating the user-friendliness of the application.

This project was funded by Vision Health Research Network and Natural Science and Engineering Research Council of Canada grants provided to Nicole Dranitsaris as well as the Trottier-Webster Research Innovation Fund, Pediatric Research Foundation, and IGNTE grants provided to Alexandre Reynaud.

53.355 MODELING RETINOTOPIC MAPS IN AMBLYOPIA REVEALS CORTICAL REORGANIZATION ACROSS THE VISUAL HIERARCHY

Martin Szinte[1] (martin.szinte@gmail.com), Uriel Lascombes[1], Yasha Sheynin[2,3], Dennis M. Levi[2], Michael A. Silver[2,4,5], Adrien Chopin[2,6,7,8]; [1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Université, Marseille, France, [2]School of Optometry, University of California, Berkeley, CA, United States, [3]McGill Vision Research Unit, McGill University, Montréal, QC, Canada, [4]Helen Wills Neuroscience Institute, University of California, Berkeley, CA, United States, [5]Vision Science Graduate Group, University of California, Berkeley, CA, United States, [6]Département d'Etudes Cognitives, Ecole Normale Supérieure, Paris, France, [7]The Smith-Kettlewell Eye Research Institute, San Francisco, CA, United States, [8]Institut de la Vision, CNRS, INSERM, Sorbonne Université, Paris, France.

Amblyopia is a common cortical developmental disorder affecting 3.7% to 5% of the adult population. It is defined as a monocular visual acuity impairment despite healthy or corrected-to-normal optical components. Deficits associated with amblyopia include reduced contrast sensitivity, impaired foveal localization, and stereoblindness. While traditionally viewed as a developmental cortical pathology given the preserved ocular function, the neural bases of amblyopia remain debated. This study aims at understanding the neural mechanisms underlying amblyopia by modeling fMRI data collected during retinotopic mapping in strabismic, anisometropic, and mixed amblyopia patients. Population receptive field (pRF) parameters were extracted across early and intermediate visual areas and compared between amblyopic patients and controls. We found that amblyopic patients systematically exhibited larger pRF sizes across all visual areas examined, with greater effect sizes from V1 to V3 and in LO/VO. Effect sizes were stronger in foveal stimulation than in the periphery, a result compatible with disordered cortical projection theories of amblyopia. Anisometropic and mixed amblyopia patients showed larger pRF sizes than strabismic patients, unexpectedly given their clinical symptoms. These results advance our understanding of amblyopia's neural underpinnings. Foveal-peripheral differences point to abnormal cortical projections. Moreover, the increased pRF sizes, mismatching clinically-predicted patterns, suggests a more widespread plastic dysfunction. Overall, our findings point to a substantial receptive field reorganization in amblyopia, mostly in the early visual system.

This research was supported by an ANR grant to M.S., by the grants ANR-18-CHIN-0002, ANR-18-IAHU-01, ANR-10-LABX-65 and the Marie-Curie fellowship PIOF-GA-2012-327905 to AC, and by RO1EY020976 grant to DL.

53.356 INTEROCULAR CONTRAST SUPPRESSION PARTIALLY ACCOUNTS FOR IMPAIRED STEREOPSIS IN THE CENTRAL VISUAL FIELD OF INDIVIDUALS WITH AND WITHOUT AMBLYOPIA

Preeti Verghese[1] (preeti@ski.org), Aidan Gauper[1], Ruhika Nandy[1], Angela Gomes Tomaz[2], Dennis M. Levi[2], Adrien Chopin[1]; [1]The Smith-Kettlewell Eye Research Institute, [2]UC Berkeley School of Optometry

Previously we mapped stereopsis across the visual field and showed that local stereopsis is impaired in the central visual field for individuals with stereo deficits due to anisometropic amblyopia or microstrabismus. In light of studies that have demonstrated a correlation between suppression and stereopsis in normal vision and in amblyopia, we set out to examine whether visual field loci with impaired stereopsis also have increased interocular suppression. We used similar stimuli for both stereopsis and suppression tasks. A full-screen random-dot stimulus was presented dichoptically with a square target region whose size was m-scaled with eccentricity. We tested target locations along the horizontal and vertical meridians at eccentricities of 0, 2.5, 5 and 10°, while monitoring fixation with an eye-tracker. For stereopsis, the target region had a coarse disparity step

of 10 arcmin, and observers indicated whether the target was in front or behind the fixation plane. For contrast suppression, the target region had a 60%-contrast decrement in one eye and 100%-contrast in the other eye, with zero disparity. The ability to detect local decrements was compared in the two eyes. In addition to contrast perimetry, we measured decrement thresholds at all the tested locations for a subset of observers. The ratio of decrement thresholds in the two eyes provided an estimate of local suppression due to dichoptic masking. Results show a general agreement between loci showing poor stereopsis and loci showing an imbalance in dichoptic masking among the 12 individuals we tested, including 6 with amblyopia or microstrabismus. The one exception was an individual with strabismic amblyopia who had poor stereopsis but excellent contrast sensitivity at all tested locations in both eyes, consistent with McKee et al. (2023). For the remaining participants, it appears that contrast suppression partially accounts for the impairment of stereopsis in the central visual field.

NIH Grant 1R01EY034370

53.357 EXAMINING THE RELATIONSHIP BETWEEN A SIMULATED GLAUCOMA IMPAIRMENT AND POSTURAL THREAT ON QUIET STANCE

Lisa K. Lavalle[1], Nora Pourhashemi[2], Taylor W. Cleworth[2]; [1]Queen's University, [2]York University

BACKGROUND: Vision provides a crucial source of sensory feedback for balance control. Peripheral visual field deficits developed through glaucoma contribute to balance deficits and a fear of falling. Currently, there is no work that examines the direct effect of fear on balance control among glaucoma patients. Therefore, this study aimed to examine the impact of height-induced fear on balance control among healthy individuals exposed to a simulated glaucoma impairment. METHODS: 31 healthy participants stood quietly on a force plate while wearing a virtual reality head-mounted display (HTC Vive) to collect center of pressure (COP) and head displacement (HD) data. Electrodermal activity (EDA) was recorded from the right palm and surface electromyography (EMG) was collected from the right tibialis anterior (TA), medial gastrocnemius (MGast), and soleus (Sol). Trials were 60s, with two at ground level and two at 7 virtual metres above ground, each exposing participants to normal vision and a simulated glaucoma impairment. After each trial, questionnaires were completed to assess balance confidence, fear of falling, perceived stability, and state anxiety. RESULTS: At height, EDA, fear, and anxiety significantly increased, while perceived stability decreased. Anteroposterior (AP) COP sample entropy (SE) and mean power frequency were significantly greater at height with normal vision, while root mean square amplitude increased at height with the glaucoma simulation. AP HD SE was significantly greater at height with normal vision only. Mediolateral COP SE was significantly greater at height for both visual conditions. Lastly, TA activity and TA/SOL co-contraction was greater with normal vision. CONCLUSIONS: Overall, while experiencing height-induced fear with normal vision, participants developed a tighter control of upright stance (increased MPF, SE, TA/Sol activation); however, this was not observed for the simulated glaucoma conditions. Therefore, balance deficits among glaucoma patients may be mediated by fear of falling contributing to a potentially maladaptive strategy.

Queen's University Faculty of Health Sciences, NSERC, VISTA at York University.

TUESDAY, MAY 21, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Action: Representation

53.358 GOAL UNCERTAINTY BIASES MEMORY FOR OBSERVED ACTIONS

Samuel McDougle[1,2] (samuel.mcdougle@yale.edu), Zekun Sun[1]; [1]Department of Psychology, Yale University, [2]Wu Tsai Institute, Yale University

Similar actions can be taken in different contexts. For example, we may make similar reaching movements towards a glass of water either when water is the only choice in front of us, or when a decision between competing goals must be made (e.g., a glass of water versus a neighboring glass of wine). This raises an intriguing question: Does the context in which an action is observed leave traces in our memory for that action? To address this, we algorithmically generated 10 videos, each depicting an agent's hand following a unique trajectory to a target. Participants were instructed to closely observe how the agent moved their mouse. After a brief mask, they then replicated the trajectory they observed by moving their own mouse into the same target. Critically, each of the curved trajectories was presented in two distinct contexts: the agent's cursor either moved to a single target in the workspace (single-target context) or toward one of two possible targets (dual-target context). We reasoned that dual-target movements would be perceived more as "choices" while single-target movements would be perceived more as "demonstrations." Thus, we hypothesized that the presence of a second goal would repel observers' memories of the movement trajectories toward the chosen target. Analysis of reproduced trajectories supported our prediction: Participants consistently reproduced more efficient trajectories in the dual-target context than the single-target context (as indicated by significantly different initial angles and positional deviations of their movements). A follow-up experiment replicated this finding by having participants reproduce movements without a visible cursor, ruling out the possibility that the effect was driven by feedback during reproduction. The results suggest that the memory of an observed action can be shaped by the context in which that action is executed: uncertainty about an agent's intent biases our memory of their actions.

NIH - R01NS132926

53.359 THE DYNAMICS UNDERLYING THE REPRESENTATION OF OBSERVED ACTIONS AT DIFFERENT TAXONOMIC LEVELS

Tonghe Zhuang[1,2], Gregor Volberg[2], Angelika Lingnau[2]; [1]Justus Liebig University Giessen, [2]University of Regensburg

Previous studies examined the representation of actions across different hierarchical levels in the occipitotemporal and parietal cortex, yet the dynamics underlying this process remain unclear. In this study, we aimed to reveal the latencies associated with the representation of observed actions at different taxonomic levels using time-resolved representational similarity analysis (RSA) of EEG data and EEG-fMRI fusion. In our experimental setup, participants were presented with static images depicting human actions. Each image belonged to one out of three superordinate level categories (e.g. 'locomotion'), one out of six basic level categories (e.g. 'swimming') and one out of twelve subordinate level actions (e.g. 'swimming backstroke'). We established Representational Dissimilarity Models to capture pairwise differences between actions, separately for the three taxonomic levels. Results of the time-resolved RSA showed consistently highest peak latencies around 170 ms for all three taxonomic level models. The EEG-fMRI fusion provided further insights, indicating that action representations measured via fMRI in the lateral occipitotemporal cortex (LOTC) exhibited the highest similarity with EEG data around 230 ms. In summary, our findings suggest that actions across three taxonomic levels occur simultaneously, with the LOTC playing a crucial role in elucidating the spatiotemporal aspects of hierarchical organization in action representations. This research contributes to a deeper comprehension of the neural processes underlying the recognition and classification of actions in the human brain.

Angelika Lingnau was funded by a Heisenberg-Professorship (German Research Foundation, LI 2840/2-1; LI 2840/4-1). Tonghe Zhuang was funded by a PhD stipend from the Chinese Scholarship Council and a stipend from the 'Finanzielles Anreizsystem zur Förderung der Gleichstellung.'

53.360 LATERALIZED REPRESENTATIONS OF CAUSE AND EFFECT IN ACTION OBSERVATION

Moritz Wurm[1] (moritz.wurm@unitn.it), Philipp Flieger[1,2]; [1]CIMeC, University of Trento, Italy, [2]Justus-Liebig-University Gießen, Germany

Understanding object-directed actions requires the visual analysis of how body parts interact with objects and how these interactions lead to changes in the objects. However, whether, and if so, how interaction (as cause) and object change (as effect) are distinctively processed in the brain remains unexplored. Based on previous findings, we hypothesized that interaction and change are represented in a lateralized manner in left vs. right nodes of the action observation network, respectively. In four fMRI sessions, 30 right-handed participants observed videos of object-directed actions (e.g., breaking a stick), corresponding abstract animations (e.g., a triangle hitting a rectangle, causing it to break in half), and animations of interaction (triangle hits rectangle) and object change (rectangle breaks in half) in isolation. Using cross-decoding, we isolated either the interaction or change in the actions (train classifier to discriminate actions, test on interaction-only or change-only animations, respectively). As hypothesized, we found that cross-decoding between actions and interaction-only animations was stronger in left vs, right anterior inferior parietal lobe (aIPL) and lateral occipitotemporal cortex (LOTC), whereas the opposite pattern of results was found for cross-decoding between actions and change-only animations. In addition, in bilateral aIPL, cross-decoding between actions and animations depicting both interaction and change was stronger than the sum of cross-decoding between actions and interaction-only or change-only animations. This super-additive effect points toward a higher-level representation of cause-effect structures beyond representations of interaction and change as isolated components. These findings demonstrate that left and right hemispheres have distinct roles in representing the interaction between entities and the induced change, respectively, and that interaction and change are integrated to cause-effect structures in aIPL. Together, these findings shed new light on the interplay of left and right LOTC and aIPL in the physical understanding of observed actions.

53.361 THE NEURAL DYNAMICS OF NATURAL ACTION UNDERSTANDING

Diana C Dima[1,2] (ddima@uwo.ca), Jody C Culham[1], Yalda Mohsenzadeh[1,2]; [1]Western University, [2]Vector Institute for Artificial Intelligence

Humans rapidly make sense of an ever-changing visual world, extracting information about people's actions in a wide range of settings. Yet it remains unclear how the brain processes this complex information, from the extraction of perceptual details to the emergence of abstract concepts. To address this, we curated a naturalistic dataset of 95 short videos and sentences depicting everyday human actions. We densely labeled each action with perceptual features like scene setting (indoors/outdoors), action-specific features like tool use, and semantic features categorizing actions at different levels of abstraction, from specific action verbs (e.g. chopping) to broad action classes (e.g. manipulation). To investigate when and where these features are processed in the brain, we leveraged a multimodal approach, collecting EEG and fMRI data while participants viewed the action videos and sentences. We applied temporally and spatially resolved representational similarity analysis and variance partitioning to characterize the neural dynamics of action feature representations. We found that action information is extracted in the brain along a temporal gradient, from early perceptual features to later action-specific and semantic information. We mapped action-specific and semantic features to areas in parietal and lateral occipitotemporal cortices. Using cross-decoding across videos and sentences, we identified a late (~500 ms) modality-invariant neural response. Our results characterize the spatiotemporal dynamics of action understanding in the brain, and highlight the shared neural representations of human actions across vision and language.

This work was supported by BrainsCAN at Western University through the Canada First Research Excellence Fund, a Western Interdisciplinary Development Initiatives Grant, a Vector Institute for Artificial Intelligence Research Grant, and a Western Postdoctoral Fellowship.

53.362 DO PEOPLE PICK UP VARIABLES THAT REFLECT OBJECT HEIGHT RELATIVE TO ACTOR HEIGHT WHEN PERCEIVING OTHERS' MAXIMUM VERTICAL ONE DEGREE-OF-FREEDOM REACH HEIGHTS?

Nicholas A. Garcia[1], Keith S. Jones[1]; [1]Texas Tech University

Jones et al. (2023) investigated what variable(s) people pick up when perceiving others' maximum vertical one degree-of-freedom reach heights. Their results suggest participants may have picked up variables that reflect object height relative to actors' top of head heights when actors' upper bodies were visible, and object height relative to actors' leg lengths when only actors' lower bodies were visible. Leg length is highly correlated with overall height. Accordingly, Jones et al. argued participants may have picked up variables that reflect object height relative to actor height in both cases. The present research evaluated Jones et al.'s explanation. In two experiments, participants viewed point-light displays depicting an actor and object. In Experiment 1, actors were depicted with short or tall necks and markers placed on their full or upper bodies. In Experiment 2, actors were depicted with short or tall legs and markers placed on their full or lower bodies. These manipulations changed object height in relation to actors' top of head heights (Experiment 1) and object height in relation to actors' leg lengths (Experiment 2) but did not change the actors' maximum vertical one degree-of-freedom reach heights. Participants adjusted the object's height to what they perceived to be the actor's maximum vertical one degree-of-freedom reach height. In Experiment 1, judgments for the short-necked actors were lower than judgments for the tall-necked actors. In Experiment 2, judgments for the short-legged actors were lower than judgments for the tall-legged actors in the Lower Body condition, but not in the Full Body condition. These results support Jones et al.'s explanation. Further, they move us closer to identifying candidate variables that people pick up when perceiving actors' maximum vertical one degree-of-freedom reach heights and enhance our understanding about how people perceive others' affordances.

53.363 'OBJECT INTERACTION FIELDS': EVIDENCE FOR SPONTANEOUS AGENTIVE INFERENCES ABOUT OBJECT-OBJECT INTERACTIONS IN VISUAL PROCESSING

Aalap Shah[1] (aalap.shah@yale.edu), Ilker Yildirim[1]; [1]Yale University

Object-based affordances, as defined by the Gibsonian account, refer to the possibilities of interaction that an object offers to an agent. Within this conventional framework, affordances are operationalized as the agent's automatic motor-system responses on objects, such as reaching to grasp. However, our capacity to perceive interactions extends beyond such *simple agent-object relationships*, to also include how objects might interact with other objects. For instance, we almost instinctively see a hook as a place for a raincoat to hang or a flat tabletop as a surface for objects to sit on. To explore these *agentive inferences of object-object interactions*, we conducted several experiments involving a distance reproduction task. Participants were asked to recreate the distance between two objects seen in a target image: (1) a cupholder and (2) either a cup or a box. Critically, the hook on the cupholder was oriented either upward (hook-up) or downward (hook-down). This allowed us to modulate the presence or absence of possible interactions between objects. In the hook-up condition, the possible interactions are present because the cup can be hung on the hook in a stable manner. In the hook-down condition, no such interaction is possible since the cup can no longer be hung. Meanwhile, regardless of the hook orientation, no such interactions are possible in the box condition. The results were clear and striking: The effect of the hook orientation on the accuracy of the reproduced distances was significantly different between the cup and the box conditions, even when all other variables were equated for. These results held across several direct replications. This phenomenon establishes a new domain of affordances that are as automatic as simple agent-object affordances. We suggest the existence of 'object interaction fields' in the mind that identify stable configurations of multi-object systems under gravity and goal-directed action.

53.365 ROBUST DIFFERENCES IN TIME-TO-CONTACT ESTIMATION IN RESPONSE TO A POSTURAL MANIPULATION

Bjoern Joerges[1] (bjoerges@yorku.ca), Laurence R. Harris[1]; [1]York University

The classic time-to-contact literature is largely focused on the optic variables that guide human performance in interception tasks. While prior information (e.g., internalized knowledge on the effects of Earth gravity) has received some attention more recently, other types of non-visual factors have been largely disregarded. In this study, we therefore used a simple postural manipulation to investigate how vestibular and somatosensory cues might affect time-to-contact estimates. To this end, we immersed observers in a VR office environment, which provided strong structural cues on its orientation, and showed them targets travelling on parabolic trajectories in the fronto-parallel plane. These targets could travel at different horizontal and initial vertical speeds and disappeared at a random time in the second half of their motion. Participants then had to indicate by button press when these targets would have returned to their initial height. The trajectories could unfold either according to natural gravity or according to an inverted (i.e., upwards) gravity, and participants completed the experiment while standing upright and while lying supine. The virtual environment and stimulus were head-fixed, i.e., the virtual floor was always simulated at the observers' feet. While we, unexpectedly, found no differences between natural and inverted gravity targets, we did observe a consistent overestimation (i.e., a later button press) of time-to-contact when lying supine in comparison to standing upright. While other studies relate reported postural differences in estimated time-to-contact to gravity (Baurès & Hecht, 2011,Perception 40, 674-681), our findings are unlikely due to an interaction with such Earth gravity prior. Rather, they suggest that non-visual cues can have a substantial impact on estimated time-to-contact even when they are uninformative with regards to the task at hand.

53.366 SENSORIMOTOR BETA ENHANCEMENT, NOT MU SUPPRESSION, DIFFERENTIATES EMOTIONAL VS. AFFECTIVELY-NEUTRAL CONTENT

Alison M. Harris[1] (aharris@cmc.edu), Chandlyr M. Denaro[1], Catherine L. Reed[1]; [1]Claremont McKenna College

According to the influential action simulation theory of embodied cognition, we recognize others' emotions by internally simulating their actions with our own sensorimotor systems. Support for this idea comes from studies using electroencephalography (EEG), which have reported reductions in the power of the mu (8-14 Hz) and beta (16-20 Hz) rhythms over sensorimotor cortex, both when executing one's own movements and observing the actions of others. Specifically, we have previously found that observation of emotional vs. affectively-neutral whole-body movements is associated with greater mu suppression, driven by reduced responses to emotional actions, as well as increased beta enhancement while viewing neutral actions. However, recent evidence suggests that measurements of these periodic oscillations may be conflated with underlying aperiodic ("1/f-like") neural activity, which can influence the shape of the EEG power spectrum. Here we examined the relative contributions of mu and beta rhythms using high-density 128-channel EEG (N = 117) during the observation of emotional and neutral point-light displays (PLDs). To control for low-level motion, all PLDs were compared to scrambled versions of the same actions. After removing the aperiodic component from the data, we found significant mu suppression for coherent vs. scrambled PLDs, but not emotional content. However, consistent with our previous findings, there was a significant difference between emotional and neutral PLDs in the beta band (14-19 Hz), reflecting enhanced activity over frontocentral sensors for neutral movements. Together, these data suggest that neural oscillations in the mu and beta bands contribute differentially to distinct aspects of action observation.

This research was funded by NSF Award BCS #1923178

TUESDAY MORNING POSTERS IN PAVILION

Tuesday, May 21, 8:30 am – 12:30 pm, Pavilion

Eye Movements: Natural world and VR

53.401 IMPACT OF FLASHBANGS ON EYE-BEHAVIOR WHILE NAVIGATING AND PERFORMING VISUAL SEARCH TASKS IN STRESSFUL OPEN-WORLD ENVIRONMENTS

Leah Enders[1] (leah.r.enders.civ@army.mil), Russell Cohen Hoffing[1], Thomas Rohaly[2], Heather Roy[1]; [1]DEVCOM Army Research Laboratory, Aberdeen Proving Ground, MD USA, [2]DCS Corp, Alexandria, VA USA

In the field, Soldiers experience different situations that are potentially threatening and/or disorienting. One example is flashbang grenades, which are designed to incapacitate individuals by overstimulating both hearing and visual systems. Maintaining vigilance is imperative during search tasks and events like flashbangs can momentarily incapacitate visual capabilities, increasing vulnerability and reduce threat and target detection abilities. Understanding how visual capabilities are incapacitated can lead to the development of future technology. Here, we investigate how vision behavior changes directly following flashbangs to understand how these events impact visual behavior in certain environments. Twenty military personnel (including active duty and veteran personnel) freely navigated two different desktop environment themes: a neutral theme (construction zone) and a military theme (tanks, barbed-wire barriers, weapons, etc.) while searching for a vehicle target (N = 10) placed at pseudo-random intervals along the street, sidewalk, or alleyways. Both environments included on-theme background noises (e.g., jackhammer, gunfire). As participants navigated the environments, flashbangs (N=5) were triggered at different progress locations. Flashbangs included both a blinding light flash (~400ms duration) and a sharp ringing noise. We report results of significant differences in saccades, fixations, pupil dilation, and blinks between the environment theme immediately prior-to and after the flashbang. Together these findings demonstrate how flashbang events impact the ability to scan for threats and targets in the environment.

53.402 WHEN YOU ARE NOT ALONE: LOOKING AT STEPS WHEN SOMEONE IS ON THE STAIRCASE IN FRONT OF YOU

Andrea Ghiani[1] (a.ghiani@vu.nl), Sara Kabbara[1], Eli Brenner[1]; [1]Vrije Universiteit Amsterdam

Where people look when navigating staircases has generally been studied with participants alone on the staircase. In daily life people regularly encounter other people while climbing staircases. Does the presence of other people influence gaze? We previously found that carrying a tray does not influence the sequence in which steps are fixated or the number of fixated steps. The presence of another person walking in front of a participant might therefore not influence the sequence and number of fixated steps either. To find out whether it does, we measured the gaze of 26 participants walking through a two-story furniture store. They were asked to examine all the furniture, and then to examine it all again, so that they had to walk up and down the staircase between the two stories twice. On one of the two occasions, an associate intentionally walked in front of the participant on the staircase. Participants were not aware that the associate was involved in the experiment. The presence of the other person on the staircase did not influence the participants' sequences of fixations, but it did make them look at fewer steps and look closer to their feet (fewer steps ahead). As when walking with and without a tray, we found a clear correlation between the fraction of fixated steps when walking alone and when walking behind another person, confirming previous findings that revealed clear consistent differences between people in terms of the number of steps they look at.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 955590 (OptiVisT).

53.403 GAZE SEQUENCES DURING WALKING IN COMPLEX TERRAIN

Kathryn Bonnen[1] (kbonnen@iu.edu), Ziggy Waller; [1]Indiana University

Walking over complex terrains requires that humans gather information about the upcoming path to find good foothold locations. Here we examine the sequences of eye movements used in service of this task, observing how walkers shift their gaze between upcoming footholds. Methods & Data Processing: We recorded the eye movements and body movements of participants (n=8) walking over rough terrain. The data was processed to find fixations and find foothold locations (see Bonnen et al 2021). A photogrammetry technique was applied to reconstruct the depth structure of the ground (Muller et al 2023). Combining the reconstructed ground and the gaze data we calculated the gaze-ground intersections. For each fixation, we found the foothold location closest to the gaze-ground intersection. If the gaze location was less than .4 meters (or about half of a step) from the foothold location, we assigned it to that foothold. We then analyzed these sequences of fixations, focusing on gaze shifts from foothold to foothold (30% of all fixations). Results: We found that when participants shifted their gaze from one foothold to another, it was more likely to be forward down the path (70.4%, .95 CI [67.5, 73.0]), rather than backwards toward their body (7.0%, .95 CI [5.6, 8.7]. The forward gaze shifts were dominated by shifts forward by a single foothold (48.5%, or 69% of all forward shifts). They stayed fixated on the current foothold 22.6% of the time (.95 CI [20.1, 25.2]). Conclusions: When walkers fixate a series of footholds, they tend to move their gaze forward down the path, often shifting forward by a single foothold at a time. This is consistent with the idea that the biomechanics of walking strongly influences the visual search patterns used by human walkers.

53.404 WHERE DO PEOPLE LOOK WHEN THEY WALK OR RUN AT DIFFERENT SPEEDS?

Eli Brenner[1], Andrea Ghiani[1], David Mann[1], Jeroen BJ Smeets[1]; [1]Vrije Universiteit Amsterdam

When walking or running, people spend part of their time looking at the ground surface in front of them. They presumably do so to guide their future foot placement. We were interested in how far ahead they look when they do so. This distance probably depends on the time it takes to select and guide the foot to suitable positions. If this time is independent of how people walk or run, how far ahead they look will be proportional to their movement speed. But considering the fundamental role of ground reaction forces in walking and running, the time needed to guide the foot may primarily depend on the step frequency, with people looking a certain number of steps ahead. If so, how far ahead they look will be proportional to their step length. To examine to what extent speed and step length determine where people look, we asked twelve participants to walk at three different speeds, as well as to run at three different speeds. We measured their eye movements, head orientation, speed, and step frequency. We combined their gaze elevation with their eye height to estimate how far away they were looking. Our participants did not scale how far away they looked to their speed of locomotion: they reached where they were looking earlier when they were moving faster. Neither did they scale how far away they looked to their step length: they looked fewer steps ahead when moving faster. The participants appeared to look a certain distance ahead, irrespective of their speed and step length, both when running and when walking. They did look less far ahead when running than when walking.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 955590 (OptiVisT)

53.405 ASSIMILATION OF OPTIC FLOW IN SACCADIC EYE MOVEMENTS

Hongyi Guo[1] (hguo06@yorku.ca), Alexander Schütz[2], Robert Allison[1]; [1]Centre for Vision Research, York University, [2]University of Marburg

When executing a saccade during locomotion, the position of the saccade target changes during the movement. Ideally, our visual system should adjust the saccade parameters to compensate for the different optic flow fields. We asked how optic flow affects the parameters of saccades to stationary targets, a question that has not been fully answered. The results showed a strong assimilation effect of optic flow, such that saccade amplitudes were biased in the direction of the optic flow. We estimated that between 13% to 25% of optic flow was assimilated. Moreover, saccade latencies were shorter in optic flow fields. These findings indicate that the oculomotor system partially accounts for the expected displacement due to optic flow in the planning of saccadic eye movements, and are consistent with the assimilation of background motion in smooth pursuit eye movements.

This project was made possible by the CREATE-IRTG program, funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and German Research Foundation (Deutsche Forschungsgemeinschaft; abbr. DFG).

53.406 CONTROL OF HEAD-EYE FIXATION IN NATURAL TASKS

Ruitao Lin[1,2] (rlin18@ur.rochester.edu), Michele A. Cox[1,2], Ashley M. Clark[1,2], Paul Jolly[1,2], Sanjana Kapisthalam[1,2], Yuanhao H. Li[1,2], T. Scott Murdison[3], Alina Neverodska[1,2], Jie Z. Wang[1,2], Bin Yang[1,2], Zhetuo Zhao[1,2], Michele Rucci[1,2]; [1]University of Rochester, [2]Center for Visual Science, University of Rochester, [3]Reality Labs

The image on the retina is always in motion during normal fixation, as head and eye movements continually perturb the line of sight. These movements occur even when attempting to maintain steady gaze on a single point, displacing the stimulus on the retina over an area as large as the foveola. Previous research has shown that humans control the amount of retinal image motion during fixation, even though they are not aware of doing so, shaping the structure of the temporal luminance flow impinging onto the retina (Intoy & Rucci, 2020; Lin et al, 2023). Here we examined whether, during normal head-free viewing, tuning of fixational motion occurs irrespective of the eccentricity of fixation, i.e., whether fixations with the eyes at various angles within the head yield retinal motion with similar characteristics. We simultaneously measured head and eye movements using a custom apparatus that enables accurate reconstruction of retinal stimulation. This device is composed of a motion capture system (Optitrack) integrated with the oscillating field monitor (Eibenberger et al, 2016), a magnetic-induction eye-tracker designed to create three orthogonal and highly uniform oscillating magnetic fields. Subjects (N=6) wore scleral eye coils in both eyes and a tightly fitting helmet with markers while they engaged in three tasks: visual searching, object sorting, and a standard Snellen acuity test. The eccentricity of fixation varied widely across tasks, ranging from a few degrees in the acuity test to over 40 degrees in searching. Results confirmed that inter-saccadic fixational motion was tuned according to the task, emphasizing luminance modulations in a task-relevant spatial frequency range. Crucially, retinal image motion maintained consistent task-dependent characteristics irrespective of the eye's orientation relative to the head. These findings indicate that selecting the appropriate amount of retinal image motion is an important principle of head-eye coordination.

Research supported by Reality Labs and NIH grants EY018363 and P30 EY001319

53.407 SPHEER: A RICH DATASET OF TIME-RESOLVED GAZE AND HEAD MOVEMENTS IN VIRTUAL REALITY

Erwan David[1] (david@psych.uni-frankfurt.de), Melissa L.-H. Vo[1]; [1]Goethe University Frankfurt, Scene Grammar Lab, Germany

The opportunities offered by extended reality (XR) devices with embedded eye tracking capabilities have opened the door to many new studies and applications, may they be scientific, video-ludic or for training. We now share a rich dataset related to 11 experiments in which we collected both eye and head data rotations as well as positions thanks to eye tracking in virtual reality devices: the Scene Perception, Head and Eye in Extended Reality dataset. The different experiments cover several types of stimuli (360 images, 360 videos, 3D environments) and tasks (free-viewing, object search). This dataset totals more than 380 participants, and accumulates over 6 days of continuous trial time, sampled at 120Hz and 250Hz. Along with this dataset we share meta-data linking every trial to the related experimental conditions and stimulus (e.g. bounding boxes of objects in indoor scenes) in order to make the dataset as rich and useful as possible. In addition, we share a new dataset of gaze data, created specifically for the identification of gaze events in 3D. We implemented a testing protocol in which participants produced fixations, saccades, smooth pursuits and vestibulo-ocular responses, and additionally made them vary vergence distance during some of these events. Such data will be very helpful for creating new methods of identifying gaze-events in 3D, which are now starting to gather substantial interests thanks to XR devices, but have not been the focus of as much dedicated methodological work as of yet, due to common eye tracking experiments historically being set on desktop computers. We share this rich dataset in the hope that communities interested in modeling gaze (e.g., saliency, or scanpath prediction models) get a chance to create new specific and generalised models derived from a rich and broad corpus of real-world eye and head tracking data.

This work was supported by SFB/TRR 26 135 project C7 to Melissa L.-H. Võ and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project 'The Adaptive Mind').

53.408 THE BLANKING EFFECT ON DETECTING CHANGES IN NATURAL SCENES ACROSS SACCADES

Jake Ferreira[1], Yong Min Choi[1], Tzu-Yao Chiu[1], Julie D. Golomb[1]; [1]The Ohio State University

Every time we make a saccade, the image projected onto our retina drastically changes, yet our perceptual experience of the environment remains remarkably stable. Previous studies asking subjects to detect changes across saccades found that we are surprisingly poor at detecting trans-saccadic changes. However, performance markedly improved in the presence of a post-saccadic blank. This blanking effect suggests that the visual system assumes stability across saccades and accesses trans-saccadic memory only when there is extreme evidence (e.g., a blank) against stability. We investigated whether the stability mechanisms evidenced by the blanking effect are employed when attempting to detect trans-saccadic changes in natural scene images, and which signals play a role in triggering this effect. Subjects viewed two scene images presented successively while making a saccade, and performed a scene change detection task. In half of the trials, a 200 ms blank screen was displayed between the presentation of the two scenes. We systematically manipulated the degree of difference between scene images using a continuous scene wheel (Son et al., 2021), and the probability of "changed" responses as a function of scene difference was fitted to psychometric functions. In the saccade condition, we found a significantly heightened sensitivity to detect scene changes (a lower threshold) and a stronger bias to report "changed" (a higher guess rate) in the presence of a post-saccadic blank, indicating that trans-saccadic vision maintains a stable perception of complex natural scenes by assuming the visual world as stable in the absence of extreme evidence against stability. In contrast, an analogous blank had no such effect in a simulated-saccade condition where subjects remained fixated while the scene image shifted spatially to mimic a saccade-induced retinal shift, underscoring a critical role of extra-retinal signals in triggering this stability mechanism.

This research was funded by NIH R01-EY025648 (JG) and NSF 1848939 (JG).

53.409 FREE-VIEWING OF STATIC NATURAL IMAGES AND MOVIES IN MARMOSET MONKEYS.

Oviya Mohan[1] (omohan@ur.rochester.edu), Amy Bucklaew[1], Dora Biro[1], Jude F. Mitchell[1]; [1]University of Rochester

Previous work (Mitchell, Reynolds, & Miller, 2014) with marmoset monkeys viewing natural images showed that, compared to larger primates, they have a restricted oculomotor range with fixations remaining largely within +/- 10 degrees. However, that previous study used images drawn from internet sources that may have been biased towards centrally placed objects of interest. It thus remains an open question whether or not the restricted range is due to motor-based constraints, i.e. inability/physical discomfort with moving eyes beyond this range, or due to sensory factors, i.e. lack of relevant/salient stimuli in the periphery driving eye movements. To test if eye movements might expand their range for salient peripheral stimuli, we allowed two head-fixed marmosets to free-view either static natural images or videos of their own marmoset colony, containing highly salient scenes of neighbors interacting. Further, they were presented at two different viewing distances (covering either +/-20 degrees or +/-30 degrees on the horizontal). The normalized density of fixations did not differ from previous work that used static natural images drawn from the internet. Further, there was also no difference between the distributions when videos and images were shown at the closer viewing distance where salient stimuli appeared more peripherally than at the farther viewing distance. This supports the conclusion that the oculomotor range is bounded by motor constraints and is robust to manipulations of stimulus salience or eccentricity. We therefore suggest that estimation of head-gaze in marmosets, even in dynamic contexts with salient social stimuli, could provide a reasonable estimate (+/- 10 degrees) of their gaze direction. This paradigm also allows a variety of extensions to study behaviors like social monitoring and gaze-following with head-fixed free viewing primates.

Funding: NIH EY030998, NIH T32EY007125

53.410 EFFECTS OF CONVERGENCE ON BINOCULAR AND MONOCULAR PERCEPTION OF UPRIGHT IN VIRTUAL REALITY

Reem Almagati[1] (ralmagati@berkeley.edu), Raul Rodriguez[1], Jorge Otero-Millan[1,2]; [1]Herbert Wertheim School of Optometry and Vision Science, University of California, Berkeley, [2]Department of Neurology, Johns Hopkins University

When we converge to look at a near target our eyes also typically undergo incyclotorsion where their top poles rotate towards the nose. Tilting the head triggers Ocular Counter-Roll (OCR) which causes both eyes to rotate in a direction opposite to the head. The amount of OCR is <10% of the head tilt, resulting in a tilted retinal image. Studies have reported even smaller OCR with convergence. This interaction reflects a conflict between our evolutionary older balance system and the more recent binocular and stereo system. Here we utilized Virtual Reality

(VR) technology to study how our perception of upright is affected by these torsional eye movements. Subjects performed a Subjective Visual Vertical (SVV) task using the FOVE0 VR headset at three head positions: upright, 20 degrees right-ear down and 20 degrees left-ear down. While they fixated on a head-fixed central target, subjects reported whether a line originating from the target, tilted between -12 to 12 degrees, was to the left or right of world-upright. The fixation target was binocularly viewed while the stimulus line was presented either monocularly or binocularly at two vergence distances: 0.25m (near) or 1.5m (far). Near and far conditions were blocked while the other conditions were randomized. We found a mean difference in the perceived upright between the left and right eye viewing conditions across all head positions: -2.7 ± 0.3 degrees for near and -0.7 ± 0.3 degrees for far. That is, perceived upright was tilted more towards the left in the left eye viewing condition. These results show a bias in the monocular perception of upright consistent with the expected incyclotorsion at near. Our findings suggest that we may not fully account for the torsional eye position when assessing the orientation of a stimulus presented monocularly in a VR setting.

Project was funded by NEI Training grant 5T32EY007043-43 and NEI R00EY027846

53.411 EYE MOVEMENTS DURING FREE VIEWING TO MAXIMIZE SCENE UNDERSTANDING

Shravan Murlidaran[1] (smurlidaran@ucsb.edu), Miguel P Eckstein[2]; [1]University of California Santa Barbara

Introduction: The extent to which eye movements during free-viewing of scenes are influenced by low-level saliency (Parkhurst et.al 2002, Harel et al. 2007, Koehler et al. 2014), local semantic meaningfulness (Henderson et al., 2017, Peacock et al., 2019), or other processes is debated. Here, we hypothesize that during free-viewing, humans direct their eyes to regions that maximize scene understanding rather than locally salient or meaningful regions. Methods: For each image (n=36) we created a scene understanding map (SUM) that assesses the contribution of individual objects to observers' (n=110) scene descriptions (global understanding of the scene) by digitally removing each object from the image and having eighteen raters evaluate the similarity of descriptions to manipulated and original images. We compared the predictions from SUM and other models like saliency (Graph-Based Visual Saliency), DeepGaze, and local meaningfulness to human (n=50 per task) fixations during free-viewing (FV) and scene-description (SD) tasks. Images were presented for 2 seconds while eye position was measured. Results: In both the scene description (SD) task and free viewing (FV) tasks, fixations to the regions most critical to scene understanding (top-highest region in SUM) were significantly higher than those to the top predictions of DeepGaze (pSD=0.0035,pFV=0.0025, with a significant difference starting with the 6th fixation, (pSD=0.013,pFV=0.044)), local meaningfulness (pSD=0.00001, pFV<0.00001, with a significant difference starting with the 4th and 3rd fixation (pSD=0.003, pFV=0.019)) and GBVS saliency(pSD<0.00001, pFV<0.00001, with a significant difference starting with the 4th fixation (pSD=0.037, pFV=0.007)) models. Conclusions: Our findings suggest that during free-viewing, humans do not execute eye movements to low-level saliency or locally meaningful regions but to image regions that maximize the global understanding of the scene.

53.412 MONITORING AROUSAL LEVELS AMONG A WARSHIP CREW: SHIFT- AND TIME OF DAY-VARIATIONS OF SACCADIC VELOCITY

Leandro L. Di Stasi[1,2] (distasi@ugr.es), Jesús Vera[1], Beatriz Redondo[1], Ignacio Grueiro Méndez[3], Jaime J. Mas Esquerdo[4], Eduardo A. Gómez Quijano[5], Carolina Díaz-Piedra[1,6]; [1]University of Granada, Granada, Spain, [2]Joint Center University of Granada-Army Training and Doctrine Command (CEMIX UGR-MADOC), Granada, Spain, [3]Buque de Aprovisionamiento de Combate (BAC) "Cantabria" (A-15), Ferrol, A Coruña, Spain, [4]Escuela Militar de Sanidad, Academia Central de la Defensa, Madrid, Spain, [5]Sección de Técnicas de Apoyo a la Decisión, Estado Mayor de la Armada, Madrid, Spain, [6]Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, US

Background: Due to the highly demanding nature of managing a warship, crewmembers are required to work in a rotating, very challenging shift schedule for very prolonged periods. The resulting short sleep duration and poor sleep hygiene can make difficult to maintain optimal arousal levels. Operational safety may be affected as a result of not optimal (i.e. reduced) arousal levels (i.e. sleepiness/fatigue). Here, we studied arousal variations of the crew of a replenishment oiler operated by the Spanish Navy before/after the shift and along the 24-hour cycle (morning, evening, and night shifts) for ten consecutive days. Methods: We conducted the study on board the Cantabria (A-15) during an international warfare training exercise (FLOTEX22). Twenty-six crewmembers (4 females, 34.67±7.62 years old) were assessed approximately every 4/6 hours. To assess oculomotor indices, we used the Fitness Impairment Tester 2000 (750 Hz, Pulse Medical Instruments Inc., US). The saccadic peak velocity was our main variable. In addition, we assessed intraocular pressure (Icare tonometer TA01, Tiolat Oy, Finland), objective/subjective sleep parameters, subjective levels of arousal and workload, as well as reaction times/errors with a psychomotor vigilance test (Fit-Alert, Miinsys-Optimal Solution S.A., Chile). Results: As expected, sleep time was always less than six hours, with great variability (and considerable sleep restrictions) depending on the shift. Although perceived workload was similar between shifts, fatigue (both subjective and physiological) was greater and performance worse at night shifts. Both saccadic peak velocity and intraocular pressure decreased after all shifts and were especially lower during the night shift. Conclusions: Our data confirm that saccadic peak velocity is a sensitive index of operator arousal levels. Overall, our findings support the viability of an objective, long-term, periodic monitoring of arousal levels in applied military settings. Real-time arousal assessment can support the monitoring of operator status for designing interventions to improve operational safety.

The study was funded by a CEMIX UGR-MADOC grant (Ref. 5/2/20 F2F), by NAVANTIA S.A. (Ref. 5169), and by Miinsys-Optimal Solutions SpA (Ref. 5400). We thanks all the BAC Cantabria crew for their help and support. We also thank D. Soler and F. Angioi for their help during data curation.

53.413 GENERALIZING FIXATION PREDICTIONS WITHIN AND ACROSS DATASETS: TOWARDS A UNIVERSAL MODEL OF FREE-VIEWING FIXATIONS

Matthias Kümmerer[1] (matthias.kuemmerer@bethgelab.org), Harneet Singh Khanuja[1], Matthias Bethge[1]; [1]University of Tübingen, Tübingen AI Center

Predicting free-viewing fixation locations has a long history both in vision science and in computer vision. Recent high-performing models are deep learning based models that are trained on an eye movement dataset such as MIT1003, and subsequently evaluated on benchmarks such as the MIT/Tuebingen Saliency benchmark, which assess model performance on one or multiple datasets. An important challenge that has only marginally been addressed so far, is the desire for saliency models to generalize across different domains, correctly predicting fixation densities for any image and recording setup. In this work, we combine a substantial range of eye movement datasets, including MIT1003, CAT2000, COCO Freeview, FIGRIM, NUSEF, OSIE and others to create a large-scale compound dataset that we envision to grow further over time aiming for maximal size and diversity. On this dataset, we train a fixation prediction model, which is an extended and improved variant of DeepGaze IIE, combining multiple pretrained deep backbones in a joint readout architecture. After training on all or a subset of these datasets, the model is evaluated on the validation splits of all datasets. Our best model improves state-of-the-art by a significant margin on many commonly used benchmark datasets, including MIT300, CAT2000 and COCO Freeview. Our modeling paradigm allows us to assess to which degree gaze patterns from one dataset generalize to other datasets, to which degree using multiple datasets creates synergy effects due to the larger diversity in the data, or to which degree different datasets show conflicting patterns. For example, we find that different datasets require different rescalings of local priority values in a way that is partially, but not fully, explained by different presentation times. Such analyses hint at underlying mechanisms that need to be understood and incorporated into models for building fixation models which are reliably applicable in diverse contexts.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG): Germany's Excellence Strategy - EXC 2064/1 - 390727645 and SFB 1233, Robust Vision: Inference Principles and Neural Mechanisms.

53.414 EASYEYES — VALIDATING A NOVEL METHOD FOR ACCURATE FIXATION IN ONLINE VISION TESTING

Maria Pombo[1], Jan W. Kurzawski[1], Augustin Burchell[1], Nina M. Hanning[2], Simon Liao[1], Najib J. Majaj[1], Denis G. Pelli[1]; [1]New York University, [2]Humboldt Universität zu Berlin

Compared to in-lab testing, online methods allow easier and faster testing of large, more diverse populations. Many psychophysical measurements, including visual crowding, require accurate eye fixation. Accurate fixation is classically achieved by testing only experienced observers who have learned to fixate reliably, or by using a gaze tracker to restrict testing to moments when fixation is accurate. However, both approaches are impractical online as online observers tend to be inexperienced, and online gaze tracking, using the built-in webcam, has a low precision (±4 deg). EasyEyes open-source software reliably measures peripheral thresholds online with accurate fixation achieved in a novel way, without gaze tracking (Kurzawski & Pombo et al., Frontiers in Human Neuroscience 2023). Observers are tasked with using their cursor to track a moving crosshair, and at a random time during successful tracking, a brief peripheral target is presented. Then the observer responds by identifying the target. To evaluate EasyEyes fixation accuracy and thresholds, we tested 12 naive observers in three ways in a counterbalanced order: first, in the lab, using gaze-contingent stimulus presentation; second, in the lab, using EasyEyes while independently monitoring gaze using EyeLink 1000; third, online at home, using EasyEyes. We find that crowding thresholds are consistent and individual differences are conserved. The small root mean square (RMS) fixation error (0.6 deg) during target presentation gets around the need for gaze tracking. Thus, this method enables fixation-dependent measurements online, for easy testing of larger and more diverse populations. Within our sample (N = 12), one observer (S9) peeked. S9 had the highest RMSE between the crosshair and cursor and the most frames with unsuccessful tracking, suggesting that peeking and tracking behavior are associated. We are now assessing whether the accuracy of cursor tracking is a good predictor of "peeking" as a way of detecting peeking.

Funding: NIH grants R01-EY027964 (DP), R01-EY031446 (NM) and P30-EY013079 (core), and a Marie Skłodowska-Curie individual fellowship by the European Commission (898520) (NH). Development of EasyEyes software was developed in part with funds provided by Meta through a sponsored research agreement.

53.415 BRINGING COLOR INTO FOCUS: ACCOMMODATIVE STATE VARIES SYSTEMATICALLY WITH THE SPECTRAL CONTENT OF LIGHT

Benjamin M Chin[1] (bechin@berkeley.edu), Martin S Banks[1], Derek Nankivil[2], Austin Roorda[1], Emily A Cooper[1]; [1]University of California, Berkeley, [2]Johnson & Johnson Vision Care

Humans bring the visual world into focus by changing the power of the lens in their eye until the retinal image is sharp. Light in the natural environment, however, can almost never be focused perfectly because it contains multiple wavelengths that refract differently through the lens. How does the visual system determine the wavelength to put in best-focus? We compared possible strategies used to focus light containing different proportions of long and short wavelengths. Under a 'switching' strategy, an observer would accommodate (focus their lens) to whichever wavelength has the highest luminance. In contrast, under a 'weighting' strategy, the accommodative response would be a weighted sum of the luminances across visible wavelengths. We measured the dynamic accommodative responses of eight participants with an autorefractor recording at 30Hz. On each trial, an observer viewed a three-letter word (24 arcmins per letter) against a black background on an OLED display for six seconds. Halfway through the trial, a focus-adjustable lens generated a step change in the optical distance of the stimulus, synchronized to a change in stimulus color (the proportion of long and short wavelength subpixels). We then fit participants' accommodative changes with both the 'switching' and 'weighting' models separately. The Akaike Information Criterion showed that for all but one subject, the likelihood of the data was greater under the 'weighting' model. Increasing luminance of long wavelengths caused the eye to accommodate nearer, while increasing luminance of short wavelengths caused it to accommodate farther. This is remarkable because it implies that people may bring wavelengths into best focus that are weak or even absent from the visual stimulus. Using these data, we aim to develop an image-computable model that can predict how the eye accommodates to the complex spectral and spatial patterns encountered during natural vision.

Tuesday, May 21, 8:30 am – 12:30 pm, Pavilion

Motion: Detection

53.416 OCULAR-FOLLOWING RESPONSES (OFRS) TO BROADBAND VISUAL STIMULI OF VARYING MOTION COHERENCE.

Boris M Sheliga[1] (bms@lsr.nei.nih.gov), Edmond J FitzGibbon[1], Christian Quaia[1], Richard J Krauzlis[1]; [1]Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health

Manipulations of the strength of visual motion coherence have been widely used to study behavioral and neural mechanisms of visual motion processing. Here we asked how changing the strength of motion coherence in different spatial frequency (SF) bands of a broadband stimulus impacts Ocular-Following Responses (OFRs). We recorded horizontal OFRs in three human subjects using synthesized broadband stimuli: a sum of 1D vertical sine wave gratings (SWs) whose SF ranged from 0.0625 to 4 cpd in 0.05 log2(cpd) steps. Every 20 ms a proportion of SWs—from 25% to 100%—shifted in the same direction by ¼ of their respective wavelengths (motion) whereas the rest of SWs were assigned a random phase (flicker) or shifted by ½ of their respective wavelengths (counterphase) or remained stationary (stationary): 25-100% motion coherence. The magnitude of the OFRs decreased as the proportion of 'not-in-motion' SWs and/or their contrast increased. The effects were SF-dependent: for flicker and stationary SWs, SFs in the range of 0.3-0.6 cpd were the most disruptive; with counterphase SWs, low SFs were more effective. The data were quantitatively well described by a model which combined two factors: (1) an excitatory drive determined by a weighted sum of moving SF components scaled by (2) a SF-weighted contrast normalization term. All weight functions were SF-dependent. The model functions for motion and counterphase were inverted (log)cumulative Gaussians whose offset and sigma were the same; the amplitude was smaller for the counterphase function. The model functions for flicker and stationary SWs were (log)Gaussians whose offset and sigma were the same; the amplitude was smaller for stationary SWs. The differences in the model weight scaling are consistent with the known dependence of SF components' weights upon the stimulus temporal frequency (Quaia et al. 2017; Sheliga et al. 2020).

Supported by the Intramural Program of the National Eye Institute at the National Institutes of Health (NIH).

53.417 PERCEIVED COHERENCE OF GLOBAL MOTION IS HIGHER IN PERIPHERY THAN IN CENTRAL VISION

Quan Lei[1]; [1]Wichita State University

Perceptual sensitivity to global motion depends on many factors including stimulus eccentricity. Little research has investigated how the appearance of global motion stimuli varies across the visual field. This study compared the perceived coherence of global motion between central and peripheral vision and examined the potential association between apparent coherence and motion sensitivity. Participants viewed random dot kinematogram stimuli consisting of 120 dots moving in a 10° circular field (contrast: 60%, speed: 6°/s, lifetime: 50 ms). There were four global motion directions: left, right, up and down, and seven coherence levels varying between 5% and 95%. The stimuli were either presented at the central fixation or 10° to the left or right of fixation with a duration of 300 ms. The participants first judged the direction of global motion, followed by a judgement of motion coherence in comparison to two learned reference levels (low: 5% and high: 95%) in a bisection task. There were considerable individual differences in motion sensitivity as measured by coherence threshold in direction judgment, some participants having lower but others having higher sensitivity in the periphery compared to central vision. Despite the variability in sensitivity, all participants reported a higher level of perceived motion coherence in the periphery compared to central vision, as indicated by the bisection point in coherence judgment. The higher perceived coherence in the periphery was consistent with faster reaction time and with the self-reported relative ease of direction judgment for peripherally presented stimuli. However, there was not a systematic association between perceived coherence and motion sensitivity across participants. The findings suggest that the perceived coherence of global motion stimuli and the sensitivity to such stimuli might be governed by distinct mechanisms. The higher perceived coherence in the periphery is attributable to a smooth motion prior applied where the visual resolution is low.

This research was partially supported by the Cognitive and Neurobiological Approaches to Plasticity (CNAP) Center of Biomedical Research Excellence (COBRE) of the National Institutes of Health under grant number P20GM113109.

53.418 IMPAIRED PERCEPTION OF ISOLUMINANT RED-GREEN CONTRAST MODULATION STIMULI: EVIDENCE FOR A MAGNOCELLULAR PATHWAY MECHANISM.

Ana Ramirez Hernandez[1] (ana.ramirezhernandez@mail.mcgill.ca), Curtis Baker[2], Ari Rosenberg[3]; [1]McGill University, [2]McGill University, [3]University of Wisconsin-Madison

Parasol retinal ganglion cells form the magnocellular pathway and support perception of achromatic visual stimuli at high temporal frequencies (TF) and low spatial frequencies (SF). Parasol cells also exhibit nonlinear behavior like that of Y-type retinal ganglion cells in the cat, giving a second harmonic response at high spatiotemporal frequencies, and therefore can be considered "Y-like" cells. Previous neurophysiology showed that responses to contrast modulation (CM) stimuli, composed of a high SF grating (carrier) whose contrast is modulated by a low SF sinewave (envelope), are driven by the nonlinear subunits of Y-like cells. Recent human psychophysics has shown that direction discrimination of CMs at high spatiotemporal carrier frequencies may reflect nonlinear processing of Y-like cells (Ramirez et al, 2022). Since Y-like cells do not process color, CM motion direction discrimination should be impaired at isoluminance. It is commonly known that motion is absent or reduced at isoluminance. However, residual motion performance might be mediated by the

luminance pathway, as indicated by impaired chromatic motion when adding luminance noise. Healthy normal subjects monocularly viewed luminant yellow-black (Y-B) or isoluminant red-green (R-G) CMs presented at the center of a CRT monitor while fixating eccentrically. They reported the direction of the moving envelope. Within each block of trials, envelope contrast for Y-B and R-G CMs was varied with the method of constant-stimuli. Superimposed 1-D luminance (Y-B) noise was added at different contrast levels. We found that subjects could discriminate motion direction for both kinds of stimuli in the absence of masking noise. Luminance (Y-B) noise maskers affected performance of Y-B CMs only when above the masker's own detection threshold, while R-G CMs were impaired at substantially lower noise contrasts. These results are consistent with our hypothesis that CM responses are driven by Y-like neurons in the magnocellular pathway.

Supported by Canadian NSERC grant RGPIN-2023-03559 to C.B. and National Institutes of Health Grant EY035005 to AR.

53.419 NON-FEATURE TRACKING IN STEREOSCOPIC MOTION DISCRIMINATION AT SHORT DURATIONS

Ichasus Llamas-Cornejo[1] (illamas@ucm.es), Omar Bachtoula[1], Ignacio Serrano-Pedraza[1]; [1]Department of Experimental Psychology, Universidad Complutense de Madrid, Madrid, 28223, Spain

In this study we tested the hypothesis that stereoscopic (cyclopean) motion is processed by a low-level mechanism that does not track features (Patterson, 2002). We conducted a motion discrimination experiment using stereo and luminance stimuli while we measured eye movements. The stimuli were luminance and stereo gratings with spatial frequency of 0.2 and 0.4 c/deg, drifting at temporal frequencies of 1 and 4 Hz and two orientations, vertical and horizontal. The horizontal disparity of stereo stimuli was 2.4 arcsec and for luminance stimuli the contrast was 0.8. Proportion of correct responses were estimated using the method of constant stimuli testing five durations between 63 and 398 ms. To measure eye movements, we used the EyeLink 1000 in binocular mode with a temporal resolution of 2000 Hz. Results showed that for all conditions and durations higher than 100 msec the proportion of correct responses was higher than 0.66 for stereo, and for durations higher than 63 ms the proportions were higher than 0.85 for luminance. For vertical luminance stimuli, we found that eyes only moved in the direction of stimulus motion for durations higher than 150 ms; and for stereo, we found small eye movements (<0.2 deg) in the direction of motion only for the longest duration (398 ms). Thus, no feature tracking was found for direction discriminations at short durations in both luminance and stereo. For horizontal luminance stimuli we found the same pattern of eye movements, however, for stereo we found that independently of the stimulus motion (upwards or downwards), eyes always moved downwards about 0.2 deg. Thus, considering all conditions, our results suggest that stereoscopic motion is processed by a low- level mechanism that does not track features.

Supported by grant PID2021-122245NB-I00 from Ministerio de Ciencia e Innovación (Spain) to ISP.

53.420 TEMPORAL CHARACTERISTICS OF PERCEIVED MOTION FLOW OF NATURALISTIC MOVIES

Yung-Hao Yang[1] (yunghaoyang@gmail.com), Taiki Fukiage[2], Zitang Sun[1], Shin'ya Nishida[1,2]; [1]Cognitive Informatics Lab, Graduate School of Informatics, Kyoto University, [2]NTT Communication Science Labs, Nippon Telegraph and Telephone Corporation*

Visual motion perception involves the computational processing of signals evolving in space and time. Prior research using simple artificial stimuli suggested that human motion processing has sluggish temporal characteristics, integrating instantaneous motion over several tens of milliseconds or longer. Nevertheless, the dynamics of perceived motion in naturalistic scenes remain unexplored. Drawing on a recently proposed method (Yang et al., iScience, 2023), we explored temporal characteristics of the perceived optical flow in naturalistic scenes. Five movie clips sourced from a slow-flow version of the MPI Sintel Dataset, each featuring a large transient change in the middle of the clip, were presented at 60FPS within a circular aperture. Amidst the movie presentation, a tiny dot was momentarily flashed at the center of the aperture to indicate the spatiotemporal location of the target that the observers had to report by matching the speed and direction with a subsequently presented brown-noise field. Critically, to assess the temporal dynamics of perceived flow, the probed dot flashed -66.7ms, -33.3ms, 0ms, +33.3ms, or +66.7ms from the physical transient change. To control stimuli bias, we played each clip in eight ways: forward or backward in time, and four spatial flip conditions. The perceived vector, averaged across trials, shows gradual changes over time around the physically abrupt motion transition. This mainly reflects trial-by-trial temporal uncertainty of perceived direction transition. The temporal pattern of the averaged perceived vector can be well described by Gaussian temporal blurring of the ground-truth vector sequence, with [center, FWHM (Full width at half maximum) of Gaussian (ms)] = [+3.5, 119.7] for group data, and [+12.7, 78.1], [-40.2, 151.5], [-17.7, 117.4], [+30.7, 84.4] for individual data. The results indicate that the temporal window of motion processing, within which perceived motion cannot be accurately aligned in time with physical motion, is approximately 100 ms for naturalistic movies.

This work has been supported by JSPS Kakenhi 20H00603 and JP20H05957.

53.421 DO CNNS TRAINED ON SELF-MOTION VIDEOS DEVELOP SENSITIVITY TO 1ST- AND 3RD-ORDER MOTION?

Zhenyu Zhu[1] (zhenyu_zhu@brown.edu), Thomas Serre[1], William Warren[1]; [1]Brown University

At least two classes of motion information play a role in locomotor control: 1st-order motion energy, such as moving high-contrast texture, and 3rd-order feature-tracking, such as moving object boundaries (Lu and Sperling 1995). Previous literature showed that human heading responses when following a virtual crowd are dominated by 3rd-order motion and weakly influenced by 1st-order motion, revealed when surface texture moves in the Same or Opposite direction as object boundaries (the phi illusion) (Zhu and Warren VSS2023). In this project, we test whether units selective for both 1st and 3rd-order motion emerge in a state-of-the-art Convolutional Neural Network (CNN) model of motion responses in the primate dorsal stream. DorsalNet (Mineault et al. 2021) is a 5-layer CNN trained to estimate self-motion parameters in simulated drone videos. We tested the model's heading estimates respectively on three virtual crowd displays used in Zhu and Warren's (VSS2023) human experiments. In the CONTROL display, DorsalNet layers, like humans, show no differences between Same and Opposite conditions, while responses significantly increase with the number of moving objects for both (Model and Human: p<0.01). In the TEXTURE DISPLACEMENT display, DorsalNet, like humans, shows significant differences when texture motion is coherent (Same > Opposite; Model and Human: p<0.01), but not when motion is incoherent due to small or large displacements. Critically, in the BLURRED BOUNDARIES display, blurring object boundaries reduces the response to 3rd-order motion, increasing the difference between the Same and Opposite conditions in humans (p<0.01), but not in the model. These results demonstrate that DorsalNet has developed a 1st-order motion energy mechanism, which can capture some human heading responses, but not those due to 3rd-order feature-tracking.

Funding: NIH R01EY029745, NIH 1S10OD025181, NIH T32MH115895

Tuesday, May 21, 8:30 am – 12:30 pm, Pavilion

Visual Search: Attention, phenomena 2

53.422 DIFFERENTIAL EFFECTS OF LEARNED AND CUED FEATURE-BASED ATTENTION ON VISUAL SEARCH SLOPES

Douglas Addleman[1,2] (addleman@gonzaga.edu), Viola Stoermer[2]; [1]Gonzaga University, [2]Dartmouth College

Knowledge about non-spatial features of targets can speed visual search when either explicitly provided by an external cue or incidentally learned through experience. Whether these search advantages are due to attentional guidance towards the relevant feature or to post-selective processes is not well understood, particularly for learned feature-based attention. Here, we manipulated visual search set size to test whether incidental learning (Exp. 1) and explicit cues (Exps. 2 & 3) guide attention towards likely target features. Participants searched for a target C (left or right gap) among either three (set size 4) or seven (set size 8) distractor Cs (top or bottom gaps). In Exp. 1 (N=64), targets–unbeknownst to participants–appeared more often in one of the four potential item colors (66% of trials). We found overall search advantages for the frequent target color, replicating previous work, but this advantage did not interact with set size (p = .207), indicating that learning about target colors did not reliably guide attention. Exp. 2 (N=63) used the same paradigm but explicitly cued participants with the likely target color before each trial. There was evidence for guidance from a modest interaction between set size and cue condition (p < .001, ηp2 = .11). Exp. 3 (N=25) tested whether more predictive cues would elicit stronger guidance. Participants completed alternating search blocks containing either 100% predictive color cues or no cues. Here, cues clearly guided attention, with shallower search slopes on cue trials (slope = 54ms/item) than non-cued trials (slope=66ms/item; p < .001, ηp2 = .41). Thus, the amount of guidance was modulated by the validity of the cue and was only reliably present during cued, but not learned attention. Overall, this suggests that different processes underlie how incidental learning and explicit cueing benefit feature-based attention, with cueing more readily affecting attentional guidance than experience.

This research was supported in part by NSF Grant BCS-1850738.

53.423 CROSS-CULTURAL VARIATIONS IN VISUAL SEARCH: EXPLORING ATTENTION DEPLOYMENT STRATEGIES AND NOVEL PRIMING ON SEARCH ASYMMETRY

Yoshiyuki Ueda[1] (ueda.yoshiyuki.3e@kyoto-u.ac.jp), Tsai Chia-Chun[2], Hikari Takebayashi[3], Jun Saiki[1], Su-Ling Yeh[2,4]; [1]Kyoto University, [2]National Taiwan University, [3]Ritsumeikan University, [4]National Humanities Center

Various cultures exhibit different efficiencies when searching for the same simple geometric figure. In Canada and the United States, it is more efficient to search for a long line among short lines than vice versa (typical search asymmetry: Treisman & Gormican, 1988), whereas in Japan and Taiwan, searching for a long line among short lines is equally efficient as searching for a short line among long lines (Tsai et al., 2021; Ueda et al., 2018). One explanation for this variation is the default deployment of attention; search asymmetry might be observed after changing attention deployment even for Japanese and Taiwanese participants. To investigate this hypothesis, we modulated participants' attention deployment before a visual search using a Navon task, in which participants were presented with a large letter composed of smaller letters and responded to either the large (i.e., global attention priming) or small letter (i.e., local attention priming). The results showed that local attention priming did not change search performances, maintaining no search asymmetry for both Japanese and Taiwanese. However, global attention priming led to the opposite-direction search asymmetry (i.e., a short line among long lines was searched more efficiently compared to vice versa) only for Taiwanese. Post-hoc analysis revealed that the opposite-direction search asymmetry is specific to participants who showed longer reaction times. These novel findings suggest that Japanese and Taiwanese participants default to local attention deployment in visual search, while those with longer reaction times employed a different search strategy with global attention priming. The emergence of opposite-direction search asymmetry raises new questions about how scenes with multiple objects are perceived and how individuals determine their attentional deployment in such scenarios.

JSPS KAKENHI (20H00107), National Science and Technology Council in Taiwan (MOST 110-2410-H-002-130-MY3 and MOST 111-2223-E-002-008)

53.424 EXPLORING DYNAMICS OF ATTENTION IN A VIRTUAL REALITY FORAGING TASK: EFFECTS OF TARGET SIMILARITY, INTER-TARGET SPACING, AND TIME PRESSURE

Anna Madison[1], Chloe Callahan-Flintoft[1], Anthony Ries[1]; [1]DEVCOM Army Research Laboratory

Visual foraging tasks are used to understand the dynamics of visual attention in contexts with multiple targets, providing insight into search behavior in complex environments. The presence of multiple targets allows for aspects of attention to be studied as targets are selected one after another. It has been posited that the size of the functional viewing field (FVF), or how much visual information is processed at a given moment, can vary with search difficulty and be indexed by eye movement metrics. In this experiment, we manipulated factors that influence search difficulty using a virtual reality (VR) foraging task while measuring eye and head movements. Participants searched for enemy targets among two friendly forces, where the camo patterns indicated identity. The two friendly force camo patterns varied in similarity to the enemy targets (low- and high-similarity). Targets appeared at varying inter-target spacing (5 or 10 DVA), and participants searched under varying time pressure conditions (blocked: 45 sec or 18 seconds). We found that target similarity, inter-target spacing, and time pressure impacted search and eye movement metrics that index the size of the FVF. The results suggest participants had an overall larger FVF in the low-similarity camo condition compared to the high-similarity camo condition. Specifically, we found in the high-similarity camo search condition, participants made more fixations accompanied by smaller saccades, whereas in the low-similarity camo condition, participants made fewer fixations with larger saccades. Overall, these findings suggest fluctuations in the FVF can be indexed by eye movement behavior in a complex, immersive VR visual foraging task.

53.425 INTERPLAY OF EXPLICIT KNOWLEDGE AND MOTIVATIONAL FACTORS IN THE USE OF ATTENTIONAL CONTROL STRATEGY

Mackenzie J. Siesel[1] (siesel.24@osu.edu), Tianyu Zhang[1], Andrew B. Leber[1]; [1]The Ohio State University

In daily life, people can choose among multiple visual search strategies to find a target of interest (e.g., car in a parking lot). Lab-based studies have revealed that people's chosen strategies are surprisingly suboptimal. While several factors may be responsible, lacking explicit knowledge of the optimal strategy is a key factor predicting suboptimality. Here we examine how explicit knowledge and motivational factors interact to jointly predict strategy. Specifically, will individuals persist in using optimal strategy instructions even when the behavioral benefits are reduced? We further assessed the frequency with which people monitor the utility of their strategy: if they knowingly abandon the optimal strategy when it is perceived to be less useful, how quickly will they reinstate it if the behavioral benefits return? We used the Adaptive Choice Visual Search (Irons & Leber, 2018) to assess strategy usage. In this task, two targets are presented in both a red and blue subset, but only one must be found. The optimal strategy in this task is to search through the less numerous color subset. For the first and third phases of the experiment, we used the standard paradigm in which the subsets were presented in a 2:1 ratio on all trials. In the critical second phase, to disincentivize the optimal strategy (i.e., reduce its behavioral benefits), we mixed color subset ratios such that 30% of trials had the typical 2:1 ratio and 70% had a 1.1:1 ratio. Even with explicit strategy instruction, optimality was significantly decreased in the mixed ratio blocks. However, optimality rapidly returned to its Phase-1 level when the standard ratio was restored in Phase 3. These findings show that explicit strategy knowledge is sometimes insufficient to elicit the optimal strategy, but this choice likely reflects the judicious – and frequent – reevaluation of the strategy's expected behavioral benefits.

This work is supported by NSF BCS-2021038 to ABL.

53.426 A COMMON FEATURE IMPROVES GUIDANCE EFFICIENCY DURING HYBRID SEARCH

Dengxinyi Wei[1,2,3], Zhiheng Zhou[1,4], Joy J. Geng[1,2]; [1]Center for Mind and Brain, University of California Davis, Davis, CA, USA, [2]Department of Psychology, University of California Davis, Davis, CA, USA, [3]Department of Psychology, Yale University, New Haven, CT, USA, [4]College of Psychology, Sichuan Normal University, Chengdu, Sichuan, China

Imagine searching for all the ingredients for a cake at the grocery store. This type of "hybrid" visual search is common in daily life and involves searching for many items from memory at once. Hybrid search involves two stages: visual search for possible targets to select, and then search through memory for a matching target. While we can successfully seek numerous objects concurrently, RT is relatively slow and increases logarithmically with the number of items in the target set. In two studies (N = 147) we tested the hypothesis that having a common feature across targets such as the color "red" (e.g., if searching for apples, strawberries, and tomatoes) will facilitate hybrid search by reducing the target template for visual guidance into one-dimensional "common feature" search. In experiments 1a and 1b, we confirmed that color, as a common feature among targets, was used to guide attention during visual search, as shown by slower RTs when a "common feature" distractor was present. In experiment 2, participants memorized twenty "target" objects in various colors (red, blue, yellow, green) and categories (fruit, toys, animals, furniture, flowers). Before each test trial, a target list of 1, 3, or 5 items was shown. Items were either color-similar (e.g., all red objects) or mixed. Only one of the possible targets was present on a single trial. We found main effects of common color and target set size, but no interaction. While the logarithmic increase in RT due to memory search set size appeared to remain, RTs in the common color condition were overall shorter and did not show serial position effects related to the target list. Finally, individuals that self-reported using common feature for search were faster overall, suggesting it is a strategic shortcut that reduces working memory load and maintenance.

We would like to thank Xinger Yu for helpful comments and concept development. This work was supported by National Institutes of Health (NIH) grant R01MH113855-01 to J.J.G.

53.427 BRIEF ADAPTATION LEADS TO FASTER REACTION TIMES IN A FACE SEARCH TASK

Idris Shareef[1] (ishareef@unr.edu), Michael A Webster[1], Alireza Tavakkoli[2], Fang Jiang[1]; [1]Department of Psychology, University of Nevada, Reno, USA, [2]Department of Computer Science and Engineering, University of Nevada, Reno, USA.

Adaptation to faces can strongly bias the perceived characteristics of subsequently seen faces, but has been found to have relatively weak effects on tasks such as face discrimination. We examined the impact of prior adaptation on the salience of face stimuli, by using a visual search task. Participants adapted for 12 seconds to grayscale images of Asian faces that slowly morphed between 3 different identities. To control for low-level aftereffects, the adaptor was shown at the center of the screen and was 1.25x larger than the test faces, which were arranged in a circle in the periphery. The test array included images of 5 Asian distractor faces and a 6th target face formed by morphing between a White and Asian face. Target salience was controlled by varying the morph level from 20% to 50% White in steps of 10%. In a second, 2-AFC discrimination task using 2-down 1-up staircases, we also measured discrimination thresholds for 3 reference faces that contained 100%, 75% and 50% of the Asian to White face morph, respectively. Prior adaptation to the Asian face images induced significant aftereffects, shifting the Asian-White category boundary in the morph toward the Asian direction. This adaptation also reduced reaction times for detecting the target face in the search task for all of the tested morph levels, but had minimal effect on discrimination thresholds or search accuracy. Our results provide evidence that face ethnicity adaptation heightens the salience of target faces that differ along the ethnicity dimension, potentially because the adaptation renormalizes the adapting ethnicity so that it appears more neutral and therefore less distinctive.

FA9550-21-1-0207

53.428 COGNITIVE RELEVANCE IS NOT ENOUGH TO FACILITATE SEARCH WHEN A DISTRACTOR BECOMES A TARGET.

M Pilar Aivar[1] (mariapilar.aivar@uam.es), Laura Cepero[1], Elena Sanz[1], Lidia Sobrino[1], Ángela Pérez[1], Victoria Plaza[1]; [1]Universidad Autónoma de Madrid

Repeated visual search tasks are useful to investigate under which conditions incidental memory for targets and distractors improves search. In a previous study, using letters as targets, we analyzed whether previous fixations facilitated search. Two groups of participants searched for the same 12 letters six times on the same display (72 trials). Targets were presented in a different order to each group, so that a 'critical' letter appeared either in each block, or for the first time after 55 trials. Surprisingly, first search RTs for that letter were similar in both groups, even though they differed considerably in the number of previous incidental fixations. Since incidental retention of visual information is higher for natural objects (Williams, Henderson & Zacks, 2005) we decided to replicate the study using pictures instead of letters. A single search display made of 24 different colored objects was used. Twelve objects (fruits and toys) were presented as targets and each target was searched for 6 times. We compared the results of three groups of participants. The first group found all 12 objects in each block of trials. The second group had to find the 'critical' object (an apple) for the first time after 55 trials. The third group also first searched for the apple after 55 trials, but saw a slightly different display, with a raspberry in the location of the apple, up to that point. Therefore, the three groups differed in the number of previous incidental fixations on this target object. Interestingly, we found that RTs for the first search of the apple were similar in the three groups, even though participants often made direct saccades to the target objects and could recall many details of the search display. Even when target objects were cognitively more relevant, incidental memory did not seem to improve search.

Research supported by grant PID2021-125162NB-I00.

53.429 COMPARING APPLES TO ORANGES TO BANANAS: A BIG DATA APPROACH TO UNDERSTANDING THE JOINT INFLUENCES OF STIMULUS PROPERTIES, TRIAL HISTORY, AND INDIVIDUAL DIFFERENCES

Audrey Siqi-Liu[1] (audrey.liu@gwu.edu), Emma M. Siritzky[1], Chloe Callahan-Flintoft[2], Justin N. Grady[1], Kelvin S. Oie[2], Stephen R. Mitroff[1], Dwight J. Kravitz[1]; [1]The George Washington University, [2]DEVCOM Army Research Laboratory

Research into the mechanisms of visual search has often explored how trial-level differences in the stimuli (e.g., set size, target salience) affect performance. However, performance is also influenced by trial history effects (e.g., priming, hysteresis) and individual differences (e.g., variation in task capacity). Each of these three factors (stimuli, history, individual differences) has been examined independently, but understanding their comparative importance and how they may interact is highly informative in a wide range of applied scenarios, for example in making resource allocation decisions between user-interface design, task training, or personnel selection, respectively. Unfortunately, it has traditionally been extremely difficult to evaluate the comparative contribution to task performance of each factor simultaneously, as this requires a large dataset with sufficient variation in trial-by-trial stimuli attributes, exposure to task conditions, and participant characteristics. The current study leveraged a massive dataset of visual search performance (~3.8 billion trials, ~15.5 million individuals) from a mobile game version of an airport security visual search task (Airport Scanner, Kedlin Co.) to quantify and compare the variance in performance accounted for by three factors of performance: 1) trial-by-trial stimuli search array features (e.g., current trial target identity and array set size), 2) trial history specific to each individual's experience of the task (e.g., cumulative exposure to target), and 3) individual differences in task performance aptitude (e.g., participant-specific target hit rate). Each of these three factors was found to strongly contribute to performance, but, importantly, the nature and magnitude of their influences varied. In particular, individual differences were found to be an extremely large, and relatively stronger, predictor of performance variance. Precise quantifications of each factor's comparative contribution across several task contexts are provided.

US Army Research Laboratory Cooperative Agreements #W911NF-21-2-0179, #W911NF-23-2-0210, & #W911NF-23-2-0097

53.430 CONFIDENCE IN A VISUAL SEARCH TASK DEPENDS ON TARGET VISIBILITY

Alexander C. Schütz[1] (a.schuetz@uni-marburg.de), Hui Men[1]; [1]University of Marburg

In many visual search tasks, searchers decide about when to terminate search themselves. This allows them to search exhaustively and to report their decision only once they have reached a desired level of confidence. Here, we studied what determines searchers' decision about the target's presence or absence and their confidence in this decision when they cannot control the duration of their search. We combined a time-limited visual search task with a confidence forced-choice task. Each trial contained two subsequent search intervals of 800 ms, in which participants had to search for a Gabor target in a noise background. The Gabor could be present or absent in each interval independently. At the end of each trial, participants first had to report if the Gabor was present or absent, separately for both intervals and finally to choose for which interval they felt more confident. To manipulate the difficulty of the search, the radius of the noise background was varied between 3° and 7.5° in one of the two intervals. Since the Gabor target was small and required high-acuity foveal vision, the hit rate to successfully detect it depended on the minimum distance between gaze and the Gabor during the search. When gaze did not come close to the Gabor, the hit rate and the false alarm rate were of similar magnitude and were correlated across participants, suggesting that these hits were lucky guesses rather than true detections. Confidence choices on average preferred target-present over target-absent decisions. They also depended on the minimum distance between gaze and the Gabor and they were correlated with the hit rate within participants. These results indicate that participants' decision about the target's presence and their confidence in that decision are influenced similarly by how well the target was visible.

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101001250).

53.431 DISSECTING THE MISTAKE: INVESTIGATING SEPARABLE MOTOR AND COGNITIVE PROCESSES OF POST-ERROR SLOWING

Sarah B. Malykke[1] (sarahmalykke@gwu.edu), Audrey Siqi-Liu[1], Emma M. Siritzky[1], Justin N. Grady[1], Dwight J. Kravitz[1], Stephen R. Mitroff[1]; [1]The George Washington University

People typically slow down following a mistake, a phenomenon known as post-error slowing. The causes of such slowing are not fully understood, and a key question is whether it is merely incidental, maladaptive, or strategically useful for producing subsequent correct responses. Both non-strategic (e.g., motor biases, perceptual distraction) and strategic/cognitive processes (e.g., increased response caution, variability in bias) have been proposed to contribute to post-error slowing. However, it is difficult to delineate these processes when relying on traditional response time measures (e.g., keypresses) that provide a single measure that encompasses both non-strategic and cognitive processes. The current study employed a "touch-and-swipe" response measure to better understand how individuals adjust their response timing post-error. In an object sorting task embedded in an airport security mobile game (Airport Scanner, Kedlin Co.), participants tap on objects ("touch time") and then swipe them to the top or bottom of the screen ("swipe time"). Previous work with this paradigm (Kramer et al., 2021) suggests that touch time relates more to motor responses and swipe time more to cognitive processes. Here, both touch and swipe time revealed post-error slowing—both subcomponents of total response time showed longer responses following an error. Likewise, both subcomponents were influenced by whether the swipe direction (upward or downward) was consistent with the previous trial with faster responses for repeated swipe directions. Notably, the swipe direction effect was larger for the touch than the swipe time. Finally, the swipe time—but not the touch time—demonstrated an interaction between previous trial accuracy and swipe direction wherein larger post-error slowing occurred when swipe motion repeated. The different outcomes for the touch and swipe response time subcomponents highlight the multifaceted nature of post-error slowing and reveal a distinct cognitive and strategic driver of the slowing.

US Army Research Laboratory Cooperative Agreements #W911NF-21-2-0179, #W911NF-23-2-0210, & #W911NF-23-2-0097

53.432 DISSOCIATING PARALLEL VERSUS SERIAL MECHANISMS RESPONSIBLE FOR THE INEFFICIENCY OF CONJUNCTION SEARCH - A GOLIATH TASK!

Rebecca Nako[1], Nick Berggren[1], Martin Eimer[1]; [1]Birkbeck College

Visual search is less efficient when a target is accompanied by multiple distractors with target-matching features, but the cause for these costs remains contentious. According to serial models, target-matching distractor features interfere with attentional guidance by frequently attracting attention before it is allocated to the target. According to parallel models, these features generate attentional biases concurrently with the target, thereby delaying the moment when biased competition is resolved in favour of the target object. Because these alternatives are notoriously hard to dissociate based on behaviour alone, we employed N2pc components as electrophysiological markers of attentional biases in visual cortex. In two search experiments, lateral targets that were defined by a colour/shape conjunction could be accompanied by two partially target-matching distractors on the opposite side. Any attentional capture by these opposite-side competitors (OS-Cs) should be reflected by an early N2pc with opposite polarity relative to the subsequent target N2pc. The presence of OS-Cs slowed search and delayed N2pcs to target objects. Target N2pcs were further delayed on OS-C trials with slow versus fast target reports. However, this median split did not reveal any sign of early attentional capture by OS-Cs even on slow trials. This was confirmed in Experiment 2, where the number of OS-C displays was increased to enable RT-based quartile and octile splits. Target N2pc onsets and target report delays remained closely associated, but there was no sign of any reversed early N2pc, not even for the slowest 12.5% of all trials. Our results strongly support a parallel biased competition account of attentional selectivity in visual search, but provide no evidence for serial attentional reallocation processes as the main mechanism responsible for inefficient conjunction search.

This work was funded by a grant by the Economic and Social Research Council, UK (ES/V002708/1).

53.433 IMPLICIT ASSOCIATIONS BETWEEN TARGET LOCATIONS AND AVERAGE STIMULI SIZES

Jeunghwan Choi[1] (abcdef0518@naver.com), Sang Chul Chong[1,2]; [1]Graduate Program in Cognitive Science, Yonsei University, [2]Department of Psychology, Yonsei University

People can find a target quickly using the regularity acquired from prior search experiences. For instance, previous studies showed that repeated spatial arrangements (Chun & Jiang, 1998) and spatial probabilities of a target (Geng & Behrman, 2005) could facilitate a visual search. These studies found that people could learn cues which facilitated visual search processes. In this study, we tested whether people could learn and use the task irrelevant regularity of an ensemble property to enhance a visual search performance. The stimuli consisted of 40 differently sized circles filled with oriented sinewave gratings. The target was an outlier tilted close to vertical and was tilted either clockwise or counterclockwise from the vertical orientation. The distractors were tilted close to horizontal. The task was to find the target and to report its orientation. We manipulated the average size of the stimuli to predict the location of a target. Specifically in the training phase, the hemifield of the target location was determined by the average size of presented stimuli. For example, when the average size of presented stimuli was larger than the mean size of possible stimuli size range, the target always located within a left hemifield. When the average size of presented stimuli was smaller than the mean size of possible stimuli size range, the target always located within a right hemifield. The participants did not know about these regularities. In the subsequent test phase, there were two conditions: a congruent and incongruent conditions depending on the consistency of the learned relationship between the target locations and the average sizes. We found that participants found a target more quickly in the congruent condition than in the incongruent condition. This result suggests that people can learn and use average sizes as a cue to the target location.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A2C3004133).

53.434 FORAGING IN THE WILD: DOES TARGET DENSITY GUIDE SEARCH IN INTERACTIVE LEGO-BASED FORAGING?

Marian Sauter[1]; [1]Ulm University

Foraging theory suggests that the efficiency of resource-seeking behavior is influenced by the spatial distribution and availability of those resources. Applying this framework, this study employs LEGO bricks to investigate how variations in target density influence foraging efficiency among humans. Participants were tasked with identifying specific LEGO bricks placed on trays, which represented different foraging zones with varying target densities. The primary hypothesis of the study was that target-rich zones would lead to more efficient foraging, evidenced by shorter reaction times. This hypothesis was strongly supported by our findings. This outcome aligns with the findings of earlier non-interactive visual foraging studies. Additionally, the target-rich zones were biased towards one of the sides of the search space, with the expectation that this bias can be exploited over time, resulting in decreased response times (statistical learning). Contrary to expectations, there was no consistent improvement in foraging efficiency over time: While participants were able to identify the location-bias for target-rich zones, this awareness did not consistently lead to quicker foraging, indicating that insights from computerized visual foraging might not fully translate to interactive environments. This study thereby offers a perspective on the complexities of foraging behavior 'in the wild', contributiong to our understanding of how humans adapt their search strategies in interactive environments with varying resource prevalence.

Graduate and Professional Training Center Ulm

Tuesday, May 21, 8:30 am – 12:30 pm, Pavilion

Decision Making: Perceptual decision making 3

53.435 A NOVEL BEHAVIORAL PARADIGM REVEALS THE NATURE OF CONFIDENCE COMPUTATION IN PERCEPTUAL DECISION MAKING

Kai Xue[1] (kxue33@gatech.edu), Medha Shekhar[1], Dobromir Rahnev[1]; [1]Georgia Institute of Technology

A central goal of research in perceptual decision-making is understanding the computations underlying choice and confidence. However, revealing these computations requires knowledge of the internal representation upon which the computations operate. Unfortunately, how traditional stimuli (e.g., Gabor patches and random dot kinematograms) are transformed into internal representations of evidence remains unknown, hindering the building of computational models. This study introduces a new behavioral paradigm where subjects discriminate the dominant color in a cloud of differently colored dots. Critically, we show that the internal representation for these stimuli can be described with a simple, one-parameter equation: the representation for n dots follows a Gaussian distribution with a mean of n and SD of alpha*n. In other words, the free parameter alpha controls observer sensitivity, and the SD of internal activations scales linearly with numerosity of dots. We first demonstrate that this one-parameter model explains decision data in complex, 3-choice tasks across two experiments with up to 12 conditions featuring different dot number combinations. Critically, we use this paradigm to test three popular theories of confidence: (1) the Bayesian Confidence Hypothesis (BCH), which assumes that confidence reflects the probability of being correct, (2) the Positive Evidence (PE) model, which assumes that confidence reflects only choice-congruent (i.e., positive) evidence, and (3) the Difference model, which assumes that confidence reflects the evidence difference between the highest and the second-highest signal. We find that the Difference model provides the best fit, followed closely by BCH. In contrast, the PE model provides very poor fits, trailing the Difference model by over 4,000 AIC points in both Experiments 1 and 2. These results establish a new paradigm where a single free parameter can characterize the internal representation across a potentially unlimited number of conditions, enabling the comparison of different theories of perceptual decision making and confidence.

This work was supported by the National Institute of Health (award: R01MH119189) and the Office of Naval Research (award: N00014-20-1-2622).

53.436 APPLYING EL GRECO FALLACY TO SERIAL DEPENDENCE

Natalia Tiurina[1], Yuri Markov[2], David Pascucci[3]; [1]TUD Dresden University of Technology, Dresden, Germany, [2]Goethe University Frankfurt, Frankfurt am Main, Germany, [3]Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Perceptual judgments are influenced by prior history, a phenomenon known as 'serial dependence.' It is believed that serial dependence reflects how the brain integrates perceptual priors, stemming from recent experience, with current sensory input, aligning with Bayesian and predictive coding principles. However, it is debated whether high-level perceptual priors influence lower-level perception (i.e., what is perceived) or if they impact only later processing stages (i.e., decision-making). To address this question, we employed the logic of the El Greco fallacy, which posits that systematic changes in perception should manifest in both the stimulus being perceived and the reporting tool. In each trial, participants reported whether sequentially presented pair of oriented Gabor patches (the target and test) had the same orientations. The test was either identically oriented to the target or rotated away by 8°. If serial dependence influences early vision, it should affect the perception of the target and test. Consequently, participants would more frequently judge the target and test as identical when they are identical, irrespective of the orientation shown in the preceding trial because the bias from the preceding trial would affect both stimuli similarly. Our preliminary findings suggest participants tended to choose the test rotated toward the preceding trial's target stimulus. In a second experiment, we controlled for the temporal distance between the target and the test, presenting the test at the same or different to the target location. Nevertheless, we found a consistent bias in perceptual judgments, aligning with the observations in Experiment 1. We propose that applying the El Greco Fallacy can potentially clarify the processing stages affected by serial dependence. Our preliminary findings suggest that the locus of serial dependence effects extends beyond early visual processing, offering important constraints on Bayesian and predictive processing perspectives of this phenomenon.

Funded by the Saxon State Ministry of Science, Culture and Tourism (SMWK) as part of support for profile-defining structural units of the TUD in 2023/24 (NT)

53.437 EYE-GAZE PATTERNS AND MOTIVATIONAL FACTORS INFLUENCE INTUITIVE PHYSICAL JUDGMENTS

Ren Calabro[1], Meriel Doyle[1], Kannon Bhattacharyya[1], Yuan Chang Leong[1]; [1]University of Chicago

When reasoning about the world around them, people rely extensively on their inferences about physical scenarios. Examples of how these inferences can be flawed are well-documented, including in real-world optical illusions. In the current work, we test the hypothesis that our intuitive physical judgments can change when we are motivated to see a particular outcome in a physical scenario. While undergoing eye-tracking, participants (N = 57) were presented with images of block towers that varied in stability and were rewarded for correctly judging whether each tower would fall or remain standing under the influence of gravity. We also incentivized participants (N = 29) to judge towers as stable or unstable using monetary bonuses; the bonuses did not depend on the participants' responses, so the reward-maximizing strategy was to respond accurately. We hypothesized that people would be motivated to respond that a given tower was stable versus unstable when financially incentivized to make a certain stability judgment. We found that on average, participants were more likely to judge the tower as stable when motivated to see it as stable than when motivated to see it as unstable. Moreover, we found that subjects' fixation patterns were significantly correlated with the average fixation patterns of other subjects who made the same choice, suggesting that how people sample information contributes to their judgments about physical scenes. Participants' fixation patterns were also significantly correlated with the average fixation patterns of other subjects who had the same motivation, indicating that motivation may influence how people sample information when making judgments. Further analysis will assess whether participants focus on a tower's mechanical failure points and/or look where they predict the blocks may fall when they judge a given tower to be unstable. This study may shed light on potential limitations of human physical scene understanding.

53.438 HOW HUMANS SOLVE COMPLEX PROBLEMS BASED ON VISUAL INTUITION: AN EXPLORATORY BAYESIAN FEATURE PYRAMID MODELING APPROACH

Limin Ye[1,2] (lmye@m.scnu.edu.cn), Liqiang Huang[3], Ke Zhou[4], Ming Meng[1,2]; [1]Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education, China, [2]School of Psychology, South China Normal University, [3]Department of Psychology, The Chinese University of Hongkong, [4]School of Psychology, Beijing Normal University

Humans often rely on intuition to facilitate solving complex problems swiftly. However, what cognitive mechanisms underlie intuition? Here we explore this question by modeling large-scale behavior data, i.e., over 360 million valid behavioral responses that were collected through an online game simulating the Traveling Salesman Problem (TSP) on a smartphone APP platform. TSP is a classic combinatorial optimization problem, where a salesman aims to find the shortest path traversing specified cities (nodes) and returning to the starting point. Through the TSP we investigate human decision-making capabilities in scenarios where optimal solutions are hard to compute as the number of nodes increases. The findings were three-fold. First, as the number of nodes increases, the computational difficulty of solving the TSP rises exponentially, but human efficiency in solving the problem does not decrease proportionally, suggesting that humans indeed base their choices on intuition rather than exhaustive calculation. Second, humans tend to connect adjacent nodes sequentially and avoid route crossings. Performance in this task correlates with the visual perception of the overall structure, with better scores achieved when the convex hull angle of the formed route is larger. We therefore hypothesize that in solving complex problems, individuals often plan based on visual construction of simple mental representations. Finally, we have trained a deep learning neural network to learn human algorithms, achieving a 95% accuracy in predicting human solutions. Using it as a benchmark, with the nearest-neighbor heuristic as the baseline, we simulated visual physiological structures and established an exploratory Bayesian feature pyramid model based on global centralized adaptive feature modulation. This model closely approximates the predictive capabilities of neural networks and human performance, exhibiting both robustness and interpretability. This work provides a novel framework for understanding the cognitive mechanisms of humans in solving the TSP based on visual intuition.

The research was supported by grants from the STI2030-Major Projects (2021ZD0204200), the Sino-German Center for Research Promotion (M-0705), and the National Natural Science Foundation of China (31871136).

53.439 ILLUMINATING IRRATIONALITIES IN MULTI-ALTERNATIVE PERCEPTUAL DISCRIMINATION: INSIGHTS FROM PSEUDO-OPTIMAL EVIDENCE ACCUMULATION

Kiyofumi Miyoshi[1] (miyoshi80@gmail.com), Yosuke Sakamoto[2]; [1]Kyoto University, [2]Osaka University

Our perceptual experience involves making decisions from a range of alternatives, and the development of comprehensive decision-making theories would greatly benefit relevant fields of vision science. Overcoming a notable hurdle in this pursuit involves addressing non-normative irrationalities that frequently emerge in decisions involving multiple options. This study specifically examines behavioral irrationalities in three-alternative dot-numerosity discrimination, focusing on the violation of the independence of irrelevant alternatives axiom (IIA violation) and the phenomenon known as the dud-alternative confidence boost. The IIA violation often manifests in the relative choice rate between the strongest target stimulus and the second strongest distractor stimulus; the relative choice rate fluctuates in response to the intensity of the least strong dud stimulus, which presents a significant challenge for broad families of decision-making theories. The dud-alternative confidence boost characterizes a puzzling scenario wherein the inclusion of a third alternative in the choice lineup leads to an irrationally heightened confidence level, surpassing that observed in a simpler two-choice problem. We propose a foundation for understanding such irrational choice behaviors through pseudo-optimal sequential evidence accumulation. Our approach considers independent evidence accumulation for three alternatives, where the difference between the first- and second-place accumulators serves as the fundamental decision variable. The dynamics of this decision variable naturally explain the IIA violation without the need for additional mechanistic assumptions. Furthermore, the early development of this decision variable immediately after stimulus onset effectively captures the trend of the dud-alternative confidence boost. To further extend the applicability of our model, we have combined the proposed decision algorithm with a deep neural network, which allows for the generation of testable behavioral predictions based on input image lineups. Leveraging this strength, we present a comparative analysis of human observers and neural networks regarding non-normative multiple-choice behaviors.

53.440 ORDER CONSTRAINED ANALYSES OF EYEWITNESS MEMORY ACCURACY

Andrea Yaoyun Cui[1] (yaoyunc2@illinois.edu), Meichai Chen[1], Michel Regenwetter[1]; [1]University of Illinois at Urbana-Champaign

In psychological research, there are several commonly reported patterns in the relationships between decision accuracy, confidence, and response times. Firstly, there is a negative correlation between the similarity of the target and lures and response accuracy, as noted in studies like Huang et al. (2021). Secondly, studies report a positive relationship between confidence and accuracy, as mentioned in research by Brewer & Wells (2006). Thirdly, response time and accuracy are inversely related, as reported in De Boeck & Jeon's (2019) work. We submit these regularities to a novel test using data from Horry and Brewer (2016)'s Experiments 1 and 2. In these experiments, participants first studied a target face, then tried to identify it from two items (Experiment 1) or four items (Experiment 2), and rated their confidence. We employed order-constrained inference using the QTEST software (Regenwetter et al., 2014; Zwilling et al., 2019). This approach translates verbal hypotheses into testable statistical ones without introducing unnecessary assumptions. At the individual participant level, we quantified evidence in favor or against each of the three hypotheses using Bayes factors (against an unconstrained baseline). We also assessed each of the three hypotheses jointly across participants using group Bayes factors. The Bayes factors obtained from the analyses indicated "decisive" evidence in support of the three hypotheses in Experiment 1 and "strong" to "very strong" evidence against the three hypotheses in Experiment 2, with notable individual variations. The lack of support in Experiment 2 might not be a flaw of the hypotheses themselves but could stem from the stimulus design or display settings not effectively reflecting the intended manipulation. This highlights the need for careful testing of stimulus design as well as display settings in visual attention research.

53.441 STIMULUS DISTRIBUTIONS AFFECT UNCERTAINTY SAMPLING APPROACHES TO ADAPTIVE ESTIMATION OF CLASSIFICATION IMAGES

Rabea Turon[1] (rabea.turon@tu-darmstadt.de), Lars Reining[1], Thomas S. A. Wallis[1,2], Frank Jäkel[1]; [1]Technical University of Darmstadt, [2]Centre for Mind, Brain and Behaviour (CMBB), Universities of Marburg, Giessen and Darmstadt, Germany

In binary decision tasks, it is often assumed that humans perform template matching by comparing a presented stimulus to a template in their mind. In effect, the template matching process defines a decision boundary in stimulus space. Many different methods have been proposed to measure decision boundaries and relevant features (reverse-correlation, bubbles, sparse regression, etc.), but to work well, all require exhaustive experimental testing. One way to measure decision boundaries efficiently is to adaptively sample stimuli to be maximally informative. A popular strategy for adaptive sampling in machine learning is to use "uncertainty sampling", where stimuli are chosen that are close to the current estimate of the decision boundary (e.g. as used in active learning for support vector machines). Here we show that this strategy can fail to adequately constrain the decision boundary, producing biased estimates compared even to random sampling, depending on the stimulus distribution and observer model. We use simulations in two dimensions to show that when class distributions are far apart in the stimulus space, uncertainty sampling repeatedly samples only a few stimuli close to the decision boundary, even for deterministic observers. This poorly constrains the direction of the decision boundary (a problem that will become more acute in higher dimensions). Uncertainty sampling does work when stimulus distributions densely populate the area around the decision boundary. However, in some psychophysical settings this might not be possible (for example, in cases where only a single fixed pool of stimuli is available such as sampling faces from a limited set of identities). We propose a solution to this problem by minimizing the entropy of the posterior distribution over the parameters of a multidimensional psychometric function. This implicitly constrains uncertainties over the whole stimulus space, and samples informative samples across the available stimulus pool.

Funded by the Hessian research priority program LOEWE within the project "WhiteBox".

53.442 THE IDENTIFIABILITY OF BAYESIAN MODELS OF PERCEPTUAL DECISION

Xue-Xin Wei[1] (weixxpku@gmail.com), Michael Hahn[2]; [1]UT Austin, [2]Saarland University, Saarbrücken, Germany

Introduction: Inferring the underlying computational processes from behavioral measurements represents a fundamental approach in psychology, neuroscience, and cognitive science. This problem may be ill-posed in that different neural processes and implementations can give rise to the same behavior. In perceptual science, an important normative modeling framework is the Bayesian observer model, which models perceptual decisions in terms of the likelihood function, the prior belief, and the loss function. Surprisingly, it remains unclear to what extent these modeling components can be recovered from behavioral data. Method: Using a combination of theory, numerical simulation, and analysis of behavioral data, we systematically investigate the problem of inferring components of Bayesian observer models from psychophysical tasks, including both continuous estimation and two-alternative forced choice (2AFC) tasks. Results: Our theoretical results guarantee in-principle identifiability under broadly applicable conditions, without any a priori knowledge of the prior distribution or stimulus encoding. In particular, when the loss function is known, the prior and encoding can be systematically identified. Interestingly, prior and loss function can be systematically confounded when the loss function is unknown. Importantly, this can be resolved by having behavioral responses based on multiple noise levels. We also find that identifiability can be achieved through the manipulation of either internal noise or stimulus noise. Numerical simulations and applications based on published behavioral datasets validate that the predictions of this theory apply in realistic settings. Conclusion: Our work leads to a systematic characterization of the identifiability of a large family of Bayesian observer models. Crucially, our results demonstrated that reliable recovery of the model components often requires having data from multiple noise levels. Our results have broad implications in interpreting existing data and in designing future psychophysical experiments.

53.443 NEW LOOK AT COMPUTING META-D' FOR CONFIDENCE JUDGEMENTS

Amélie Goertz[1] (goertz.cm.amelie@gmail.com), Ádám Koblinger[1], József Fiser[1]; [1]Central European University

In the standard signal detection (SDT) framework, assuming that observers' decisions (Type 1 decision) and confidence (Type 2 decision) about decisions are based on the same information, their sensitivity (d') to sensory stimulus can be assessed both from decision accuracy and confidence reports. However, ample empirical evidence indicates that d' derived from these two sources are not equal suggesting that Type 1 and 2 decisions rely on at least partially distinct information. Based on this insight, several studies explored the use of d' derived from Type 2 decisions independently (meta-d') to characterize metacognitive performance. The resulting single core algorithm used by most popular methods is built to follow the logic of standard SDT without explicitly defining a normative framework. By developing a normative generative model of metacognition and through theoretical analyses and simulations, we found that the core algorithm does not fit the natural extension of the classical SDT-based generative model. It provides correct measures -according to the natural extension- only in the case when no added or subtracted noise is assumed during the confidence judgment compared to the decision stage, i. e. when d' = meta-d'. For example, at a typical value of d'=1.16 if meta-d' deviates from d' by 10%, the core algorithm will predict as much as 30% deviation. As a result, using the core algorithm eliminates the rigorous link between the descriptions of Type 1 and Type 2 decisions and in turn, the fundamental logic of the M-ratio-based metric using meta-d'/d' is called into question. In contrast, our analysis also provides a computational method of meta-d' that restores the link while it adheres to the normative generative framework. In conclusion, we identified a significant flaw in the popular method of treating Type 2 decisions and provided a normatively justified algorithm for assessing metacognitive performance.

53.444 INTEGRATING VISION AND DECISION-MAKING MODELS WITH END-TO-END TRAINABLE RECURRENT NEURAL NETWORKS

Yu-Ang Cheng[1] (yuang_cheng@brown.edu), Ivan Felipe Rodriguez[1], Takeo Watanabe[1], Thomas Serre[1,2]; [1]Brown University, [2]Carney instititue for Brain Science

Historically, research in visual perception and perceptual decision-making has been pursued independently. Models of visual perception have primarily focused on developing neurocomputational mechanisms for visual processing, particularly in object and face recognition. These models, however, largely approximate only human accuracy levels, not fully utilizing reaction time data. In contrast, decision-making models have sought to replicate both accuracy and reaction times in human behavior, but they do not adequately address underlying visual processing mechanisms. Here, we bridge this gap and introduce an integrated end-to-end trainable recurrent neural network model. First, we optimize a vision module, a convolutional neural network, for a well-known perceptual decision-making task, i.e., the random dot motion task (Britten et al., 1992). We show that fitting a straightforward nonlinear reaction time function (Goetschalckx et al., 2023) to the vision module outputs fails to capture the distributions of human reaction times for the same task. However, fitting the drift-diffusion model (Ratcliff & Rouder, 1998), a traditional cognitive model significantly improves the goodness of fit. We further turn to a discrete-time recurrent neural network (RNN) approximation of the Wong-Wang circuit (Wong & Wang, 2006) for decision-making, which we optimize end-to-end together with the vision module using human behavioral data. We show that this combination offers a better fit for experimental data. In addition, analyzing the weights of the resulting model yields novel insights about the underlying integration process's time course and the image features driving these decisions. Our integrated RNN model of vision and decision-making represents a first step towards a complete computational model of perceptual decision-making.

Tuesday, May 21, 8:30 am – 12:30 pm, Pavilion

Temporal Processing: Duration, atypical, timing perception

53.445 ENDOGENOUS ATTENTION DOES NOT PROLONG THE SUBJECTIVE DURATION OF PERIPHERAL STIMULI - DIRECTING ATTENTION IN THE OPPOSITE DIRECTION SHORTENS IT

Alina Krug[1] (alina.krug@uni-ulm.de), Lisa Eberhardt[1], Anke Huckauf[1]; [1]Ulm University

A number of studies have shown that increasing stimulus eccentricity decreases the perceived duration of short stimuli, yet increases for longer stimuli. This might be due to endogenous shifts of spatial attention towards the peripheral stimulus, that have been shown to prolong perceived duration. Aim of the present study is to investigate whether endogenous attentional shifts affect the perceived duration of peripheral stimuli. In a temporal bisection task, participants categorized the duration of a short peripheral stimulus (20-220 ms) presented at 3° or 9° of eccentricity left or right from fixation as short or long. An eyetracker ensured constant fixation. In Experiment 1, a central arrow cue either reliably indicated the position of the subsequent peripheral stimulus (valid cue) or did not convey information regarding its position (neutral cue). In Experiment 2, the cue either correctly indicated the position of the subsequent stimulus in 75% of the trials (valid cue) or pointed towards the opposite direction (invalid cue). For both experiments, perceived duration decreased for stimuli presented at 9° of eccentricity compared to 3°, replicating previous studies. With regard to cueing, reaction times were significantly shorter for the valid cues compared to the neutral (Experiment 1) and invalid cues (Experiment 2), indicating that the attentional manipulation was successful. However, a significant difference in perceived duration was only found in Experiment 2, where validly cued stimuli were perceived as longer than invalidly cued stimuli. Among others, the results are discussed with regard to the notion that directing attention to a stimulus may not increase perceived duration, but that directing attention away from it may decrease it.

53.446 A STUDY OF CRITICAL FUSION FREQUENCY AND DUTY RATIO WITH MULTIPLE LIGHT STIMULI

Kotaro Oikawa[1], Ruggero Micheletto[1]; [1]Yokohama City University

In recent years, many studies have focused on the importance of the temporal domain in the visual system. Critical fusion frequency (CFF) is the threshold value at which a flickering stimulus appears to be a steady stimulus. Therefore, the CFF has been extensively used to evaluate the visual temporal processing. However, there are unexplained aspects of this perceptual phenomenon. Our previous results have shown that by measuring the duty ratio dependence of CFF in detail, the profile of the CFF has an asymmetric characteristic curve. In this study, we examined the effect of multiple flashing stimuli on the CFF-duty ratio profile. In our experimental approach, CFF values were measured for each duty ratio using a 3x3 grid of flashing LEDs. An analog signal output device was used to drive them and control the flashing frequency, switch the duty ratio, and record the perceived CFF threshold value. The 3x3 grid stimuli are also synchronized by being driven by a transistor circuit. As a result, we observed that the asymmetry of the CFF-duty ratio profile tended to mitigate to symmetry in the 3x3 grid stimulus compared to previous experiments with a single LED flashing stimulus. We quantified the degree of symmetry by normalizing the measurements in the two presented stimuli and found that the 3x3 grid stimuli were approximately the same or more symmetric than the single stimulus. Specifically, for the 3x3 grid stimuli, a decrease of about 5% in the normalized CFF value, a factor affecting symmetry, was observed for curve duty ratios ranging from 20% to 40%. Furthermore, this result fits our theoretical model and the symmetry parameter appears in our mathematical formalism reported in our study.

53.447 A NOVEL TASK TO MEASURE TEMPORAL INTEGRATION IN THE HUMAN VISUAL SYSTEM

Michele Deodato[1] (md5050@nyu.edu), David Melcher[1]; [1]New York University Abu Dhabi

Visual temporal integration refers to the brain's ability to combine and process visual information over time. This mechanism plays a crucial role in various aspects of visual perception, including motion detection, object recognition, and the perception of temporal patterns. Understanding the temporal integration capabilities of the human visual system provides insights into how our brains construct a unified representation of the visual world from discrete visual inputs. Here, we developed a new task to measure visual temporal integration. Participants were presented with the image of a horizontal sine wave pattern. On each frame the image was recreated and random noise added. The task required the merging of consecutive frames presented in rapid succession, in order to discount the noise and construct the percept of a coherent sine wave pattern. The amount of noise added was systematically varied along the image's vertical axis and participant were asked to report the vertical position at which they could no longer see the pattern in the flickering noise. This value indicates the amount of information that can be integrated from successive frames. Importantly, on separate trials the duration of the frames was varied from 2 to 1000 ms to obtain a temporal threshold of integration. We found that the visual integration for each frame duration followed a sigmoid function that reached an asymptote at around 120 ms, consistent with previous reports of the temporal dynamics of the human visual system. Notably, some participants showed a sharp decrease in performance when the frames duration was ~100 ms. This effect is consistent with visual entrainment of the brain alpha rhythm to the stimuli, providing further evidence for the role of the alpha rhythm in visual temporal integration. Indeed, alignment of alpha integration cycles with single images would prevent integration of multiple stimuli and impair visual performance.

53.448 TEMPORAL SEGMENTATION PRINCIPLES IN VISION AND AUDITION

Linda Garami[1] (garamil@ceu.edu), József Fiser[1]; [1]Central European University

Our sensory system automatically and unconsciously extracts patterns at varying complexity from the environment and these patterns bias sensitivity to changes, accuracy during recognition and also guide attention. A growing body of evidence shows that sensory integration starts at low-level cortical areas as neural correlates of temporal integration (chunking) are already present in the primary sensory cortices. Based on this, we hypothesized that certain chunking principles and the resulting perceptual biases are similar across modalities. To test our hypothesis, we focused on a well-known auditory chunking principle called the Iambic-Trochaic law (ITL). Established in language processing, ITL says that longer syllables in a sequence signal word ends and similarly, a longer duration of a tone is understood as a closing element of a perceptual chunk. ITL is considered a basic law since it has been found across multiple species, but it is unknown whether it generalizes across modalities. We implemented a Short-Short-Long stream segregation go/no-go paradigm for human participants in an identical manner in the visual and auditory modalities and found that although the general performance was lower in vision, biases in sensitivity to deviations showed the same bias in the two domains (N=17). Extending the paradigm to a Short-Long-Long stream, we replicated the basic effect while clarifying that this sensitivity bias could neither be explained by the repetition of individual elements nor by the absolute duration of the individual elements alone in either sensory domain. Instead, this bias depended on an unconscious chunking process that integrated effects related to both duration of the individual elements and the internal structure of the stream such as the number and arrangements of the long elements. Our results support the existence of domain-general non-linguistic grouping principles incorporating structural features of ever-increasing complexity that could ultimately give rise to high-level chunks leading to object concepts.

This work has been supported by the grant of CEU-ITI.

53.449 DETECTION AND IDENTIFICATION OF ONE-DIMENSIONAL NOISE STIMULI: EFFECTS OF TEMPORAL SPECTRUM

Annabel Wing-Yan Fan[1] (annabel.fan@mail.mcgill.ca), Alex S Baldwin[1]; [1]McGill

Although contrast detection models can also account for stimulus identification performance, neuroimaging and behavioural studies have found partially distinct mechanisms underlying the tasks. We examined the effect of stimulus duration on detection or identification tasks with filtered dynamic noise targets and the interaction between the temporal spectrum of the dynamic noise and the benefit of temporal integration. Detection and identification (left- or right-oblique target) thresholds were measured for one-dimensional spatially pink noise presented for 50, 200, or 1000 ms in a circular envelope. The noise was either static or dynamic. In the dynamic case, it had either a white or pink temporal spectrum. In the detection task, a noise target appeared in either the first or second interval (with the other interval blank). In the discrimination task, both intervals contained a noise stimulus, one being left- and the other right-oblique. The observer reported which interval contained the left-oblique target. Ten participants with healthy vision performed all experiment conditions. Thresholds were analysed with a three-way rm-ANOVA (task x temporal noise x noise duration). Threshold RMS contrast significantly varied depending on the task ($F_{1,9} = 16.7$, $p < 0.01$), temporal spectrum ($F_{2,18} = 21.0$, $p < 0.0001$), and duration ($F_{2,18} = 246.1$, $p < 0.0001$). Identification thresholds were lower than detection thresholds. Thresholds also decreased with stimulus duration. Comparing between constant and dynamic noise conditions, thresholds were highest for stimuli with white temporal spectra, followed by pink and constant noise. We find that temporal integration improves sensitivity with increasing stimulus duration, though for white noise it is likely that linear combination over short durations reduces the effective contrast. Although we find greater sensitivity in the identification task, we do not find any differences in the effects of the temporal aspects of the stimuli compared to simple detection.

This work is funded by the Vision Health Research Network of the FRQS, FRQS, Fiera Capital Awards for Diversity, Equity, and Inclusion in Health Care, and an NSERC Discovery Grant.

53.450 THE ROLE OF ALPHA OSCILLATIONS IN PERCEIVING VISUAL DURATIONS

Audrey Morrow[1] (aumorrow@ucsc.edu), Montana Wilson[1], Michaela Geller-Montague[1], Sara Soldano[1], Sabah Hajidamji[1], Jason Samaha[1]; [1]University of California Santa Cruz

An individual's alpha-band oscillation (IAF) has long been proposed as a temporal clocking mechanism for estimating event durations, however mixed findings and limited use of task types and duration ranges have hampered strong conclusions. This EEG study measured performance on a battery of timing tasks to evaluate whether IAF is associated with the estimation and discrimination of visual event durations. In a temporal estimation task, participants reported the duration of a single stimulus presented between 300-1200ms. In a temporal discrimination task, participants were asked to report which of two stimuli was longer: a standard stimulus (100, 600, or 1200ms) or a comparison stimulus (50-150% of the standard). Stimuli also varied in whether they were constant luminance ("static" condition) or varied randomly in luminance within a trial ("dynamic" condition). Finally, the critical flicker frequency (CFF) was measured since several studies have shown the CFF to be correlated with IAF in clinical populations. We found that the average absolute error in duration estimates was moderately negatively correlated with IAF, suggesting that higher IAF is related to better accuracy when estimating visual durations. Psychometric functions fit to the discrimination task data revealed IAF had a low to moderate positive correlation with slope parameter, indicating that higher IAF is related to greater sensitivity in duration discrimination, with stronger effects for static stimuli. Interestingly, the CFF was moderately correlated with duration discrimination sensitivity for the longer standards (600ms and 1200ms), despite a lack of overall correlation between IAF and CFF. These results seem to suggest that IAF plays a role in duration perception, but to different degrees depending on the type of stimuli. Further analyses will measure alpha frequency at the trial level to examine whether moment-to-moment fluctuations in IAF bias duration perception.

53.451 THE HUMAN TEMPORAL DELAY FUNCTION

Callista Dyer[1] (callista@sas.upenn.edu), Johannes Burge[1,2,3]; [1]Department of Psychology, University of Pennsylvania, PA, [2]Neuroscience Graduate Group, University of Pennsylvania, PA, [3]Bioengineering Graduate Group, University of Pennsylvania, PA

Two major aspects of temporal processing are temporal blur (or smear), which determines sensitivity, and temporal delay, which sets the effective latency of processing. Temporal sensitivity is well-known to improve with eccentricity and follow an inverted U-shaped function with spatial frequency. Much less is known about temporal processing latencies. Here, with data having sub-millisecond resolution, we present the human temporal delay function, which shows how processing latencies change with spatial frequency across the visual field. We developed stimuli that leverage a visual illusion called the Pulfrich effect. While subjects fixated, stimulus-induced interocular delays caused the rotating stimulus to appear tilted top-back or bottom-back relative to the screen. The task was to report the apparent tilt. Like Kelly (1984), we used different stimulus-types to probe processing: (i) a rotating annulus of eight radial Gabors, (ii) a rotating ring textured with bandpass-filtered 1/f noise, with passbands matched to the Gabors, and (iii) a set of Gabors with static envelopes and drifting carriers. (Different stimuli can identify processing (in)efficiencies and support disparate experimental goals.) We measured processing delays across a 4.5-octave spatial frequency range (0.33-8cpd), at eccentricities ranging from 1.0deg to 6.0deg. The speed of rotation (2-12deg/sec) and the stimulus width(s) (20-120arcmin) scaled with eccentricity. Processing delays decrease dramatically with eccentricity, change systematically with spatial frequency, and are consistent across stimulus-types. Further, sensitivity is described by inverted U-shaped functions similar to those characterizing previous work. Unlike response-time measures—which are noisy, often influenced by decision biases, and always impacted by motor contributions—the current paradigm has the advantage that temporal delays manifest as stereo-depth effects, so subjects need not make explicit estimates of any temporal aspect of the stimulus. The current results show how processing delay changes with spatial frequency across the visual field, a fundamental but understudied aspect of visual processing.

This work was supported by the National Eye Institute and the Office of Behavioral and Social Sciences Research, National Institutes of Health Grant R01-EY028571 to J.B.

53.452 DISCRIMINATING CORTICAL RESPONSES TO DIFFERENT VISUAL FLICKER FREQUENCIES IN DYSLEXIA

Josiane Mukahirwa[1] (hijosie@udel.edu), Qianli Meng[1], Keith Schneider[1]; [1]University of Delaware

Introduction:The magnocellular system theory of developmental dyslexia suggests a pervasive dysfunction of the magnocellular system throughout the brain. Therefore, we sought to test the distribution of magnocellular cortical input by measuring brain responses to visual flicker in dyslexic subjects. We used MVPA to discriminate different frequencies and compare the results between dyslexic and normal readers. Methods: Using fMRI at 3T, we imaged whole-brain responses in 8 subjects with dyslexia and 11 normal readers as they observed epochs (3–6 s) of black and white checkerboards, contrast reversing at frequencies from 2–60 Hz. After preprocessing (motion correction, temporal filtering, and smoothing), we performed a trial-wise estimate for individual subjects. The results were used to perform multi-voxel pattern analyses (MVPA). A support vector machine classifier was used for a whole brain searchlight analysis to explore brain areas that displayed significant discrimination probabilities between the lowest and highest frequency bins. Results: The normal readers displayed high discrimination in the visual cortex, in particular, left V5/MT (FWE, $p < .05$, TFCE), while individuals with dyslexia exhibited discrimination only in the posterior visual cortex and not in V5/MT. The between-group difference was not significant in our small sample. Discussion: The results suggest that the subjects with dyslexia may have weaker magnocellular input to visual cortex. In future analyses we plan to examine between-group classifiers and region of interest analyses.

53.453 HIGH PERCEPTUAL LOAD ATTENUATES THE MAGNITUDE OF INTENTIONAL BINDING FOR AUDIOVISUAL EVENTS

De-Wei Dai[1], Po-Jang (Brown) Hsieh[1]; [1]National Taiwan University

The intentional binding (IB) effect is the perceived shortening of time between a voluntary action and its sensory outcome. This study investigates the IB effect for different audiovisual sensory outcomes. Experiment 1 examines whether the perceived causality of sensory events affects the IB effect's magnitude. It employs a 2 x 2 repeated measures design: Action (Passive/Agency) and Event (Collision/Pass), using a modified IB task (Libet clock). We discovered that the IB effect for the sensory outcome is eliminated when the outcome includes both audio and visual modalities, regardless of perceived causality. The second experiment aims to determine whether this attenuation is due to multisensory integration or the divided attention nature of the paradigm. Experiment 2 employs a 2 x 3 repeated measures design: Action (Passive/Agency) and Event (Audio Only/Audiovisual Integrated/Audiovisual Irrelevant). The focus is on whether the IB magnitude (i.e., the difference between Passive and Agency) changes with the type of sensory event. Participants' primary task is to judge the onset of a collision sound. In the Audiovisual conditions, participants also assess the trajectory of two colored discs on the screen. In the Audiovisual Integrated condition, two discs launch toward center fixation and bounce away upon contact, with a simultaneous collision sound depicting a causal event. In the Audiovisual Irrelevant condition, two discs launch in pseudo-randomized directions, unrelated to the target sound. Results show that the IB magnitude is significantly weaker in both Audiovisual Integrated and Audiovisual Irrelevant conditions compared to Audio Only. The attenuation in the Audiovisual Irrelevant condition suggests that the effect is due to high perceptual load rather than multisensory integration.

53.454 LIGHT ADAPTATION HAS SIMILAR EFFECTS ON RESPONSE TIMES GENERATED BY THE L-, M-, AND S-CONES, FOR BOTH INCREMENTS AND DECREMENTS.

Yesenia Taveras-Cruz[1] (taverascruz.y@northeastern.edu), Rhea T. Eskew, Jr.[1]; [1]Northeastern University

Visual system response kinetics change based on the adapting ambient light level, test stimulus properties (e.g., test polarity), and how a test stimulus is being processed by post-receptoral mechanisms (e.g., summation or differencing of the L-, M-, and/or S-cones). Light adaptation shortens the time constants of the photoreceptor response, reducing integration time, regulating sensitivity of the visual system: the visual system speeds up as a function of light level. Beginning in the retina at the cone-bipolar synapse, visual information is organized into parallel systems, signaling either a light increment (ON) or light decrement (OFF); the OFF pathway has been shown to be faster. We studied the effects of light adaptation as a function of test polarity and cone class by measuring response time (RT) distributions at low (0.31 or 1.21 Log Td) and high (4.13 Log Td) adapting background luminances. Stimuli were 2X threshold two-dimensional blurred rectangles calculated to be cone-isolating increments or decrements, to selectively modulate individual cone classes and activate either ON or OFF pathways. We found that: (a) The visual system speeds up similarly for all cone classes; average median RT decreased by ~90 msec as light level increased by ~3.4 log units. (b) Decrement RTs were faster than increment RTs by ~18 msec at both light levels. (c) At both light levels, latency is ~3 times longer for S-cones compared to the L- and M-cones. The simplest conclusion is that the speeding up of the visual system as luminance increases is predominantly due to changes at the photoreceptor level, such that increments and decrements see a similar decrease in RT. Despite the dramatically different post-receptoral pathways carrying S-cone signals, the temporal advantage of decrement perception occurs for S- as well as L- and M-cones, likely stemming from the faster kinetics of the OFF bipolar cells.

Supported by NSF BCS-2239356

53.455 SPEED AND TIME DISCRIMINATION WITH BOUNCING DISK STIMULI

Anthony Bruno[1] (anthony_bruno@brown.edu), Leslie Welch[1]; [1]Brown University

Our previous work showed that a bouncing disk, a moving stimulus that accelerates and decelerates, improved precision on a time discrimination task relative to when time intervals were defined by discrete visual flashes. However, time interval and speed were perfectly correlated, so participants could have used either information to respond correctly. To disentangle the relationship between speed and time with our stimuli, we measured participants' speed and time discrimination performance with bouncing disk stimuli akin to those used in our prior research. In this study, participants completed two tasks. In the speed task, participants reported whether a test stimulus moved faster or slower relative to the standard stimulus (15 deg/s). Since the stimuli accelerated and decelerated, their average speed was used as the criterion for the speed discrimination. In the time task, participants reported whether a test time interval, defined as the time between the disk's bounces, was shorter or longer relative to the standard time interval (550 ms). Crucially, all bouncing disk stimuli shown to participants were identical in both tasks. The speed and time dimensions were sampled relatively equally by varying the height of the disk's trajectory. For all participants, precision (threshold) in the speed discrimination task was at least a factor of three worse than in the time discrimination task. This result suggests that participants in our previous experiment were indeed making time judgments, not speed judgments. Further, the bouncing disk improved time discrimination despite it not providing good information for speed discrimination. A high quality time representation from a moving stimulus was achieved without a high quality speed representation.

53.456 THE EFFECTS OF PHYSICAL EFFORT ON TIME PERCEPTION

Li Yang[1] (lyang147@ucr.edu), Rachel Francis[1], Hyung-Bum Park[2], Weiwei Zhang[1]; [1]University of California, Riverside, [2]University of Chicago

The interaction between time perception and physical effort plays a crucial role in our daily activities, especially in contexts such as sports where precise timing is essential. The present study therefore aims to assess the effects of a simple(est) physical effort (i.e., isometric handgrip) on time perception with concurrent time reproduction and handgrip tasks. Specifically, the participants experienced a lapse of time at various levels while concurrently exerting grip force on a hand dynamometer at levels proportional to their maximum voluntary contraction (MVC). Immediately after, they pushed buttons to reproduce the perceived time interval. On the one hand, isometric handgrip can induce physical arousal, thereby accelerating time and leading to overestimated durations (i.e., arousal hypothesis). On the other hand, handgrip may directly impair time perception given the potential competition for magnitude-based processes underlying the two concurrent tasks, resulting in underestimated durations (i.e., Magnitude hypothesis). Our preliminary data revealed a significant increase in estimated duration under high physical effort compared to lower physical effort, supporting the arousal hypothesis. Some alternative accounts (e.g., response bias and effects on motor responses) were ruled out given that the handgrip effect was absent in a time production task where the required time duration was indicated by a numerical cue, instead of the actual passage of time. Together these findings suggest that physical effort can distort perceived time.

53.457 TEMPORAL DYNAMICS GAP BETWEEN POSITION TRACKING AND ATTRIBUTE TRACKING

Yen-Ju Chen[1] (ra51011daniel@gmail.com), Zitang Sun[1], Shin'ya Nishida[1,2]; [1]Cognitive Informatics Lab, Graduate School of Informatics, Kyoto University, Japan, [2]NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, Japan

Continuous tracking is a promising psychophysical paradigm to quickly estimate the temporal dynamics of the visuomotor system. It asks observers to mouse-track a randomly changing target spatial position (position tracking: PT) or attribute value (attribute tracking: AT). The impulse response can be computed from the cross-correlation in velocity between stimulus and response. To understand the visuomotor mechanisms underlying continuous tracking, we examined the effects of task (PT or AT) and stimulus. We used eleven stimuli: a Gaussian blob changing either in position (for PT) or in luminance (for AT); a color-modulated concentric Gabor (red-green or blue-yellow) changing in position or chromatic contrast; a luminance-modulated concentric Gabor (one of eight spatial frequencies) changing in position or luminance contrast. The results show that stimulus parameters affected the estimated impulse response. PT and AT responses were weakly correlated with regard to the effects of stimulus in either peak latency or impulse response width, which suggests a partial commonality of the underlying mechanism. However, the AT response was approximately 2.6 times slower in latency and 5.0 times broader in width than the PT response. Where does this difference come from? It cannot be ascribed to rapid automatic vision-hand mapping in PT since even when compared with an anti-PT condition where observers had to move the mouse in the opposite direction of the visual stimulus movement, the AT response was 1.7 times slower and 2.3 times broader. Another task difference is that PT does not need to evaluate the attribute value, but AT does. To test this factor, we measured the simple reaction time to the detection of a single stimulus change and found it agreed well with PT's peak latency. The results suggest that slower responses for AT than for PT can be ascribed to the extra process to access the attribute content.

Supported by JSPS Kakenhi 20H00603, 20H05957.

Tuesday, May 21, 8:30 am – 12:30 pm, Pavilion

3D Perception: Size, shape, distance

53.458 A TWO-TIERED MODEL OF THE PERCEPTION OF ASPECT RATIO IN BINOCULARLY VIEWED SURFACES

Andrew Rubner[1] (andrew.rubner@gmail.com), Manish Singh[1]; [1]Rutgers University

Standard Bayesian models of the perception of aspect ratio in binocularly viewed surfaces predict that observers will accurately perceive the aspect ratio of a surface given reliable binocular-disparity cues to slant. However, there is psychophysical evidence that this prediction is not borne out. Specifically, Hibbard et al. (2012) showed that observers underestimate the aspect ratio of surfaces by about 13% when the binocularly specified surface slant is 0. This has been taken to show that the standard Bayesian model is inadequate. We agree but argue that a different kind of Bayesian model can explain the recalcitrant data. We propose a two-tiered model based on an explanation of previous data that shows that line length is (mis)perceived as a function of image orientation (Howe and Purves 2002). According to that argument, the misperception of line length is driven by the correlation between line image orientation and its 3D slant (from natural scene statistics). In our model, the misperception of length in the vertical direction leads to an overestimation of aspect ratio in the first tier. Moreover, this first tier is informationally encapsulated insofar as the bias it introduces cannot be corrected by cues to slant available to higher-level modules. The information from the first tier is then fed into a second tier that uses Bayes' theorem to compute an estimate of surface aspect ratio given the biased image aspect ratio plus binocularly specified slant. This two-tiered model can explain why observers underestimate surface aspect ratio despite having binocular-disparity cues to slant. The model thus shows how percepts that appear inconsistent with Bayes can nevertheless be understood in terms of a hierarchy of modules, even if the processing within each module is Bayesian. This model also explains how percepts that are not projectively consistent with the image data are possible.

53.459 CONTOUR INTERACTION, SPATIAL EXTENT, AND SIZE CONTRAST IN A VISUAL SIZE ILLUSION

Ryan E.B. Mruczek[1] (rmruczek@holycross.edu), Botamina Girgis[1], Emma Stowell[1]; [1]College of the Holy Cross

Previous research has documented multiple contextual influences on perceived size in classic visual illusions (e.g., Ebbinghaus and Delboeuf), including contour interactions (e.g., Todorović and Jovanović, 2018, Acta Psychologia), spatial extent (Kirsch and Kunde, 2021, Vision Research), and size contrast (e.g., Roberts et al. 2005, Perception). Such studies typically focus on one factor, as multiple factors are difficult to disentangle (but see Roberts et al., 2005). However, the postulated contextual mechanisms are not mutually exclusive. We sought to quantify the relative contributions and potential interactions of size contrast, contour interactions, and spatial extent on illusory changes in perceived size. Participants reported which of two squares ("standard" or "target") appeared larger. The standard (1.5°) was either isolated (control) or surrounded by four inducers, each composed of a single parallel line and a single perpendicular line (forming the shape of plus or a tack). Across five experiments, we systematically manipulated the separation between the central square and (1) the perpendicular line, (2) the innermost point of the parallel line, and (3) the outermost point of the parallel line (1.5°, 3.25°, of 5° in all cases). This combination of stimulus configurations allowed us systematically explore changes in contour interactions (between parallel edges of the target and inducers), spatial extent (defined as the outermost extent of the inducers), and size contrast (based on the overall size/length of the inducers). The point of subjective equality (PSE) was extracted from psychometric curves (target square size range: 1.2° to 1.8°). We found two main factors influencing perceived size: (1) contour interactions for perpendicular contours near the target made the central target appear larger and (2) larger spatial extents made the central target appear smaller. Overall, our results are consistent with a contour interaction and spatial extent influencing perceived size, with no specific contribution from size contrast.

Department of Psychology and Neuroscience Program, College of the Holy Cross

53.460 THE ROLE OF UNCERTAIN PERSPECTIVE INFORMATION IN RECOVERING 3D SYMMETRICAL SHAPES

Mark Beers[1], Zygmunt Pizlo[1]; [1]University of California, Irvine

It is known that 3D shape reconstruction from a single 2D perspective image of a symmetrical object is unique, but 3D reconstruction from an orthographic image yields a one parameter family of possible 3D shapes. Last year, we showed that subject's perceived 3D shape from a perspective image tended to be closer to veridical when compared to reconstruction from an orthographic image. In our new experiment, we tested how the human visual system incorporates perspective information. On each trial in the experiment, a perspective image is shown on the left side of a monitor, and on the right side of the screen a subject-adjustable 3D shape is shown rotating. The subject adjusted the aspect ratio of the mirror-symmetrical rotating 3D shape on the right until the 3D reconstruction was as close as possible to their 3D percept produced by the static 2D image on the left. From trial to trial, object shown and simulated distance were adjusted. As the distance of an object becomes greater and the retinal image becomes smaller, perspective information becomes less reliable, and the subject's percept becomes less veridical. Within this framework, we tested two classes of objects: (i) synthetic symmetrical polyhedral objects and (ii) symmetrical or approximately symmetrical, real-world objects selected from the ModelNet-40 dataset. The computational model uses regularization to perform 3D shape reconstruction. The cost function has terms that penalize departure from 3D compactness of the 3D shape's convex hull. In addition, the cost function incorporates a measure of the reliability of perspective information. The cost function biases the 3D reconstruction towards veridicality when perspective information is reliable. With subjects, reconstruction of natural objects is more accurate than reconstruction of synthetic polyhedral objects. This result suggests the presence of additional shape constraints, such as a second mirror symmetry.

53.461 TOWARD A THEORY OF PERSPECTIVE PERCEPTION IN PICTURES

Aaron Hertzmann[1] (hertzman@dgp.toronto.edu); [1]Adobe Research

We propose hypotheses for 3D shape and space perception in pictures, in which eye fixations and foveal vision play a central role. Perspective describes how painters and cameras produce 2D picture arrangements from observations of 3D space. Many past theories and experimental studies focus solely on linear perspective. Yet, these theories fail to explain many important perceptual phenomena in perspective, including the effectiveness of other approaches to projection; few classical paintings strictly obey linear perspective, nor do the best distortion-avoidance techniques for wide-angle computational photography. We hypothesizes a two-stage process for 3D vision, in which shape is locally interpreted for a current eye fixation in a picture, according to foveal vision, and a global interpretation of a picture arises from spatial relationships between these regions. This framework could provide new insight for understanding pictorial perception across many types of depiction and suggests new studies to explore these hypotheses.

53.462 DISCRIMINATING ABSOLUTE DEPTH OF CONSECUTIVELY PRESENTED TARGETS IN THE REAL ENVIRONMENT WITH BINOCULAR DISPARITY AND RELATIVE HEIGHT INFORMATION

Yiya Chen[1] (chen.12112@osu.edu), Zijiang He[2], Teng Leng Ooi[1]; [1]The Ohio State University, [2]University of Louisville

Coupled with binocular disparity, the background surface provides a reference frame for perceiving the absolute location of a target in midair in the intermediate distance range (2-25 m). However, other depth information, namely, relative height (angular declination/elevation), is also present in the natural environment. We examined how consistent and conflicting binocular disparity and relative height information affected observers' ability to discriminate the absolute distances of two midair targets. Paired, dimly-lit targets (0.2 degree, 805 msec) in a dark room were presented from two target arrays located 4.5 m and 5.75 m from the observer. The first array had six vertically arranged targets at different relative heights (declined or elevated 0.66-0.90 m from an eye-level of 1.2 m) and the second array had one target. The relative binocular disparity between the two arrays was 10 arc min. The observers' task was to report in a two-interval forced choice (2IFC) procedure whether the target in the first or second array was nearer. The main comparisons tested were (i) binocular versus monocular viewing and (ii) texture background versus dark (no-texture) conditions, respectively, to examine the interaction between binocular disparity and visible background surface. The texture background was a 2x4 parallel array of dimly-lit elements either on the floor or ceiling. The average results (n=9) reveal absolute depth discrimination was significantly better with binocular than monocular viewing ($p<0.001$). Performance was superior in the presence of texture background compared to the dark, but only with binocular viewing ($p=0.003$) and not monocular viewing ($p=0.933$). The influence of relative height increased as the observers' binocular depth ability weakened ($r2=0.8913$, $p<0.001$); and with monocular viewing, it was stronger when the texture background was present ($p=0.012$). This indicates both binocular depth and relative height information contribute to locating a midair target. Notably, binocular disparity has a higher weighted contribution.

NIH R01EY033190

53.463 DISTANCE PERCEPTION IN A DEPTH CUE-IMPOVERISHED ENVIRONMENT AFFECTED BY INTENDED NAVIGATIONAL GOAL

Lizhu Yan[1], Lingling Bai[1], Teng Leng Ooi[2], Zijiang He[1]; [1]University of Louisville, [2]The Ohio State University

The visual system integrates the intrinsic bias, an internal model of the ground surface, with external depth information to form a ground surface representation. When walking a short distance from one's home-base in the dark, the intrinsic bias remained anchored to the home-base through path-integration, indicating utilization of an allocentric reference frame (Zhou et al 2023, eLife). Here, we investigated whether the ground reference frame is also allocentric when one walks toward a goal-target in a depth cue-impoverished environment. During the experiment, the observer saw a goal-target (1.5 cd/m2) at 3.5 m amongst an array of 2x8 parallel texture elements (2.25-9.25 m) on the floor. The goal-target was extinguished 5 sec later, and the observer walked to the remembered goal-target location. But before the observer reached the goal-target location, after having walked for either 1.25, 1.75 or 2.75 m, the texture elements were extinguished. This signaled the observer to stop walking. Ten seconds later, a test-target (0.16 cd/m2) was presented (2 sec) for the observer to judge its location using the blind walking-gesturing task. Four test-target locations were measured: three distances (4.50, 5.75, 7.00 m) with 0.14 m elevation and one distance (5.75 m) with 0.5 m elevation. The average results (n=6) show judged test-target distances were longest when the observers' initial walk toward the remembered goal-target location was stopped earlier (1.25 > 1.75 > 2.75 m); that is, when they stopped further from the goal-target ($p<0.001$). This indicates that upon seeing the goal-target, the intrinsic bias shifted toward the goal-target. Thus, as the spatial interval from where the observers stopped walking (to the goal-target) to the shifted intrinsic bias gradually shortened (at 1.75 and 2.75 m), the perceived distance of the test target decreased. Altogether, our data reveal the visual system utilizes an allocentric ground surface representation during navigation.

NIH R01EY033190

53.464 INVESTIGATING WHAT OPTICAL TEXTURE PROPERTY IS USED FOR RELATIVE DISTANCE PERCEPTION

Pin Yang[1] (youngpink678@gmail.com), Geoffrey P Bingham[1]; [1]Indiana University Blomington

Bingham et al. (2022) investigated which property of the texture on a support surface is used to perceive relative distance and found it was the texture image shape. We further investigated the use of image shape by manipulating texture image size and the aspect ratio of optical texture elements. Methods: Participants matched the distance (TD) of a target by moving a hand avatar with both appearing on different side-by-side textured support surfaces in a virtual display. While texture elements on the left remained 3cm squares, elements on the right varied in three conditions: (1) Identical Texture, right elements same as left; (2) Different Size (DS), right element sized 0.6/0.8/1/1.2/1.4 relative to left; (3) Aspect Ratio (AR), image width or height altered so image width/height = 0.6/0.8/1/1.2/1.4 on the right. Results: multiple regressions showed that in the width AR manipulation, TD was significant ($p < 0.001$, $F = 243.98$), but neither AR ($p > 0.07$) nor the interaction ($p > 0.3$) was. In the height AR manipulation, TD was significant ($p < 0.001$, $F = 869.17$) as was AR ($p = 0.009$, $F = 3.64$), but the interaction ($p > 0.7$) was not. In the DS manipulation, TD was significant ($p < 0.001$, $F = 904.84$) as was DS ($p < 0.001$, $F = 5.64$), but the interaction ($p > 0.2$) was not. Conclusions: Variations in the size of the texture elements influenced relative distance perception. However, when manipulating aspect ratio, changes in the width were not significant, while the image height was, which is consistent with Chen & Saunders (2020) showing that compression of optical texture was used in slant perception. Reconsidering Bingham et al. (2022), we concluded that the relative compression of optical texture (image height) is used as monocular information about relative distance.

53.465 INTERPLAY OF EXOCENTRIC AND EGOCENTRIC INFORMATION IN DISTANCE PERCEPTION FOR VISUOMOTOR TASKS

Chaeeun Lim[1] (chaeeun_lim@brown.edu), Dhanraj Vishwanath[2], Fulvio Domini[1]; [1]Brown University, [2]University of St Andrews

Imagine two different visuomotor tasks directed at a cup in front of you: either reaching to touch the front or back of the cup separately or grasping it from front-to-back. Intuitively, the first task could be solved with egocentric information (an estimate of distance), while the second task would additionally require exocentric information (an estimate of depth). Indeed, a previous study suggested that reaching to points on two separated rods involved only egocentric information such as vergence angle and accommodation, whereas grasping both rods involved exocentric information specified by cues such as relative disparity. Here, we asked whether this dissociation persists when the observer is asked to either reach to or grasp the front and back of the same object (a binocularly viewed virtual paraboloid). To manipulate exocentric information independently from egocentric information, we added a texture gradient to disparity information that, according to previous studies examining cue combination, should result in a greater perceived exocentric depth. Contrary to the idea that reaching relies solely on egocentric cues, our findings reveal that reaching distance to points on the same object is influenced by both egocentric and exocentric information, mirroring the pattern observed in grasping tasks. Strikingly, participants consistently overestimated the depth interval between the front and back of the object in both reaching and grasping. This overestimation intensified with an additional texture cue, particularly at greater viewing distances, aligning with predictions of a summative model of cue integration (Intrinsic Constraint model). Paradoxically, when subjects reached at the same point in space, its position was underestimated when it represented the object's front and overestimated when it represented the back. The present findings suggest that the visual system does not exclusively rely on egocentric information to estimate reaching distance for points on the same extended object, but it also incorporates exocentric information.

This material is based upon work supported by the National Science Foundation under Grant No.2120610.

Tuesday, May 21, 8:30 am – 12:30 pm, Pavilion

3D Perception: Virtual and augmented reality

53.466 VISUAL FACTORS INFLUENCING TRUST AND RELIANCE WITH AUGMENTED REALITY SYSTEMS

Gerd Bruder[1] (bruder@ucf.edu), Michael P Browne[2], Zubin Choudhary[1], Austin Erickson[1], Hiroshi Furuya[1], Matthew Gottsacker[1], Ryan Schubert[1], Greg Welch[1]; [1]University of Central Florida, [2]Vision Products LLC

Augmented Reality (AR) systems are increasingly used for simulations, training, and operations across a wide range of application fields. Unfortunately, the imagery that current AR systems create often does not match our visual perception of the real world, which can make users feel like the AR system is not believable. This lack of belief can lead to negative training or experiences, where users lose trust in the AR system and adjust their reliance on AR. The latter is characterized by users adopting different cognitive perception-action pathways by which they integrate AR visual information for spatial tasks. In this work, we present a series of six within-subjects experiments (each N=20) in which we investigated trust in AR with respect to two display factors (field of view and visual contrast), two tracking factors (accuracy and precision), and two network factors (latency and dropouts). Participants performed a 360-degree visual search-and-selection task in a hybrid setup involving an AR head-mounted display and a CAVE-like simulated real environment. Participants completed the experiments with four perception-action pathways that represent different levels of the users' reliance on an AR system: AR-Only (only relying on AR), AR-First (prioritizing AR over real world), Real-First (prioritizing real world over AR), and Real-Only (only relying on real world). Our results show that participants' perception-action pathways and objective task performance were significantly affected by all six tested AR factors. In contrast, we found that their subjective responses for trust and reliance were often more affected by slight AR system differences than would elicit objective performance differences, and participants tended to overestimate or underestimate the trustworthiness of the AR system. Participants showed significantly higher task performance gains if their sense of trust was well-calibrated to the trustworthiness of the AR system, highlighting the importance of effectively managing users' trust in future AR systems.

This material includes work supported in part by Vision Products LLC via US Air Force Research Laboratory (AFRL) Award Number FA864922P1038, and the Office of Naval Research under Award Numbers N00014-21-1-2578 and N00014-21-1-2882 (Dr. Peter Squire, Code 34).

53.467 INVESTIGATING FINGER-TAPPING AND PUPILLOMETRY AS POTENTIAL INDICATORS OF PRESENCE IN VR

Sean Hinkle[1] (sean.hinkle@ucf.edu), Shayna Soares[1], Robert Bacon[1], Corey Bohil[2]; [1]University of Central Florida, [2]Lawrence Technological University

Spatial presence in VR, the feeling of 'being there,' is linked to outcomes in clinical, training, education, and entertainment applications. Overreliance on survey measures has hampered the field and prevented progress with few generalizable alternatives. The field has tested physiological, neuroimaging, and behavioral measures in search of continuous and objective indicators of presence. In two studies we evaluated finger-tapping and pupillometry as potential indicators of presence. We predicted that variance in inter-tap-intervals (ITIs) and pupil size would predict presence, and that a neural-net classifier would be able to identify high versus low presence conditions at the individual subject level. In Experiment 1, participants walked the "virtual plank" tapping to a rhythm at heights or on the ground to manipulate presence. Surveys confirmed that heights manipulated presence ($p = .04$). ITI variance did not follow this pattern ($p = .375$). A feedforward neural-net classifier was trained on tapping and pupillometry data at the individual level. For finger-tapping, the classifier identified the condition of four-second windows of finger-position data at 77% accuracy. Pupillometry data yielded 70% accuracy, but a lighting confound weakened our conclusions. In Experiment 2, participants watched two 360-degree videos twice, with or without sound, to manipulate presence while controlling global luminance. Each video was analyzed separately. Surveys confirmed that sound increased presence for both videos (all $ps < .05$), but pupil variance did not follow this pattern. The neural-net classifier was unable to replicate the high accuracy of Experiment 1, with accuracies of 57% and 55%. Our results demonstrate that finger-tapping is a promising indicator of presence in VR and is especially sensitive when analyzed via neural-net classifier. While results for pupillometry are mixed, we believe that pupillometry and other eye-tracking metrics merit further investigation with more refined machine-learning methods, potentially in combination with finger-tapping.

53.468 GAIN ADAPTATION IN VIRTUAL REALITY

Xue Teng[1,2] (xueteng@yorku.ca), Laurie Wilcox[2], Robert Allison[1,2]; [1]Department of Electrical Engineering & Computer Science, York University, [2]Centre for Vision Research, York University

Humans rely on both visual and kinesthetic cues to register self-motion and navigate through the world. Normally, these sources of information are consistent with the perceived motion through the environment. Virtual Reality (VR) often introduces visual/kinesthetic inconsistency due to scale differences between the physical space and the simulated virtual space (motion gain). Large amounts of left-right self-motion gain has been found to compress apparent distance and monocular depth. In the present study, we asked whether observers adapt to exposure to extended periods of high or low motion gain. In the adaptation phase, observers played a VR game. They moved laterally to intersect targets with their body; their virtual motion was scaled to be either 0.67, 1 or 2 times their physical motion. These three adaptation conditions were presented in separate sessions, each starting with 5 minutes of initial adaptation, followed by testing interleaved with three 2-minute top-up adaptation periods. During the test phase, observers swayed left-right over 20 cm to the beat of a 0.5 Hz metronome and indicated if the virtual environment moved 'more or less than they did'. Using a method of constant stimuli, we measured the PSE for each gain in separate blocks. Results from 18 observers showed that there were no consistent differences between the PSEs obtained in the three gain conditions. Neither increasing the adaptation duration nor testing monocularly affected this pattern of results. These data suggest that while observers are sensitive to differences between their intended movement and that rendered in VR, they do not appear to adapt to a constant mismatch in the current setup. Ongoing experiments are evaluating if this is also true for forward-backward motion. Lack of adaptation to mismatches between self and world motion may be an important evolutionary strategy as such distortions could signal a hazardous situation (e.g. poisoning).

Thanks to NSERC Canada for support under an Alliance grant, ALLRP 570802 - 21 in partnership with Qualcomm Canada.

53.469 SIZE CONSTANCY IN VIRTUAL REALITY

Yu Chan Kim[1] (kimuchan1@pusan.ac.kr), Sung Jun Joo[1]; [1]Pusan National University

Size constancy enables us to have a stable size representation of an object despite changing retinal angular size at different viewing distances. However, it is unknown whether size constancy is maintained in virtual reality (VR), in which ample depth cues are available as in the real world. In the present study, we aimed to test size constancy in VR by assessing whether the size estimation of an object is affected by viewing distance. In our experiment, naïve observers (n=22) viewed a standard sphere (diameter of 40 cm) at 2 m and a test sphere at various viewing distances (2-12 m) simultaneously. They adjusted the size of the test sphere presented with a random diameter so that it looked equal to the physical size of the standard sphere. The results showed that at a near distance (~8 m), size constancy was preserved: the estimated size of the test sphere was not different from the size of the standard sphere. However, size constancy was progressively impaired as the distance increased above 8 m and our observers' PSEs were systematically reduced (10 m: 36.28 cm, $p = 0.012$, 12 m: 33.23 cm, $p < 0.001$), which means that they overestimated the size of the test sphere. Our finding suggests that size constancy in VR might be affected by the well-known underestimation of distance in VR.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (NRF-2022R1I1A2067731 and NRF- 2022R1A2C3004133).

53.470 RECONSTRUCTING RETINAL IMAGES FROM NATURAL OUTDOOR WALKS WITH 3D GAUSSIAN SPLATTING

Daniel Panfili[1] (dan.panfili@utexas.edu), Nathaniel Powell[2], Mary Hayhoe[3]; [1]University of Texas at Austin

Understanding the visual signal incident on the retina is crucial for understanding what visual information is used to guide specific actions. Retinal image calculation requires a gaze vector, eye position in a world-centered coordinate frame, and a representation of the environment. Previously, we calculated a 3D representation of the environment (together with luminance and chromaticity) using photogrammetry methods (Muller et al, 2023). This allowed more accurate calculation of retinal motion, and investigation of the visual information used for path planning. However, this reconstruction method faced challenges in accurately depicting scenes with reflective, transparent, or homogeneous surfaces, and struggled with high-spatial frequency elements (e.g., fine-grain textures). Consequently, the resulting retinal images lacked several natural scene features, notably the sky, bodies of water, and fine tree limbs.

This limitation renders the data unsuitable for contexts where image realism is important, like perceptual straightening paradigms (Henaff el al, 2019). To enhance realism, we adopted a new rasterization technique called 3D gaussian splatting (Kerbl et al, 2023). 3D gaussian splatting converts point clouds into a smooth, continuous representation by applying Gaussian-based kernel blending, emphasizing visual information from nearby points in the cloud while gradually decreasing influence with distance. This technique provided a more accurate portrayal of the environment, closely resembling the perspectives captured by the scene camera used for reconstruction. As a result, it generated higher fidelity retinal images compared to photogrammetry meshes. However, this technique, while promising for stimulus presentation, is still in its nascent stage and lacks some critical features like collision calculation. Overall, 3D gaussian splatting emerges as a promising tool for stimulus presentation, though additional refinement is necessary to more accurately reflect the visual image properties.

NIH grant EY05729

53.471 TEST ON THE EFFECTIVENESS OF VIRTUAL NOSE IN REDUCING CYBERSICKNESS

Sai Ho Yip[1], Adrian K. T. Ng[2], Henry Y. K. Lau[3], Jeffrey A. Saunders[1]; [1]University of Hong Kong, [2]University of Bristol, [3]United Kingdom Atomic Energy Authority

It has been reported that rendering a virtual nose in a head-mounted display (HMD) can reduce cybersickness. A virtual nose could be a promising low-cost intervention against cybersickness. However, the only peer-reviewed and published evidence for a virtual nose effect is from a single low-powered experiment. In this study, we attempted to replicate the virtual nose effect. Subjects performed a virtual navigation task with an HMD and reported their motion sickness using the Fast Motion Sickness (FMS) scale. Two conditions, Nose-Present vs. Nose-Absent, were tested on two separate days. For the Nose-Present condition, the virtual nose had to be placed further from the natural position to ensure its visibility, as in the previous study. Potential habituation effects were controlled by counterbalancing the order of conditions. Contrary to the previous study, we did not detect a significant difference between FMS scores with and without a virtual nose, and our results suggest that any benefit was not as large as previously reported. Our results do not support the hypothesis that a virtual nose can effectively reduce cybersickness.

This research was supported by a grant from the Hong Kong Research Grants Council (GRF 17407914)

TUESDAY AFTERNOON POSTERS IN BANYAN BREEZEWAY

Tuesday, May 21, 2:45 – 6:45 pm, Banyan Breezeway

Face and Body Perception: Emotion

56.301 AGE-RELATED DIFFERENCES IN DETECTING FACIAL EMOTION CHANGES: UNVEILING THE IMPACT OF DYNAMIC STIMULI ON SENSITIVITY

Su-Ling Yeh[1,2,3,4,6] (suling@ntu.edu.tw), Po-Cheng Huang[1], Yu-Shan Chen[1], Yi-Chuan Chen[5]; [1]Department of Psychology, National Taiwan University, Taipei, Taiwan, [2]Graduate Institute of Brain and Mind Sciences, National Taiwan University, Taipei, Taiwan, [3]Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan, [4]Neurobiology and Cognitive Science Center, National Taiwan University, Taipei, Taiwan, [5]Department of Medicine, Mackay Medical College, New Taipei City, Taiwan, [6]National Humanities Center, Research Triangle Park, North Carolina, USA

Facial expressions play a pivotal role in understanding emotions. Yet, older adults often face recognition challenges, possibly due to static images used in studies that do not align with their daily experiences. This study explores whether dynamic and emotionally meaningful stimuli enhance older adults' sensitivity to detect changes in facial emotions compared to their younger counterparts. A psychophysical method was developed to measure participants' sensitivity to facial emotion changes, with judgments focused on whether the emotion displayed at the start of a video clip differed from that at the end. Emotional valence (happy, sad) and emotion-change direction (onset, offset) were manipulated, resulting in four conditions: happy-onset, happy-offset, sad-onset, and sad-offset. Detection sensitivity, quantified by the reciprocal of the threshold for changes in facial emotion, revealed that older adults had higher sensitivity to both the onset and offset of happy expressions compared to younger adults and their own sensitivity to the onset and offset of sad expressions. This emphasizes the impact of positive emotion on older adults, regardless of the direction of emotion change, aligning with the socioemotional selectivity theory. Conversely, younger adults exhibited higher sensitivity to the onset of sad expressions than the offset and higher sensitivity to the offset of happy expressions than sad ones, reflecting the combined impact of emotional valence and emotion-change direction. This study implies that older adults may outperform younger counterparts in detecting changes in others' facial emotions when exposed to dynamic and emotionally meaningful positive stimuli. In conclusion, these findings prompt a reconsideration of the prevailing assumption that aging universally implies cognitive deterioration, highlighting the nuanced impact of age-related differences in visual perception.

This work is supported by the National Science and Technology Council in Taiwan (MOST 110-2410-H-002-130-MY3 and MOST 111-2223-E-002-008)

56.302 BEAT: BERKELEY EMOTION AND AFFECT TRACKING DATASET

Ana Hernandez[1] (ana.hernandez@berkeley.edu), Zhihang Ren[1], Jefferson Ortega[1], Yifan Wang[1], Zhimin Chen[1], Yunhui Guo[2], Stella X. Yu[3], David Whitney[1]; [1]University of California, Berkeley, [2]University of Texas at Dallas, [3]University of Michigan, Ann Arbor

The ability to perceive the emotions of others is incredibly important when navigating and understanding the social world around us. To understand this visual perceptual mechanism, previous studies have focused on face processing leading many previous datasets to collect face-centric data. Recent research has found that the visual system also utilizes background scene context to modulate and assign perceived emotion (e.g., Chen & Whitney, 2019, 2020, 2021). Similarly, "in-the-wild" datasets have been created to include contextual information, such as CAER and EMOTIC. However, these datasets only track categorical emotions, and ignore dimensional ratings (i.e. valence and arousal) or collect ratings on static images. In this project, we propose BEAT: The Berkeley Emotion and Affect Tracking Dataset, the first video-based dataset that contains both categorical and continuous emotion annotations for a large number of videos (124 videos total). BEAT provides more insights into human emotion perception by providing both categorical and dimensional ratings for individual videos. Additionally, BEAT can help researchers better understand how emotion is processed temporally, as it is in the real world. Additionally, compared to other datasets, a large number of annotators (n = 245) were recruited to avoid idiosyncratic biases. BEAT can also be beneficial for artificial intelligence (AI) models. Using unbiased and multi-modality annotations, AI models trained on BEAT can be more robust and fair. Finally, we also release a new AI benchmark for emotion recognition multi-tasking. The BEAT dataset will help increase our understanding about how humans perceive the emotions of others in natural scenes.

56.303 DEVELOPING A NON-HUMAN PRIMATE MODEL TO DISSECT THE NEURAL MECHANISMS OF FACIAL EMOTION PROCESSING RELEVANT IN AUTISM SPECTRUM DISORDER

Shirin Taghian Alamooti[1] (staghian@yorku.ca), NaYeon Kim[2], Ralph Adolphs[2], Kohitij Kar[1]; [1]Department of Biology, York University, [2]California Institute of Technology

Understanding facial expressions is crucial for human social interaction as they convey emotions and intentions. Autistic adults often show marked differences (compared to neurotypically developed adults) in interpreting these cues, impacting communication and empathy. To aid autistic individuals, it's essential to comprehend the neural basis behind these differences. However, the diverse nature of behavioral reports in autism impedes efficient study design. Current models for interpreting facial emotion judgments overlook individual image-level sensory representations, pivotal in understanding differences between neurotypical and autistic adults. In a recent study by Kar (2022), behavioral disparities between these groups were more evident at the image level rather than through broad categorical descriptors like "happiness" or "fear." To investigate this further, we established an image-level framework using 360 diverse facial expression images from the Montreal Set of Facial Displays of Emotion (MSFDE). Through binary emotion discrimination tasks, we observed subtle yet significant differences in image-level behavioral error patterns between neurotypical and autistic individuals. Addressing the challenge of heterogeneity, we pinpointed shared variances in our developed image-level metrics, serving as a critical behavioral benchmark. To delve into the neural underpinnings, we conducted extensive neural recordings in the inferior temporal (IT) cortex of rhesus macaque monkeys. Initial findings align with previous predictions (Kar 2022), indicating stronger correlations between macaque IT-based decodes of facial emotion responses and neurotypical behavior compared to autistic behaviour. Our study aims to create an innovative framework merging non-human primate neural investigations with the autistic behavioral phenotype. By focusing on shared variances in image-level behavioral metrics, we aim to identify more sensitive neurobehavioral markers

CFREF, Brain Canada, SFARI

56.304 ETHNICITY AND PAIN RECOGNITION: UNRAVELING CONFUSION PATTERNS IN FACIAL EXPRESSIONS

Marie-Pier Plouffe-Demers[1,2,3], Chaona Chen[3], Angélica Pérez Motta[2], Valentina Gosetti[3], Caroline Blais[2], Rachael E. Jack[3]; [1]University of Quebec in Montreal, [2]University of Quebec in Outaouais, [3]University of Glasgow

Ethnic minorities expressing pain are often under-diagnosed and under-treated (Cintron et al., 2006). The misinterpretation of pain signals conveyed through facial expressions across different ethnicities might be a key factor contributing to these variations (Dildine et al., 2023). Current theories of social perception suggest that such misclassifications could be due to ethnic stereotype knowledge that could amplify confusions between pain and similar-looking facial expressions, like anger or disgust (Hugenberg et al., 2004; Kunz et al., 2019; Roy et al., 2015). However, no study has explicitly examined the interaction between these factors. Here, we addressed this question by first examining whether face ethnicity influences the interpretation of facial expressions of pain as other negative emotions. Using a generative model of the human face, we displayed 40 facial expressions of pain, sadness, anger, fear and disgust on each of 40 face identities varying in ethnicity (Black, East Asian, White) and sex (male, female). Participants (30 White Western; sex-balanced; 18-35 years) classified each stimulus in an alternative forced choice task (5AFC). Combining a within-subject bootstrap analysis (10,000 resamples) with the Bayesian estimate of population prevalence (Ince et al., 2021), we investigated the effect of ethnicity on accuracy and confusion patterns between pain and other negative emotions. Preliminary results (n = 22) show that, across face ethnicities, facial expressions of pain are the least accurately identified and generate more systematic confusions with disgust for black compared to white faces. To explore these confusions further, we plan to conduct a second complementary discrimination task where participants will detect the presence of target emotions (measured using d-prime, n = 30). By investigating the role of face ethnicity in interpreting facial expressions of pain, our study aims to shed light on how and why disparities in pain perception arise and provide potential insights into mitigating these effects.

This work was supported by a European Research Council Starting Grant (Rachael E. Jack), a Canada Research Chair in Cognitive and Social (Caroline Blais), and by graduate scholarships from the Social

Sciences and Humanities Research Council and Mitacs Globalink (Marie-Pier Plouffe-Demers)

56.305 GENDER DIFFERENCES IN THE ENCODING AND DECODING OF PAIN FACIAL EXPRESSIONS.

Arianne Richer[1] (rica30@uqo.ca), Camille Saumure[2], Daniel Fiset[1], Zoé Glardon[2], Caroline Blais[1]; [1]University of Quebec in Outaouais, [2]University of Fribourg

Facial expressions are crucial for assessing others' affective states. However, pain facial expressions (PFE) are poorly recognized, often confused with other negative affective states (Kappesser & Williams, 2002) and less easily perceived in women's faces (Riva et al., 2011). Studies have revealed various configurations of PFE (Kunz & Lautenbacher, 2014). Yet, it is unclear whether some of these configurations are easier to recognize than others–such differences may in part explain the disparities in perceived pain as a function of face gender. This study explored potential gender differences in the configurations of PFE (encoding) as well as their perception by external observers (decoding). We used the Delaware Pain Database (DPD; Mende-Siedlecki et al., 2020), containing 225 pictures of White individuals posing PFE. To investigate potential differences in PFE encoding between men and women, we used OpenFace to measure the activation levels of 17 action units (AUs) in those pictures. A principal component analysis indicated five main groups of AUs with correlated activations. Most importantly, the first component included AUs typically associated with PFE, and was more prominent in PFE of men than women. To verify if PFE are decoded differently as a function of face gender, we used ratings openly available within the DPD. Each picture in the DPD has indeed been rated on the perceived intensity of six basic emotions and pain. A mixed ANOVA 7 (affective states) by 2 (genders) indicated significant main effects of affective states and gender, as well as an interaction between both factors. T-tests indicated that fear and sadness were perceived as significantly higher in women's PFE, while pain was perceived as significantly higher in men's. These findings emphasize gender-specific disparities in PFE, their potential overlap with other affective states, and underscore the potential contribution of both encoding and decoding in observed gender differences.

The present study is supported by the Canada Research Chair in cognitive and social vision to Caroline Blais (# 950-232282)

56.306 IMAGES OF FACIAL EXPRESSIONS WITH HARDER TO RECONSTRUCT REPRESENTATIONS ARE EVALUATED AND REMEMBERED AS MORE INTENSELY EMOTIONAL

Nicholas Fagan[1] (nicholas.fagan@yale.edu), Qi Lin[2], James Gross[3], Amit Goldenberg[4], Steve Chang[1], Ilker Yildirim[1]; [1]Yale University, [2]RIKEN Institute, [3]Stanford University, [4]Harvard University

If faces hold a special place in our minds, emotional faces stand out as extraordinary. Not only do faces expressing emotions engage rich inferences that deeply impact how we navigate the social world, they also lead to greater attention, better memory, and faster reactions. Yet, the computational basis of how we evaluate and remember emotionality of faces remains unaddressed. Here, we hypothesize that instead of domain-specific mechanisms optimized for emotional expressions, processing the emotionality of faces arises from a general mechanism of a perception-to-memory interface. To test this hypothesis, we use a recent computational formulation of the classic level-of-processing theory — the idea that memory strength is determined by the depth of perceptual analysis. This formulation used a sparse coding model (SPC) to compress feature embeddings of natural scene images, and showed that images with harder to reconstruct representations are more memorable. Here we train the SPC model to compress feature embeddings of images of faces, with the feature embeddings obtained from a pretrained face recognition deepnet. We hypothesized that the remembered emotionality of faces would be positively correlated with the depth of processing, operationalized by the magnitude of reconstruction residuals of individual face images under this model. We tested this prediction by exploiting a recent phenomenon of emotionality judgments: Humans overestimate the average emotionality of sequentially presented faces. We find that this "sequential amplification" effect falls from our model: Simply averaging the reconstruction error of the representations of faces in a sequence reproduces the amplification effect, including finer-grained effects of valence (greater amplification for negative vs. positive). Moreover, reconstruction error correlates highly with the emotionality of face sequences and singletons. Crucially, these results don't occur using vision-only models (i.e., without reconstruction error) or non-face deepnets. These results ground evaluation and remembering of emotional expressions in a general perception-to-memory interface.

56.307 INTERACTION BETWEEN THE PREFRONTAL AND VISUAL CORTICES SUPPORTS SUBJECTIVE FEAR

Vincent Taschereau-Dumouchel[1,2] (vincent.taschereau-dumouchel@umontreal.ca), Marjorie Côté[1,2], Shawn Manuel[1,2], Valevicius Darius[1,2], Hakwan Lau[3]; [1]Department of Psychiatry and Addictology, Université de Montréal, [2]Centre de Recherche de l'Institut Universitaire en Santé Mentale de Montréal, [3]RIKEN Center for Brain Science

It has been reported that threatening and non-threatening visual stimuli can be distinguished based on multi-voxel patterns of hemodynamic activity in the human ventral visual stream. Do these findings mean that there may be evolutionarily hardwired mechanisms within early perception, for the fast and automatic detection of threat, and maybe even for the generation of the subjective experience of fear? In this study, we provide evidence using human neuroimaging that the ventral visual stream may represent affectively neutral visual features that are statistically associated with fear ratings of participants, without representing the subjective experience of fear itself. More specifically, we show that patterns of hemodynamic activity predictive of a specific "fear profile" (i.e., fear ratings reported by a given participant) can be observed in the ventral visual stream whether a participant reports being afraid of the stimuli or not. Further, we found that the multivariate information transmission between ventral visual areas and prefrontal regions distinguished participants who reported being subjectively afraid of the stimuli from those who did not. Together, these findings support the view that the subjective experience of fear may depend on the relevant visual information triggering implicit metacognitive mechanisms in the prefrontal cortex.

56.308 PERCEIVED AMBIGUITY OF FACIAL EXPRESSION -CONTRIBUTION OF THE COMBINATION OF FACIAL PARTS OF DIFFERENT EMOTIONS-

Takashige Hamano[1] (thlab5603@gmail.com), Atsunori Ariga[1]; [1]Chuo Univsercity

Although facial expressions serve as important emotional cues in our daily interaction with others, we sometimes have a difficulty in reading others' emotions when their facial expressions are ambiguous. Previous research has investigated the ambiguous facial expression based on the assumption that it conveys some sort of emotions, requiring observers to determine which single emotion (e.g., happiness or sadness) it expresses. However, we could indeed perceive more than one emotion from the other's facial expression or leave it "ambiguous" without an identification of emotions. There is currently no empirical research on how we perceive the ambiguity of the facial expression, which was the focus of this study. We first made line-drawing faces comprised of eyebrows, eyes, and mouth, which expressed happiness, anger, and surprise respectively. Next, we shuffled facial parts and made 27 facial stimuli in total. Participants rated the ambiguity of the facial expression, which was defined as faces expressing more than one emotion or being difficult to be emotionally identified, as well as the strength of perceived expressions of six basic emotions for each stimulus. First, the perceived ambiguity increased as the number of emotions involved in a face. Second, even though the number of emotions contained in a face was same, the perceived ambiguity increased when it involved the angry expression. Third, the perceived ambiguity decreased when perception of the happiness and surprise were salient. These results suggest that the perceived ambiguity depends on the quantity and quality of emotions involved. Given that the anger contributes to the perceived ambiguity and that it depends on the saliency of the happiness (or positive expression) and surprise (or neutral expression), we are likely to perceive the ambiguous expression as somewhat negative, rather than neutral offset of multiple emotions.

This research was financially supported by JSPS KAKENHI Grant Numbers JP22K03210 to AA.

56.309 YOUR SMILE, MY SUCCESS: FACES ASSOCIATED WITH POSITIVE EMOTION FACILITATE COGNITIVE FLEXIBILITY AND STABILITY

Rebeka Almasi[1] (almasi@gwu.edu), Jini Tae[2], Yoonhyoung Lee[3], Myeong-Ho Sohn[1]; [1]The George Washington University, [2]Gwangju Institute of Science and Technology, [3]Yeungnam University

The emotional valence of surroundings affects behavior of the beholder. While previous research investigated the effect of task-relevant affect cues, many of the applications of emotion rely on more subliminal, task-irrelevant cues—for example, making rooms calming or bright colors to influence mood without having people explicitly say how happy the walls make them. Advertisers, too, are aware that irrelevant emotional cues influence decision-making, as in the case of associating their products with positively-valenced stimuli. Specifically, in the recent literature, emotional valence of task-relevant stimuli affects cognitive control. In an attempt at greater ecological validity, our research determines whether previous research showing facilitation of cognitive stability by negative emotion and cognitive flexibility by positive emotion holds for task-irrelevant emotion. In Experiment 1, participants responded to the apparent gender of a face ignoring a congruent or incongruent gender word. Always frowning actors, as opposed to always smiling actors, prompted a larger congruency effect, a hindrance to distractor inhibition, implying that positively-associated faces facilitated cognitive stability. In Experiments 2 and 3, participants switched between identifying age and gender. The smaller task switch cost, a marker of increased cognitive flexibility, for smiling faces shows further facilitation of cognitive control. Our results imply that faces with positive emotion can promote better cognitive stability and flexibility than faces with negative emotion, as demonstrated in a gender Stroop task (that requires maintaining a high level of cognitive control to reduce distractor interference) and a task-switching paradigm (that necessitates cognitive flexibility to switch between tasks). Our very ability to use controlled and flexible attention may depend on our emotional associations with stimuli in the environment.

56.310 VISUOCORTICAL, AUTONOMIC, AND BEHAVIORAL DYNAMICS DURING THE GENERALIZATION OF ACQUIRED SOCIAL FEAR

Jourdan Pouliot[1] (jourdan.pouliot@ufl.edu), Richard Ward[1], Caitlin Traiser[1], Payton Chiasson[1], Faith Gilbert[1], Andreas Keil[1]; [1]University of Florida

When acquiring fear through aversive conditioning, biological organisms tend to generalize the acquired fear response to stimuli that resemble the conditioned threat cue (i.e. the CS+). Applied to interpersonal interactions, the generalization of a CS+ identity to similar faces is the basis of identifying potentially dangerous conspecifics in such an environment. The visuocortical correlates of fear generalization have been extensively studied using simple visual stimuli such as shapes and Gabor patches, but little work has examined complex stimuli such as faces. The present study investigated how neural and psychophysiological responses generalize across facial identities during Pavlovian aversive conditioning. Subjects were presented seven flickering faces varying along a similarity gradient, while measuring electroencephalography (EEG) and pupil diameter. One of these faces (CS+) was repeatedly paired with a loud noise (US). Various measures of affect-biased attention, including the steady-state visual evoked potential (ssVEP), endogenous alpha power, pupil diameter, and self-reported valence and arousal were collected. Trials were separated into early and late periods to analyze the effect of contingency learning on conditioned responses. Data were fit to a priori weights via nonparametric frequentist and Bayesian methods to examine the shape of the observed tuning functions. We found that across all faces, occipitoparietal ssVEP power fit a sharpening pattern resembling a Difference-of-Gaussians. Contrary to expectations, fit to this sharpening pattern was greatest during early trials and declined later in acquisition. In contrast, alpha responses during late trials showed a broad generalization pattern in occipitotemporal regions associated with processing of facial features. Lastly, pupil and evaluative responses exhibited shallow generalization during early trials, and this tuning function narrowed as learning progressed. These results indicate that (1) conditioned generalization of faces prompts similar changes as observed with simple visual cues; (2) that visuocortical, autonomic, and behavioral indices change at different temporal rates.

This research was supported by NIH grant R01MH125615.

Face and Body Perception: Wholes, parts configurations, features

56.311 EFFECTS OF PROTOTYPICALITY ON PERCEIVED FACIAL UGLINESS

Cheng Hsuan Chen[1] (r09122005@ntu.edu.tw), Chien Chung Chen; [1]National Taiwan University

It is suggested (Eco, 2007) that the lack of beauty is not ugly, which requires the involvement of negative emotions, but unremarkable. We tested whether beauty and ugliness exist on the same perceptual dimension by observing how facial prototypicality affects facial ugliness and beauty. The prototype faces were created by averaging frontal view faces of 36 individuals. We then interpolated between the prototype faces and the original faces to create faces with various degrees of prototypicality and extrapolated out of the original faces to create caricatures and out of the prototype faces to create anti-faces. We then presented participants with faces of varying degrees of prototypicality and asked them to rate each face on its perceived ugliness, beauty, and symmetry with a 6-point Likert scale. The prototype faces were considered the most beautiful and the least ugly, while the caricatures were deemed the ugliest and least attractive. The non-existent anti-faces received the equivalent rating as its corresponding normal face, suggesting that the evaluation of attractiveness is based on the deviations from prototype faces. In general, prototypicality was positively correlated with perceived beauty but negatively correlated with perceived ugliness. Prototypicality contributes to the perception of ugliness. Descriptions of a face as 'unattractive' were comparable to descriptions of it as 'ugly.' The results suggest that the spectrum of aesthetic judgment encompasses both beauty and ugliness, challenging existing theories on the aesthetic difference between ugliness and averageness previously proposed (Eco, 2007).

NSTC (Taiwan) 112-2423-H-002-002

56.312 IDENTIFYING OTHER-RACE FACES: IT'S LESS IN THE EYES.

Anthony Proulx[1] (anthonyproulx21@gmail.com), Isabelle Charbonneau[1], Justin Duncan[1], Vicki Ledrou-Paquet[1], Caroline Blais[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais

Most studies in face recognition have focused on how individuals identify faces within their own ethnic group, highlighting the crucial role of the eye region in face identification (Vinette et al., 2004; Butler et al., 2010; Royer et al., 2018). Nevertheless, the general population encounters difficulties in identifying individuals from a different ethnicity, a phenomenon known as the "other-race effect" (ORE; Meissner & Brigham, 2001). Despite decades of investigation, the perceptual mechanisms associated with ORE remain inadequately understood. It is plausible that different perceptual strategies are employed in identifying other-race faces. In this study, 21 White participants initially learned to identify 8 Black and 8 White faces. Subsequently, they were tasked to recognize these same faces presented through small Gaussian apertures ("Bubbles"; Gosselin & Schyns, 2002). We also measured the extent of the ORE using an old/new recognition task. Collectively, participants exhibited an ORE, evidenced by a higher d' with own-race faces ($\mu d=0.26$, $\sigma d=0.27$, $t(20)=4.32$, $p<.001$, $d=0.94$). Regarding the Bubbles results, Pixel Tests ($p<.05$; Stat4Ci Toolbox; Chauvin et al., 2005) revealed a significant reliance on the eyes and the mouth for faces of both race. Crucially, comparison between classification images of own- and other-race faces unveiled significant differences: greater eye reliance (z-score difference of 12.52) for own-race faces and increased nose and mouth reliance (z-score differences of -5.58 and -5.94, respectively) for other-race faces. We hypothesize that, at least as a group effect, the ORE may arise from diminished eye reliance and an excessive dependence on facial features associated with ethnic information (Levin, 1996). Additional participants, including a more diverse sample (e.g., African participants), are currently undergoing testing to explore individual and cultural differences in perceptual strategies employed to recognize own- and other-race faces.

56.313 RECURRING FACE INTEGRATION IN THE "DOUBLE FACE" ILLUSION

Yingshi Ding[1,2] (yc27156@connect.um.edu.mo), Runnan Cao[2], Zhen Yuan[3], Sheng He[2], Peng Zhang[2]; [1]Faculty of Education, University of Macau, Macau, China, [2]State Key Lab of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China., [3]Faculty of Health Sciences, University of Macau, Macau.

Face recognition is a holistic process that relies on the integration of facial features, normally a rapid process completed in a fraction of a second. A "double face" with two eyes and two mouths allows multiple possibilities of face integration. Observers often feel uncomfortable and disturbed when viewing double faces, possibly due to the continuation of face integration process. Therefore, the double face illusion may serve as a useful paradigm to study the face integration process. In an eye-tracking experiment (N = 21), subjects performed a two-back task based on the identity of face images, with normal faces and double faces presented in separate runs. Results show that subjects had fewer fixations in the eye region and more fixations in the mouth region when they viewed the double face images. In addition, there were more saccades in-between the eyes and mouth regions when subjects viewed the double face images. These results are consistent with continued attempts to integrate facial features into holistic face representations. In a fMRI experiment (N = 7), BOLD activity were measured during prolonged viewing of normal face and double face images. The double faces produced significantly stronger BOLD activations in the right FFA. More importantly, the generalized Psychophysiological Interactions analysis revealed significantly stronger functional connectivity between rFFA (used as a seed ROI) and right OFA in double face compared to normal face condition. Conclusion: These results suggest a recurring integration process in a "double face" illusion, which enhanced the activation in the fusiform face area and its functional connectivity with the occipital face area.

56.314 THE EFFECT OF FIXATION LOCATION ON FACE PERCEPTION IN YOUNGER AND OLDER ADULTS

Eric Cui[1,2] (mcui@research.baycrest.org), Farhan Vaheed[3], Matthew Clark[3], Björn Herrmann[1,2], Allison Sekuler[1,2,3]; [1]Rotman Research

Institute, Baycrest Academy for Research and Education, North, ON, Canada, [2]Department of Psychology, University of Toronto, Toronto, ON, Canada, [3]Department of Psychology, Neuroscience, and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada

Information of different facial features is distributed in various facial regions. For instance, the eye region, compared to other areas, is particularly informative for identity perception. However, not all observer focus on the eye region for this purpose. There are inter-individual differences in the region in which younger individuals fixate for identity perception: some preferred the eyes, the nose for the others. Fixation patterns also differ between age groups: older adults have more fixations on the lower half of the face compared to younger adults. However, when fixation is restricted to specific regions, younger adults demonstrated the best identity perception when fixation was restricted to the eye. Where do younger and older adults look on faces, and can they show improved performance in recognizing faces when focusing on specific facial regions? The first objective of this study is to describe and compare the optimal fixation locations of younger and older observers. The second objective is to investigate how restricting fixation location interacts with age on face perception. Each trial started with the presentation of a fixation square, followed by the target face, and a face selection task. The off-face condition began with a fixation square outside the face's anticipated display area. The on-face condition started with a fixation square on the forehead, eye, nose, or mouth. Preliminary analysis showed age-related in both conditions (20 younger and 20 older adults). Restricting the fixation location seems to reduce these age-related differences, with older adults showing improved performance, particularly at the eye and nose locations. Notably, an age-related difference emerged at the nose location, where older adults benefited more from nose restriction fixation than younger adults. Both younger and older adults showed greater performance in face perception when the initial fixation was directed to the eye and the nose region.

56.315 THE INTEGRATION OF INFORMATION FROM THE LEFT AND RIGHT HALVES OF THE FACE IN HUMAN AND ARTIFICIAL NEURAL NETWORKS.

Bartholomew P. A. Quinn[1] (bpaq500@york.ac.uk), David M. Watson[1], Kira Noad[1], A. Mike Burton[1], Timothy J. Andrews[1]; [1]University of York

In natural viewing, the left and right halves of the face are often divided across separate visual hemifields, each of which initially projects to the contralateral hemisphere. Accordingly, creating unified representations of faces involves the integration of information from both hemispheres. Here, we explored how information from the left and right half of a face is combined into a single representation. First, participants were asked to make familiarity judgements on composite faces, which combined the left and right halves of a famous face and an unfamiliar face. Consistent with the traditional composite face effect (in which the top and bottom halves of faces are combined), we found that accuracy was lower and response time was higher when the composites were aligned compared to when they were misaligned. This showed that the two halves of the face were automatically combined into a holistic representation. Next, we measured the neural correlates of this hemispheric integration in natural viewing using fMRI. We found consistently higher interhemispheric (e.g. rFFA – lFFA) compared to intrahemispheric (e.g. rOFA – rFFA) connectivity between regions of the face network. However, this interhemispheric bias was absent in early visual regions (V1-V3), suggesting an important role of interhemispheric communication in higher-level perceptual processing. Finally, we compared the similarity of left and right face halves in a deep convolutional neural network (DCNN) trained to recognize faces. We found that representations of left and right face halves were independent in the convolutional layers of the DCNN. However, there were similar representations of left and right halves of faces with the same identity in the fully-connected layers. Together, these findings reveal how information from the left and right halves of faces are combined holistically in human and artificial neural networks.

56.316 THE LEFT EYE AND UPPER EYE BIASES ARE LARGELY NOT FACE-SPECIFIC.

Nicolas Davidenko[1] (ndaviden@ucsc.edu), Mika Peer Shalem, Monique Crouse, James Pedroff; [1]University of California, Santa Cruz

Faces elicit a "left gaze bias", or a tendency to fixate on the left side (often on the left eye). However, recent work has shown that tilting a face clockwise or counterclockwise disrupts this bias, revealing a different bias to fixate whichever eye is higher (termed the "upper eye bias"; Davidenko, Kopalle, & Bridgeman, Perception 2019). In the current study, we investigated whether the left eye and upper eye biases are face-specific, by measuring whether and how they manifest with and without a face context. We recorded participants' eye movements while they reported whether a pair of symbols shown on a screen were asterisks (*) or crosses (+). In one block of trials, the two symbols were presented in the eye positions of parametric face drawings shown at a wide range of orientations (0, ±15, ±30, ±45, ±60, ±90, ±135, and 180 degrees). In the other block (counterbalanced order), the pair of symbols appeared in the same positions, but with no face present (i.e. on an empty background). The two symbols always matched, so participants could perform the task by looking at either one. An analysis of first fixations revealed (1) a strong bias to fixate the left symbol (observed in 16 of 19 participants) and (2) a strong bias to fixate the upper symbol (observed in all 19 participants). Critically, these biases were highly correlated and nearly identical in the face and no-face conditions (except for a slightly larger left bias in the face vs. no-face condition at some counterclockwise orientations). Overall, our results show that the "left eye" and "upper eye" biases manifest similarly whether there is a face present or not, suggesting that these biases are largely driven by general attentional mechanisms not related to face processing per se.

We thank Moorea Welch, Ghazaleh Mahzouni, Angela Huang, Joy Churchill, Lillian Nabkel, and Rishika Isanaka for their help collecting and analyzing data.

56.317 THE TUNING OF FACE PAREIDOLIA TO ORIENTATION STATISTICS

Benjamin Balas[1] (benjamin.balas@ndsu.edu), Abigail Berger[1], Mary Casey[1], Myra Morton[1]; [1]North Dakota State University

Face pareidolia is a ubiquitous phenomenon in which observers detect face-like structures in natural textures, arrangements of common objects, or other patterns. Pareidolic face detection likely reflects observers' internal representations of typical face structure and so may be tuned to the same low-level image features as typical face detection and recognition. Presently, we investigated whether face pareidolia, like other face recognition tasks, is tuned to horizontal orientation energy. We presented participants (N=43) with a series of fractal noise images (8 unique images of 1/f noise per condition) and asked them to report any pareidolic faces they saw in these patterns. To vary orientation statistics, we first generated isotropic noise images that were not biased in favor of any orientation passband and then applied orientation filtering to create horizontal and vertical versions of each stimulus. These images were printed on paper and presented to participants in a counterbalanced blocked design. Participants were instructed to examine each image and indicate the presence of any pareidolic faces by circling them using a marker, including a cartoon hat to indicate the orientation of the pareidolic face on the page. We predicted that observers would find more pareidolic faces in isotropic and horizontally-filtered images than vertically-filtered images. We counted the pareidolic faces detected in each image by our participants and analyzed this count data using a mixed-model analysis that included orientation as a fixed effect and participant as a random effect. This analysis revealed significantly negative slopes associated with the horizontal and vertical conditions relative to the isotropic condition, indicating that orientation filtering reduces the rate of pareidolic face detection, but vertical filtering is no more detrimental than horizontal filtering. Pareidolic face detection may thus depend on richer image structure (broadband orientation energy in particular) than mechanisms for face identification and categorization.

Tuesday, May 21, 2:45 – 6:45 pm, Banyan Breezeway

Visual Search: Memory, search templates

56.318 A COMPREHENSIVE COMPARISON OF ATTENTIONAL TEMPLATES MAINTAINED IN WORKING MEMORY AND LONG-TERM MEMORY

Huixin Song[1] (huixin_song@163.com), Xinran Chen[1], Beibei Cai[1], Yingtao Fu[1], Mowei Shen[1], Hui Chen[1]; [1]Zhejiang University

Attentional templates maintained in either working memory (WM) or long-term memory (LTM) can guide visual search. However, a definitive conclusion remains elusive regarding the attentional guidance efficiency of WM and LTM templates, attributed to diverse experimental designs and measures. This study systematically compared the attentional guidance of WM and LTM templates across various contexts (independent/competitive) and template loads, employing a unified and canonical measurement - search slope. In the first section, the comparison was made in an independent context using a blocked design, preventing interference between WM and LTM templates. Results consistently indicated similar search efficiency and sensitivity to load manipulations for both WM and LTM templates (Experiments 1 and 2). Moreover, we ruled out the alternative explanation that the observed comparable search efficiency resulted from retrieving LTM templates into the WM system during search (Experiment 3). Instead, it indicated that LTM templates directly guided search with an efficiency equivalent to WM templates. The second section introduced a competitive context with a mixed design, incorporating interference or competition between WM and LTM templates. Findings revealed that compared to LTM templates, WM templates exhibited superior search efficiency and were less impaired by increased template loads, suggesting WM templates hold an advantage in the competitive context (Experiment 4). Furthermore, it was demonstrated that this advantage is a result of an unequal prioritization between WM and LTM templates in the competitive context (Experiments 5 and 6). Overall, these results suggest that attentional templates maintained in WM and LTM are fundamentally the same in their ability to guide search, but their search efficiency may be influenced by their relative priorities in contexts involving interference or competition.

This work was supported by grants from Science and Technology Innovation 2030-“Brain Science and Brain-like Research” Major Project (2022ZD0210800), National Natural Science Foundation of China (32171046, 32071044, 32200844)

56.319 ARE SEARCH TEMPLATES TARGET-OBJECT RECONSTRUCTIONS?

Seoyoung Ahn[1] (ahnseoyoung@gmail.com), Hossein Adeli[2], Gregory Zelinsky[1]; [1]Stony Brook University, [2]Columbia University

Search theory heavily relies on the concept of a template, an internal representation of the target that, via a matching process to a visual input, creates a top-down attention biasing. The search template was originally conceptualized as being specific to a given object, but over the years this definition broadened to include the features of a target category. Exploiting recent generative methods, we suggest re-conceptualizing the search template yet again, thinking of it now as a fully-generated target object residing in peripheral vision and not just a collection of features. Our approach is to generate potential target-object appearances in degraded peripheral pixels. For example, when searching for a mosquito, our attention may be drawn to any small, roundish-shaped objects because they provide an ideal canvas for generating or attaching limbs. We used a Generative Adversarial Network (GAN)-based method to reconstruct peripheral objects so that they more closely resemble the typical appearance of the target category. We quantified the extent of pixel changes this reconstruction requires and tested whether the "reconstruction cost" accounts for target guidance in both digit and natural object-array search tasks. Our model, even though not explicitly trained for target detection, exhibited remarkable performance (~%90 accuracy) in locating target objects in blurred peripheral input, outperforming an object detector baseline. Moreover, the model exhibits a strong behavioral alignment with human eye-movements collected during the same task. Our model explained attention guidance comparably or significantly better than the detector in both target-present and target-absent conditions (Our Pearson’s $r = 0.891$, $p = 0.013$, and Detector's $r = 0.911$, $p = 0.012$ for target-present; Our's $r = 0.332$, $p = 0.052$, and Detector's $r = 0.134$, $p = 0.056$ for target-absent). Our work suggests that the target template may be an internal generation of a potential search target in peripheral vision.

This work was supported in part by NSF IIS awards 1763981 and 2123920 to G.Z. and by a grant to S.A from the American Psychological Association.

56.320 ATTENTIONAL TEMPLATE ACTIVATION AND SWITCH COSTS DURING PREPARATION FOR PREDICTABLE MULTIPLE-COLOUR SEARCH

Ziyi Wang[1] (ziyi.wang5@durham.ac.uk), Martin Eimer[2], Anna Grubert1[1]; [1]Durham University, [2]Birkbeck, University of London

Visual search for known objects is guided by target templates that are held in visual working memory. Such attentional templates are activated in preparation for search. Previous studies have shown that during search for two randomly presented colours, multiple attentional templates can be activated in parallel. But can multiple templates be activated strategically when target colours are repeated versus switched predictably across trials? We presented observers with series of search displays (every 1600ms) which contained two target-colour bars (AB; red and green) and four differently coloured nontargets. Following an AABB target colour sequence, participants reported the orientation of the target-colour bar in each trial. There were clear behavioural switch costs, with delayed target reports on colour-switch trials. Between search displays, we presented a constant stream of probe displays (every 200ms) each of which contained a colour singleton in one of the two target colours (ab; red or green) and five grey items. We measured N2pc components (electrophysiological markers of attentional capture) in response to these colour probes to track the time course of template activation prior to search. In colour-repeat trials, probes that matched the upcoming target colour (aA) produced reliable N2pcs from 800ms prior to search. However, in colour-switch trials, such probes only triggered N2pcs when they immediately preceded a search display. These differences in probe N2pc patterns suggest that the behavioural swich costs are a product of delayed template activation in colour-switch trials. Probes in the search-irrelevant target colour (bA) never triggered N2pcs, suggesting that behavioural switch costs were not caused by competition between simultaneously activated templates. Overall, our results demonstrate perfect colour selectivity and strategic control (i.e., serial template activation) during a two-colour AABB search task, and link colour switch costs to the time course of search template activation.

This work was funded by research grants of the Leverhulme Trust (RPG-2020-319) awarded to AG and the Economic and Social Research Council (ES/V002708/1) awarded to ME.

56.321 MULTIPLE IDENTIFIED ITEMS CAN BE SIMULATENOUSLY TESTED AGAINST THE TARGET TEMPLATES IN MEMORY IN THE HYBRID VISUAL AND MEMORY SEARCH

Yujie Zheng[1] (yjzheng@zju.edu.cn), Jiafei Lou[1], Zhi Li[1]; [1]Zhejiang University

Most of the search tasks in daily life are a hybrid of visual and memory search. The interplay between the visual and memory search in a hybrid search task can be modeled by a three-stage architecture: 1. visual selection; 2. identification/categorization; 3. target comparison. It is suggested that the visual items are selected sequentially but they can undergo simultaneous identification/categorization, because the pace of visual selection is fast but the pace of object recognition is slow. In the present study, we examined whether multiple identified items can be simultanenously tested against the target templates in memory. In Experiment 1, observers look for either one or two target types in an RSVP stream. The SOA of items is varied to measure a threshold. For search for one target, threshold was 89 ms. When look for either of two targets, threshold is 192 ms. The threshold difference (103 ms) served as a baseline. In Experiment 2, observers look for one or two types of target in a search array. If they had to test each target type separately, we might expect a big jump in the slope of the RT x Set size function, on the order of the baseline measured in Experiment 1. However, the slope difference was only 13 ms/item. These findings suggest that multiple identified items can be simultanenously tested against the target templates in memory, which extends our understanding of the mechanisms of hybrid search.

56.322 ATTENTIONAL SELECTION IS THE GATEKEEPER TO VWM

Zachary Hamblin-Frohman[1] (z.hamblin@utoronto.ca), Jay Pratt[1]; [1]University of Toronto, [2]University of Toronto

Visual attention and visual working memory (VWM) are intertwined processes that allow navigation of the visual world. These systems can compete for highly limited cognitive resources, creating interference effects when both operate in tandem. Previous research has shown that selectively attending, compared to non-selectively attending, an item causes obligatory interference with concurrently maintained VWM information. This finding may reflect that selectively attended items are automatically encoded into VWM. The current study examines this proposal by utilizing the procedures of the memory-driven capture paradigm. If an item is stored in VWM, then attention is captured by feature-matching items in subsequent tasks, and importantly even if they are distractors. On Trial 1, participants searched for a diamond shape with a varying colour. The target diamond was presented either alone (non-selectively attended condition) or among differently coloured non-targets (selectively attended). On Trial 2, the diamond-target and non-targets were the same colour, and one of the non-targets now had a singleton colour. This distractor colour could either match the colour of the diamond target from Trial 1 or was a novel colour. If a selectively attended item is automatically encoded into VWM, then it the feature-matching distractor capture on Trial 2 (measured via Eye-movements and RTs) should be higher for the selectively attended colour (Trial 1). This capture should also be above and beyond the effects of feature priming from the non-selectively attended Trial 1 colour. The results support this finding, the difference between matching and novel distractor colours are larger for the selectively attended condition compared to the non-selective attention condition. This study displays the effects of memory-driven capture in a task where participants were never required to encode stimuli into VWM and suggests that selective attention leads to obligatory VWM encoding.

56.323 NO SIGNS OF INTERFERENCE BETWEEN LEARNED ATTENTIONAL SETS

Seth Marx[1] (seth.a.marx@vanderbilt.edu), Geoffrey Woodman[2]; [1]Vanderbilt University, [2]Vanderbilt University, Department of Psychology & Neuroscience Program, Vanderbilt Vision Research Center, Tennessee 37240, USA

Humans guide visual attention to different targets as they move through different environments (i.e., going from a school zone to a construction zone while driving). To monitor how observers relearn which target features to prioritize across such contextual shifts, we had observers perform a visual search task where target shapes are presented in different colors with distinct probabilities (e.g., 33%, 26%, 19%, 12%, and 5%). Then, we periodically shifted the color-to-probability mapping. Here we show that people rapidly relearn complex attentional priority maps, and that they show no evidence for interference as they shift from old to new attentional sets, even when the new color mappings are in direct conflict with the old. Our findings suggest that learning to control attention may be interference free, unlike other forms of visual learning.

Grants were provided by the National Science Foundation (BCS-2147064), and NEI (P30-EY08126).

56.324 SEEING EYE TO "EGG": CAN ATTENTION AND MEMORY BE IMPACTED BY "SOCIAL" OBJECTS?

Dana Hayward[1] (dana.hayward@ualberta.ca), Arianna Cook[1]; [1]University of Alberta

An oft-used distinction when studying the effect of social content on cognitive processes is to create a division between people/faces as social, and objects as non-social; however, this practice presumes that objects lack social value. In contrast, evidence from other research domains suggest that objects can communicate social information, such as CDs or cars which can communicate information about the owner's identity. Thus, here we explored whether attentional and spatial memory biases exist for objects with a social versus non-social value. 84 participants (either primed to think about objects as social or not) engaged in a visual search task and a memory task that included objects belonging to one of three categories: (i) identity objects—such as a menorah—which are affiliated with an individual's identity, (ii) situation objects—such as a board game—which are used in social situations, and (iii) neutral objects—such as a toothbrush—which are not reflective of one's identity nor used in social situations. Participants were (i) faster to locate identity and situation objects compared to neutral objects, (ii) were the most accurate at remembering the location of the situation objects, and (iii) the social prime was not necessary to show differences in attention and memory for the three object categories. These findings demonstrate that not all objects are attended to or remembered equally, and have important implications when considering the objects used for drawing conclusions about social versus non-social processes.

NSERC (DAH); URI (AKC)

TUESDAY, MAY 21, 2:45 – 6:45 PM, BANYAN BREEZEWAY

Visual Search: Mechanisms, models

56.325 COMBINING SURFACE REFLECTANCE AND MOTION CUES IN PERIPHERAL TARGET DETECTION

Yunyan Duan[1] (yunyan.duan@tu-darmstadt.de), Swantje Mahncke[1], Thomas S.A. Wallis[1,2]; [1]Technical University of Darmstadt, [2]Centre for Mind, Brain and Behaviour (CMBB), Universities of Marburg, Giessen and Darmstadt, Germany

Humans are able to search for a target in cluttered scenes more easily when it contrasts with its background. However, the way in which humans combine different visual cues to detect a pop-out target is not fully understood, particularly for peripheral viewing in displays with partial occlusions and clutter. Here, we focus on surface reflectance and motion. Inspired by so-called "dead-leaves stimuli", we present an array of ellipses with different surface reflectance (luminance and color) and motion features, and define a target as a group of spatially-localized leaves with similar visual properties as the background leaves except for the cues of interest. A trial could contain a target defined by a single cue, a target defined by two cues jointly, or two targets each defined by a single cue, leading to single-cue, cue-combination, or cue-conflict conditions. Participants clicked on the center of the target with a mouse if they detected it, or right clicked if they detected nothing. A pre-experiment using a staircase procedure was used to titrate the cue difficulty before the formal experiment. If humans independently combine the two cues in the detection task, the cue-conflict condition should be as easy as the cue-combination condition, and better than single-cue conditions. Alternatively, if humans sum information from both cues for detection, the cue-combination but not the cue-conflict condition should outperform single-cue conditions. Results showed that the presence of both cues lead to improved performance compared with single-cue conditions, with better detectability in the cue-combination than the cue-conflict condition. However, given that participants could detect either of the two targets in the cue-conflict condition, the presence of both cues also drove additional accuracy improvement than expected by pure summation. The combination of surface reflectance and motion is overall better explained by a summation model than an independent-combination model.

This work is founded by the European Union (ERC, SEGMENT, 101086774).

56.326 DEEP LEARNING AND VISUAL SEARCH: USING RAW EYE MOVEMENT DATA, CONVOLUTIONAL NEURAL NETWORKS GENERATE TARGET-LOCATION PREDICTIONS IN LINE WITH EXPERIMENTAL MANIPULATIONS

Nicholas Crotty[1] (ncrotty@trincoll.edu), Nicole Massa[2], Michael Grubb[1]; [1]Trinity College, [2]Massachusetts General Hospital

Eye tracking during visual search generates spatiotemporally rich but complex data. Traditional analysis methods typically employ metrics like the proportion of trials containing a saccade to a target (or distractor), dwell time on important stimuli, etc. However, such approaches leave out potential information contained in the raw eye data. Here, we asked if deep learning advancements might aid scientists in navigating this trade-off. A convolutional neural network (CNN) is a type of artificial neural network that can identify key features of input data, then use these features to sort (classify) unlabeled inputs into their appropriate groups. CNNs can learn from this classification process, using mistakes to help further define the portions of data that are most informative in determining what group an input belongs to. Although CNNs are commonly applied to images, they can generate predictions from other complex inputs like timeseries data. In a pre-existing dataset, participants searched for a color-defined target amongst 5 differently-colored distractors. We built a CNN that receives the raw x,y timeseries data and predicts which of the six locations contained the target on each trial. We trained the CNN on 2/3 of the data and validated on the rest. In short, the CNN performed well, predicting target location substantially above chance (67% vs. 17%). In our study, participants were pre-cued with reliable information about target color on half the trials (validity: 100%) and with unreliable information on the other half (validity: 50%). Prediction accuracies of two new CNNs, trained and validated on data from the two precue conditions separately, reflected this experimental intervention: greater classification accuracy in reliable (70%) than unreliable trials (63%). Bootstrapped error bars and subject-level null hypothesis testing confirmed the statistical reliability of this difference. These findings highlight the potential of CNNs as a novel analysis method.

NSF-2141860 CAREER Award to Michael Grubb

56.327 MODELING NEGATIVELY ACCELERATING SEARCH SLOPES IN A RELATIONAL SEARCH TASK

Rachel Heaton[1] (rmflood2@illinois.edu), John Hummel[1]; [1]University of Illinois Urbana-Champaign

Logan (1994) showed that visual search for a target defined by the spatial relations among its parts (e.g., a plus above a dash among dashes above plusses) gives rise to a steep linear search slope. In seven previously reported experiments, we (Heaton et al., 2021; Heaton, 2023) conceptually replicated Logan with slightly different stimuli (colored Xs and Os in above-below relations) and with a larger range of set sizes than those used by Logan. These experiments revealed steep but negatively accelerating functions of response time vs set size. Here we report simulations with the most recent version of the CASPER model of visual search (Heaton, 2023) that account for the negative acceleration in these search functions in terms of the role of an emergent feature in the negative space between the X and the O in our stimuli. The influence of this feature is modulated by factors that affect perceptual grouping strength. In the absence of this feature, CASPER produces steep linear functions of response time over set size. These results suggest that the visual system is capable of searching for targets defined only by spatial relations but will exploit any simple visual features that are confounded with the correct response, and in so doing highlight the difficulty of conducting search experiments that we can be confident depend exclusively on relational processing.

This research was supported by funding from Robert Wickesberg and the Department of Psychology at the University of Illinois-Urbana-Champaign.

56.328 THE ATTENTIONAL TEMPLATE THEORY OF MULTIPLE-TARGET SEARCH ERRORS

Stephen H Adamo[1] (sadamo13@gmail.com), Joseph Schmidt[1]; [1]University of Central Florida

Subsequent search miss (SSM) errors (also known as Satisfaction of Search)— worse target detection after finding a first target, is a common multiple-target search error. To explain SSMs, the Attentional Template theory predicts that a detected first target is utilized as an attentional template for subsequent targets, making observers more likely to miss dissimilar targets compared to similar targets in the same image. Across three behavioral and one EEG study, we tested the predictions of the Attentional Template theory by: 1) investigating how second target accuracy is affected by perceptual similarity between targets, 2) determining how long target similarity effects last, and 3) how the P300 is affected when the similarity between a first and second target is reduced. Observers were asked to search for up to two T and/or L-shaped targets that could differ in shape and rotation. Observers missed more second targets when they were dissimilar (e.g., a different shape and rotation), and this effect was expressed at shorter search durations (i.e., 1 second or less). For the P300, 1) observers had a larger P300 when targets were similar compared to dissimilar and when observers found one target (regardless of number or targets present), 2) a smaller P300 when targets were dissimilar compared to when targets were similar or observers found one target and 3) no difference in the P300 when only one target was found (regardless of similarity or whether one or two targets were present). Together, these findings broadly support the predictions of the Attentional Template theory and suggest that: 1) a first target, acting as an attentional template, can underlie SSMs, 2) when SSMs occur, attentional target recognition, as measured by the P300, is not allocated to second target processing, and 3) when targets are dissimilar, less attentional target recognition of second targets occurs.

National Cancer Institute (1K99CA267163-01A1)

56.329 GETTING MORE OUT OF RESPONSE TIME MEASURES: DELINEATING SEPARABLE MOTOR AND COGNITIVE SUBCOMPONENTS OF RESPONSE TIME VIA DRIFT DIFFUSION MODELING

Justin Grady[1] (justgrad@gwu.edu), Audrey Siqi-Liu[1], Sarah Malykke[1], Patrick Cox[2], Dwight Kravitz[1], Stephen Mitroff[1]; [1]The George Washington University, [2]Lehigh University

Cognitive psychology studies have used a wide range of experimental designs to make important contributions to the understanding of cognitive processing. Most studies, however, have relied on a rather coarse measure of response speed—how long it takes participants to press a key on a keyboard. While straightforward, this measure may confound cognitive processes with motor initiation, stimulus encoding, and other potential sources of noise. Recent work from our lab (Kramer et al., 2021) introduced a “touch and swipe” response as part of an object sorting task. In this task, participants tap on objects as they appear on the screen (touch time) and then make a swiping motion to sort them into one of two bins based on category membership (swipe time). Critically, this work suggests that touch time may reflect a prepotent motor response or impulsivity, while swipe time may better reflect cognitive processing time. The current study aimed to further characterize the nature of the touch and swipe time measures using drift diffusion modeling, which derives separable decision making components from response time distributions. Specifically, this study assessed whether estimates of non-decision time and drift rate (speed of evidence accumulation) scale differentially with touch and swipe time measures. Non-decision times scaled positively with relatively longer touch times, while drift rate scaled positively with relatively longer swipe times. These modeling results provide converging evidence for the hypothesis that touch and swipe times are meaningful subcomponents of response time that map onto separable underlying processes—swipe time relates to evidence accumulation, which is likely more informative for psychological experiments of cognitive performance, while touch time relates to non-decision time which is more reflective of motor components. Taken together, these results indicate that the touch and swipe method for recording responses is a potentially powerful experimental design.

US Army Research Laboratory Cooperative Agreements #W911NF-21-2-0179, #W911NF-23-2-0210, & #W911NF-23-2-0097

56.330 COMPARING CONVOLUTIONAL NEURAL NETWORKS TO TRADITIONAL MODELS OF COVERT ATTENTION DURING VISUAL SEARCH

Ansh K. Soni[1] (asoni@ucsb.edu), Sudhanshu Srivastava[1], Miguel P. Eckstein[1]; [1]University of California - Santa Barbara

Introduction: Performance degradation in Visual Search with an increased number of distractors (the set-size effect) is often used to make inferences about the properties of covert attention. Recently, studies have utilized Convolutional Neural Network (CNN) models (Srivastava, 2023; Nicholson, 2022; Poder, 2022), although their relationship to traditional models is not well understood. Our goal is to compare CNN models to traditional models of visual search. Methods: We implement six models of covert attention during visual search to predict target detection accuracy (yes/no) in feature and conjunction tasks. For each model, we adjust the two feature searches (line angle or luminance) to match model performance. These feature values were used to build the conjunction task (angle and luminance). Three of the models are image-computable, acting directly on the pixels of the images: Image computable Bayesian ideal observer (IC-BIO), a fully-trained 5-layer CNN, and a large network trained on image classification (VGG-16 network) with transfer learning. The remaining three models operated on assumed extracted feature values (normally distributed activations from the target and each distractor): a Signal Detection Theory (SDT; Green and Swets, 1966) model without or with capacity limits (SDTc; Poder, 2019) and a Guided Search Accuracy model (GSA; Wolfe 1989; Eckstein, 2000). Results: For feature and conjunction search, the CNN and VGG-16 models showed similar set-size effects to unlimited capacity models (IC-BIO/SBIO and SDT) and smaller than the models with capacity limitations (SDTc and GSA with serial attention). The accuracy degradation from feature to conjunction search was lowest to highest for: IC-BIO, CNN, VGG-16, SDT (assuming independent processing of features; Ecktein, 1998), and SDTc. Conclusion: Our findings benchmark newer CNN models against traditional search models, showing a correspondence between CNN set-size effects and Signal Detection/Ideal Observer models but distinct feature/conjunction search accuracy relationships.

56.331 BAYESIAN HEURISTIC DECISION ANALYSIS OF VISUAL SEARCH

Wilson S. Geisler[1] (w.geisler@utexas.edu), Anqi Zhang[1]; [1]University of Texas at Austin

Quantitative models of perceptual performance often include limitations due to the physical properties of the stimuli, limitations due to sensory processing, and the assumption of rational (Bayes optimal) decision processes applied to the outputs of the sensory processing. Determining the Bayes optimal decision process is important for characterizing the available information, how that information varies with stimulus conditions, and what specific computations achieve optimal task performance. Although there are tasks where humans reach near optimal performance there are many where they fall short. How is it that humans achieve near optimal performance in some cases and yet clearly fail in other cases? One approach to this question is Bayesian heuristic decision analysis (BHDA), which systematically explores the space of heuristic decision processes, using Bayes optimal decision processes as the benchmark. If there are biologically plausible heuristics that approach Bayes-optimal performance, then they provide realistic testable hypotheses for how humans can achieve near optimal performance. Conversely, if there are no such biologically plausible heuristics, then the expectation is non-optimal performance. We demonstrate this analysis on visual search tasks. In covert search with a single known target, the optimal decision rule is to weight the feature responses at each potential target location by the discriminability (d') at that location, add the log of the prior probability at that location, and then pick the location with the maximum value. We find that a wide range of simple decision heuristics closely approach optimal accuracy, even though these heuristics largely ignore the actual variation in discriminability and prior probability. These results help explain the high efficiency of humans in this task. Bayesian heuristic decision analysis also shows that efficient fixation selection in overt search can be achieved with realistic limitations in memory and posterior updating. Several other important implications of this approach will be described.

Supported by NIH grants EY11747 and EY024662

56.332 STEPPING INTO THE SAME RIVER TWICE: ARE MISS ERRORS IN VISUAL SEARCH DETERMINISTIC OR STOCHASTIC?

Aoqi Li[1] (aoqi.li@manchester.ac.uk), Johan Hulleman[1], Wentao (Taylor) Si[2], Jeremy Wolfe[3,4]; [1]University of Manchester, [2]Bates College, [3]Brigham and Women's Hospital, [4]Harvard Medical School

Observers make errors in visual search, whether in a lab experiment or a real-life task. Those errors can be categorized as “deterministic” or “stochastic”. If errors are deterministic, errors committed once will definitely be repeated again. Alternatively, errors can be “stochastic”: occurring randomly with some probability. An error would lie in between these extremes if it is likely, but not guaranteed to occur a second time. To identify the nature of miss errors in a simple T-vs-L visual search task, we presented each search display twice in random sequence. The miss rate, P1, for the first copy of the display and the miss rate, P2, for the second copy were calculated, as was the proportion of cases where both copies were missed, P12. Purely stochastic errors would predict that P12=P1*P2. Purely deterministic errors will lead to P12=min(P1,P2). If errors are a mix of stochastic and deterministic, P12 will fall between these two predictions. In Experiment 1 where the letters were clearly visible, the errors were almost completely stochastic. An error made on the first appearance of a display did not predict that an error would be repeated on the second appearance. In Experiments 2a and 2b where the visibility of the letters was manipulated, the errors became a mix of stochastic and deterministic. Lower contrast targets produced more deterministic errors. In Experiments 3a, 3b and 3c, we tested several interventions with the goal of finding a 'mindless' intervention that could effectively reduce errors without needing to know the answer in advance. An almost mindless intervention that knew the location but not the identity of items (Exp 3c), succeeded in reducing deterministic errors. This gives some insights into possible methods for reducing errors in important real-life visual search tasks, where search items may not be clearly defined and visible.

JMW was supported by NIH-NEI: EY017001, NSF: 2146617, and NIH-NCI: CA207490. AL and JH were supported by UKRI grant ES/X000443/1.

56.333 MECHANISMS OF A CONVOLUTIONAL NEURAL NETWORK THAT LEARNS TO COVERTLY ATTEND

Sudhanshu Srivastava[1] (sudhanshu@ucsb.edu), Miguel P Eckstein[1]; [1]University of California, Santa Barbara

The cueing task is one of the most prominent paradigms to study covert attention. Observers are quicker and more accurate in detecting a target when it appears with a cue (valid trials) than when it appears opposite to the cue (invalid trials). How neuronal populations across the visual hierarchy progressively represent and integrate visual information across the target, cues, and locations to give rise to the behavioral cueing effect is not well-understood. To gain a theoretical understanding of the plausible system-wide neuronal computations and mechanisms mediating the cueing effect, we analyze the response properties of 180k neurons per network across layers of ten feedforward Convolutional Neural Networks (CNN). The CNNs are trained on noisy images without any explicit attention mechanism and show human-like benefits of cues on detection accuracy. Early layers show retinotopic neurons separately tuned to target or cue with excitatory or inhibitory responses. Later layers show neurons jointly tuned to both target and cue and integrate information across locations. Consistent with physiological findings, we find increasing influence of the cue on target responses in deeper layers in the network and computational stages similar to those of a Bayesian ideal observer, but with more gradual transitions. The cue influences the mean neuronal response to the target and distractor, and changes target sensitivity with two mechanisms: integration of information across locations at the dense layer, and interaction with the thresholding Rectified Linear Unit (ReLU) in the last convolution layer. We find novel neuronal properties not yet reported in physiological studies: cue-inhibitory neurons, inhibitory cue influences on target neurons, and location-opponent cells, which are target-inhibitory at one location and target-excitatory at the other. Together, our analyses illustrate a system-wide analysis of the neuronal computations that might give rise to behavioral cueing effects and provide a theoretical framework to inform physiological studies.

Tuesday, May 21, 2:45 – 6:45 pm, Banyan Breezeway

Perceptual Organization: Neural mechanisms, models

56.334 HIGH-DIMENSIONAL LATENT MANIFOLDS AS PREDICTORS OF INDIVIDUAL DIFFERENCES IN NATURALISTIC MOVIE VIEWING

Chihye Han[1] (chan21@jhu.edu), Michael Bonner[1]; [1]Johns Hopkins University

The human visual system is adept at processing complex, high-dimensional sensory data. A prominent theory proposes that the visual system accomplishes this by transforming high-dimensional sensory inputs into simpler, low-dimensional representations. However, recent theoretical and empirical work suggests that the dimensionality of visual cortical representations may be more extensive than previously thought. We hypothesize that even low-variance dimensions in cortical population activity are critical to human vision and that individual differences in visual experience are captured by these high-dimensional codes. To investigate this possibility, we used a recent method, known as cross-decomposition, to identify the shared high-dimensional signal between pairs of individuals. We applied this method to publicly available fMRI data collected from forty participants while they viewed four short movies. We first computed cross-validated covariance spectra between subject pairs, creating a matrix that reflects individual differences in the high-dimensional latent space. We then measured the reliability of these individual difference matrices by computing their correlation across pairs of movies. Our analysis revealed a long-tailed spectrum of reliable, low-variance dimensions shared among individuals in the ventral visual stream, and we found that these high-dimensional signals were highly consistent for the same subject pairs across different movies. These findings suggest that there are rich and reliable individual differences in the high-dimensional representational manifolds that underlie naturalistic visual experience.

56.335 INVERTED VISUAL CODING ACROSS CATEGORY-SELECTIVE VISUAL AREAS

Adam Steel[1] (adamdanielsteel@gmail.com), Peter Angeli[1], Edward Silson[2], Caroline Robertson[1]; [1]Dartmouth College, [2]University of Edinburgh

Traditional models of brain function propose a shift from retinotopic to amodal, abstract coding as visual information progresses anteriorly from visual cortex towards memory structures. However, recent evidence challenges this conception, suggesting that memory-related brain areas implement a retinotopic code characterized by spatially-selective negative responses during population receptive field modeling (-pRFs), and this code structures interactions among category-selective brain areas involved in scene perception and memory (Steel*, Silson* et al., 2023, Nat. Neuro.). Here, we investigated whether -pRFs are present within or anterior to other visual-category preferring areas (beyond scene areas), or whether -pRFs uniquely appear in regions specialized for scene memory. We computed pRFs for all subjects in the Natural Scenes Dataset (Allen et al., 2022) and compared -/+ pRF concentrations in category-preferring regions in ventral temporal cortex for scenes (anterior & posterior PPA), faces (iOG-, pFus-, mFus-faces), bodies (FBA-1 & 2), and words (OWFA, VWFA-1 & 2). Importantly, for scenes, we replicated our previous observation: -pRFs were preferentially concentrated in anterior versus posterior PPA ($p<0.005$), and the lateral place memory area (individually localized using resting-state fMRI) compared to the scene perception area OPA ($p<0.001$). Next, we examined the prevalence of -pRFs in and anterior to other category-selective visual areas. Face- and body-selective areas exhibited no differences in -pRF concentration between posterior and anterior functional regions (all ps > 0.09). Interestingly, for word-selective areas, the concentration of -pRFs increased up the processing hierarchy from OWFA to VWFA-1 and 2 (ps < 0.05). Crucially, the visual field preferences of -/+ pRFs in word-preferring areas were well-matched, supporting the notion of functional linkage. We propose that the -pRFs associated with visual areas may serve visual functions that demand perceptual-mnemonic interaction across the visual field such as navigation and reading.

R01MH130529

56.336 NEUROPHYSIOLOGY OF SYMMETRY PROCESSING IN MACAQUE VISUAL CORTEX

Pauline Audurier[1] (audurier@ohsu.edu), Vincent D. Costa[1], Robert M. Friedman[1]; [1]Oregon National Primate Research Center

In a world composed of inherent regularities, symmetry intricately shapes spatial relationships, guiding visual scene organization. As a highly salient visual feature, the perception of symmetry properties has been demonstrated in different species, notably in the support of object recognition and figure-ground segregation. While the cortical network underlying symmetry processing has been extensively studied and identified in humans, functional imaging of this network in non-human primates (NHPs) has been limited to a single recent study (Audurier et al., 2021). Using fundamental symmetry stimuli generated from wallpaper type textures this study revealed a comparable rotational symmetry processing network between humans and NHPs that recruits cortical areas beyond V1, including V2, V3, V3A, and V4. To deepen our understanding of the neural responses to symmetry, we applied a dual approach combining intrinsic optical imaging (IOI) and multisite neurophysiology of primary visual cortical areas. IOI recordings were performed in anesthetized rhesus macaques in hopes of revealing symmetry selective domains at the mesoscale and to identify functional domains (e.g. orientation and color) for the functional targeting of electrodes. Subsequently, multi-site linear electrodes were implanted in V1, V2, V3, and V4 to assess the roles and dynamics among these distinct visual areas to the encoding of different symmetry conditions. Analyses of raster plots have revealed a subset of neurons that differentially encode rotational symmetry with there being potentially a greater proportion of such neurons in V4 and V2 than V3 and V1. This finding supports previous findings. Our integrative methodology aims to unravel the intricacies of symmetry processing and contribute to a more comprehensive understanding of visual perception of primates.

56.337 SPARSE NULL CODES EMERGE AND DOMINATE REPRESENTATIONS IN DEEP NEURAL NETWORK VISION MODELS

Brian S. Robinson[1] (brian.robinson@jhuapl.edu), Nathan Drenkow[1], Colin Conwell[2], Michael F. Bonner[2]; [1]Johns Hopkins University Applied Physics Laboratory, [2]Johns Hopkins University

Representations in vision-based deep neural networks and biological vision are often analyzed from the perspective of the image features they encode, such as contours, textures, and object parts. In this work, we present evidence for an alternative, more abstract type of representation in deep neural networks, which we refer to as a “null code”. Through a series of analyses inspecting the embeddings of a range of neural networks, including different transformer architectures and a recent performant convolutional neural network, we observe null codes that are both statistically and qualitatively distinct from the more commonly reported feature-related codes of vision models. These null codes are highly sparse, have a single unique activation pattern for each network, emerge abruptly at intermediate network depths, and are activated in a feature-independent manner by weakly informative image regions, such as backgrounds. We additionally find that these sparse null codes are approximately equal to the first principal component of representations in middle and later network layers across all analyzed models, which means that they have a major impact on methodological and conceptual approaches for relating deep neural networks to biological vision. In sum, these findings reveal a new class of highly abstract representations that emerge as major components of modern deep vision models: sparse null codes that seem to indicate the absence of features rather than serving as feature detectors.

This work was supported by funding from the Johns Hopkins University Applied Physics Laboratory

56.338 THE EFFECT OF PUPIL SIZE ON NEAR-THRESHOLD DETECTION IS NOT MEDIATED BY ALPHA, BETA, OR THETA POWER

Veera Ruuskanen[1] (v.h.s.ruuskanen@rug.nl), Nico Böhler[2], Sebastiaan Mathôt[1]; [1]University of Groningen, [2]Ghent University

Larger pupils are associated with improved performance in visual detection tasks. This is often attributed to changes in level of arousal as indexed by pupil size. However, changes in pupil size also affect the amount and focus of light on the retina, potentially influencing detection independently of arousal. Such optical effects of pupil size on detection are rarely studied and poorly understood. The goal of this project was to gain a better understanding of how pupil size and arousal (as measured with EEG) are related to each other and to performance in a visual detection task. We collected EEG and pupil data while participants performed a task consisting of detecting faint luminance patches briefly flashed in peripheral vision. Stimulus parameters (contrast and spatial frequency) were adjusted with a staircase procedure to fix performance at 65% accuracy. We performed a mediation analysis within a structural equation modeling framework to investigate the direct and indirect effects present in the data. The results show that larger pupils and theta suppression are both linked to better performance. Pupil size is also positively correlated with power in the alpha and beta bands, with the latter being a stronger relationship. Importantly however, the mediation analysis showed that the relationship between pupil size and accuracy is not mediated by any of the neural measures. This suggests that pupil size affects detection performance independently of the kind of arousal that is expressed as fluctuations in power in the alpha, beta, and theta bands.

56.339 ASSOCIATION BETWEEN PROFICIENCY AND IDIOSYNCRATIC BIASES IN MEDICAL IMAGE PERCEPTION

Zhihang Ren[1] (peter.zhren@berkeley.edu), Julien Vignoud[2], Zixuan Wang[1], Heath Rutledge-Jukes[4], Tejas C. Sekhar[4], Kathy Fang[3], Erik P. Duhaime[4], David Whitney[1]; [1]University of California, Berkeley, [2]École Polytechnique Fédérale de Lausanne (EPFL), [3]Golden State Dermatology, Albany, CA, USA, [4]Centaur Labs, Boston, MA, USA

Previous work has confirmed that clinicians vary in their medical image perception proficiency, impacting diagnostic performance and patient outcomes. However, previous work focuses more on global metrics of performance, such as accuracy, specificity, and so on. This indicates that some clinicians are better than others, but it makes no prediction about how a clinician will perform given a particular image. Here, we tested the possibility that individual clinicians have idiosyncratic patterns of bias at the image-level. To address this, we analyzed around 750k malignancy discrimination judgments of 7,818 skin lesion images, from a pool of 1,173 observers including those with medical image training. We utilized a deep learning encoding algorithm to cluster similar skin lesions based on deep semantic features. This allowed us to analyze individual differences locally at the image-cluster level. First, we identified significant individual differences in image-selective diagnostic performance, revealed by patterns of bias in perceptual judgments that were idiosyncratic and stimulus-specific. When visualized, these patterns of perceptual bias reveal fingerprints of perceptual bias that characterize individual clinicians. Second, we confirmed that proficiency is associated with increased agreement among medical professionals. Finally, by isolating the most ambiguous images, we found that proficiency is associated with stronger idiosyncratic biases: as images become more challenging, more proficient individual observers have increasingly unique and precise patterns of bias in their perceptual decisions. Our results suggest a potential systematic cause of diagnostic errors at the level of individual clinicians, they expose potential mitigation strategies, and they have implications for multi-reader scenarios.

This work has been supported by National Institutes of Health (NIH) under grant #R01CA236793. And we appreciate Centaur Labs offers us the diagnostic data for analysis.

Tuesday, May 21, 2:45 – 6:45 pm, Banyan Breezeway

Perceptual Organization: Parts, wholes, groups

56.340 SPATIAL TUNING OF VISUAL RESPONSES TO SYMMETRIES IN TEXTURES

Yara Iskandar[1] (misk22@my.yorku.ca), Christopher Lee[2], Sebastian Bosse[3], Peter J Kohler[1]; [1]York University, [2]Image Engine Design Inc., [3]Fraunhofer HHI

Symmetry is a biologically significant visual feature that relies on the perceptual grouping of spatially separate elements. Symmetry has been shown to play a role in numerous domains of visual perception in both humans and other animals. Brain imaging studies have revealed that several regions in the visual cortex exhibit robust and precise responses to symmetry. Here we explored the spatial mechanisms that mediate symmetry perception by measuring Steady-State Visual Evoked Potentials using high-density EEG. Our stimuli were taken from a class of regular textures, known as wallpaper groups – a set of 17 unique combinations of symmetry types that represent the complete set of symmetries in 2D images. We focused on groups PMM, which contains reflection symmetry, and P4, which contains four-fold rotation symmetry, and generated exemplars from each group based on log-domain band-limited random noise patches. This approach allows us to manipulate both the spatial frequency content of the exemplars, and the scale of the lattice structure that is repeated to tile the plane in all wallpaper groups. Across 8 conditions, we varied spatial frequency between 1 to 8 cycles-per-degree (centre frequency of the noise patch), and the lattice scale between 1/12 and ½ (relative size of lattice to overall wallpaper). Consistent with previous studies we found that symmetry-specific responses were weaker overall for rotation compared to reflection. However, responses also exhibited clear evidence of spatial tuning, with low spatial frequency and small scale lattices generally producing the biggest responses for both wallpaper groups. Interestingly, reflection (PMM) and rotation (P4) symmetry elicited clearly distinct response patterns across the spatial frequencies and lattice scales, suggesting that the two symmetry types rely on distinct cortical mechanisms. Future studies will relate these findings to responses in distinct areas along the human visual processing hierarchy.

This work was supported by the Vision Science to Applications (VISTA) program funded by the Canada First Research Excellence Fund (CFREF, 2016–2023) and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada awarded to PJK.

56.341 VISUAL RESPONSES TO LOCAL VS. FULL-FIELD TEXTURES IN THE PRIMATE SUPERIOR COLLICULUS

Divya Subramanian[1] (divya.subramanian@nih.gov), Leor Katz[1], Richard Krauzlis[1]; [1]National Institutes of Health

Visual responses in primate superior colliculus (SC) are strongly modulated by salient low-level features. Higher-order salience, such as the presence of a distinct form, may also modulate SC neurons independent of low-level features. Here we tested whether SC visual activity was higher when a stimulus in the receptive field (RF) was distinct from, rather than part of, a background while holding low-level salience fixed. We compared the visual responses of neurons in macaque SC to orientation-filtered white noise textures presented either as full-field stimuli ("background" condition) or including an orthogonally oriented patch placed in the neurons' RF ("patch+background"). To further probe the interactions between the RF and the background, we also included isolated patches on blank backgrounds and full-field textures with blank RF patches ("hole"). All textures had the same average luminance, contrast, and spatial frequency. SC neurons ($n = 65$) were modulated by higher-order salience already in their earliest response phase. Activity in the patch+background condition was 65% higher compared to the background 40-60ms after stimulus onset ($p<0.01$, Wilcoxon rank-sum test), and increased to being 2.7 times higher in a later phase (70-120ms: $p<0.0001$). Early responses to the isolated patch were almost twice as high as the uniform background ($p<0.0001$) and 20% higher than the patch+background ($p<0.05$), reflecting rapid surround inhibition from the background. Responses to the hole were lower than to the background in the early phase ($p<0.05$), but 22% higher in the later phase ($p<0.05$), suggesting a second-order discrepancy detection process even in the absence of inputs to the RF. Our results show that the primate SC rapidly detects stimuli that are distinct from the background even when their low-level features are the same. This suggests that the SC may be involved in aspects of figure-ground segmentation, an intermediate step between low-level salience and higher-order object processing.

National Eye Institute Intramural Research Program at the National Institutes of Health ZIA EY000511

56.342 INVESTIGATING LOCAL AND CONFIGURAL SHAPE PROCESSING WITH STEADY-STATE VISUAL EVOKED POTENTIALS

Shaya Samet[1] (shayasam@my.yorku.ca), James H Elder[1], Nick Baker[2], Erez Freud[1], Peter J Kohler[1]; [1]York University, [2]Loyola University Chicago, USA

The perception of object shape underlies our ability to detect, recognize and manipulate objects. Both local shape (curvature) and non-local (configural) shape contribute, and recent work has used specialized stimuli and behavioural methods to dissociate these contributions. Here we used high-density EEG to explore the cortical mechanisms involved in both local and configural shape perception. Object shape silhouettes were presented during passive viewing in an SSVEP paradigm that allowed us to isolate differential brain processing between pairs of stimulus conditions. Stimuli included natural animal-shape silhouettes (upright or inverted), synthetic maximum-entropy shapes progressively matching local curvature statistics of natural shapes but lacking global (configural) regularities (Elder et al., 2018), and stimuli in which the top and bottom half have been flipped to disrupt configural shape, named Frankenstein stimuli (Baker & Elder, 2022). Our findings so far (n = 32) reveal differential activity in occipital and temporal cortices emerging 170–280 msec post-stimulus, influenced by both local curvature and global configural shape. We find clear effects of matching the local curvature statistics on brain processing in visual cortex, especially for the variance. However, even when controlling all the local statistics, responses to natural animal shapes are still quite distinct from the curvature-matched controls. Interestingly, the differential responses to natural animal shapes compared to curvature-matched controls is subject to an inversion effect, highlighting the potential influence of semantic and holistic processing on the measured responses. It is important to note, however, that inverted animals still produce measurable differential responses compared to curvature-matched controls, suggesting that some configural properties survive the inversion. Future work, including ongoing studies with the Frankenstein stimuli, will further explore what those properties are.

This work was supported by the Vision Science to Applications (VISTA) program funded by the Canada First Research Excellence Fund (CFREF, 2016–2023) and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada awarded to PJK.

56.343 UNCONSCIOUS PERCEPTION OF CONTINUITY DURING VISUAL SUPPRESSION

Zhilin Zhang[1] (zzhang5@gradcenter.cuny.edu), Tony Ro[1]; [1]City University of New York

While it has been shown that accurate perception of visual features, such as orientation and color, can occur without visual awareness, it is unclear if higher-order stimulus attributes, such as continuity, can also be unconsciously discriminated. In the present study, we assessed whether perceptual grouping by continuity can be unconsciously processed during suppression of visual awareness from backward masking or transcranial magnetic stimulation (TMS). Subjects viewed briefly presented cross-shaped target patterns formed by one horizontal and one vertical set of parallel lines, with one set appearing to occlude the other based on perceptual grouping by continuity of the orientation of the center lines. We first ensured accurate perception of the stimuli in a target-only session without any awareness suppression. Perception of the target was then suppressed in two groups of subjects with backward visual masking or TMS to early visual cortex. Subjects reported their perception of target visibility and then the orientation of the continuous, non-occluded set. Detection of the target was suppressed by both visual masking and TMS, as compared to the no-suppression condition. When subjects were unaware of the target, accuracy of discrimination by continuity configuration was significantly above chance. There was no significant difference in unconscious discrimination performance between the visual masking and TMS condition. The present study reveals accurate perceptual organization with grouping by continuity in the absence of visual awareness of the stimulus, suggesting that gestalt feature segregation based on both local and global stimulus properties can occur without phenomenological impression of the visual item.

NSF BCS 1755477 to TR.

56.344 GAMMA-BAND SYNCHRONIZATION IN VISUAL CORTEX INDUCED BY TACS PROMOTES CONTOUR INTEGRATION

Sang Wook Hong[1,2] (shong6@fau.edu), Yosun Yoon[1,2]; [1]Department of Psychology, Florida Atlantic University, [2]Stiles-Nicholson Brain Institute, Florida Atlantic University

The Rhythm Theory of cognition posits that various human perceptions and cognitions are mediated by the oscillatory activity of the human brain. Particularly, it has been demonstrated that gamma-band oscillations are involved in the integration of visual information, playing a significant role in object perception. Recently, we demonstrated that interocular grouping during binocular rivalry was enhanced by inter-hemispheric gamma-band synchronization induced by transcranial alternating current stimulation (tACS) (Yoon & Hong, 2023). This previous result suggests that contour integration across visual hemifields may be mediated by synchronized neural oscillation in the gamma-band frequency. To directly test this hypothesis, we investigated whether gamma-band synchronization between the left and right visual occipital cortices facilitates the detection of a colinear line composed of three short bars. The target line was presented within a square window (6 by 6 visual angle) where 397 distractor bars with random orientations were simultaneously presented. The location of the target was chosen from one of eight predefined positions along an imaginary circle (2.5-degree radius). Participants were instructed to report the orientation of the target line (left-tilted or right-tilted from the vertical orientation) as accurately and quickly as possible. Gamma-band synchronization between hemispheres was induced by applying 40 Hz tACS to the left and right visual areas corresponding to O1 and O2 in the international 10-20 system. The data revealed that the response time to detect the target line presented across two hemifields indeed decreased during stimulation and post-stimulation compared to the response times measured before stimulation (baseline). In the sham control condition, the response times for detecting the target line were not significantly different between pre-stimulation and post-stimulation. We also observed decreased response times for the within-hemifield target locations after gamma-band tACS. Together, our results indicate the causal role of gamma-band oscillation in contour integration within and between hemispheres.

56.345 BRAIN RESPONSES TO SYMMETRIES IN NATURALISTIC NOVEL THREE-DIMENSIONAL OBJECTS

Shenoa Ragavaloo[1,2], Nikan Movahedi[1,2], Peter J. Kohler[1,2]; [1]York University, [2]Center for Vision Research

Symmetries are prevalent in natural and man-made objects and scenes. During natural vision, symmetries in the world are subject to perspective-distortion and rarely produce symmetrical images on the retina. Here we used high-density EEG to investigate the ability of the human visual system to overcome such distortions, by measuring responses to images of naturalistic, novel, 3D objects. Our paradigm used Steady-State Visual Evoked Potentials (SSVEPs) to isolate brain activity specific to symmetry processing. We presented images of symmetrical and asymmetrical objects under two viewing conditions: One that produced symmetries in the image plane, and another where objects were rotated such that symmetries would be distorted in the resulting image. In each stimulus cycle, an asymmetrical object image was shown followed by a second image, either symmetrical or, in a control condition, another asymmetrical image. Image pairs for each cycle were selected so activity in a deep convolutional neural network trained to classify object categories (VGG16; Simonyan and Zisserman, 2015) was similarly matched between all image-pairs. We showed 10 cycles per trial, at a stimulation frequency of 1 Hz. In a follow-up experiment, the shading cues to 3D shape were removed from the images. In the resulting 2D object silhouettes, perspective-distorted symmetry cannot be detected. We analyzed the SSVEP data in three electrode regions-of-interest over occipital cortex and left and right temporal cortex, defined based on independent data. The results showed that during passive viewing, perspective-distorted symmetry can elicit measurable symmetry-specific SSVEPs, but compared to image-plane symmetry they are weaker, more anterior and possibly more right-lateralized, consistent with responses in higher-level visual cortex. The 2D stimuli elicited similar responses to 3D for image-plane symmetry, but as expected, produced no symmetry responses for perspective-distorted symmetry. Future work will determine how task manipulations may influence responses to these stimuli.

This work was supported by the Vision Science to Applications (VISTA) program funded by the Canada First Research Excellence Fund (CFREF, 2016–2023) and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada awarded to PJK.

56.346 FROM CURVATURE TO CONTOUR: HIERARCHICAL REPRESENTATIONS OF CONTOUR SHAPES IN TERMS OF CONSTANT CURVATURE SEGMENTS

Kevin Lande[1] (lande@yorku.ca), Doug Addleman[2], Denis Buehler[3], Cameron Pham[4], Silvia Rufus[4], Nicholas Baker[4]; [1]York University, [2]Gonzaga University, [3]Ecole Normale Supérieure, [4]Loyola University

Introduction: We can discriminate a stunning variety of shapes. How does the visual system encode these different shapes? We investigated the hypothesis that the visual system forms hierarchically structured representations of contour shapes, based on primitives that represent segments of constant curvature (CC). This hypothesis implies that (i) in representing a contour, encoding of CC segments is obligatory; (ii) variation in CC segments will induce perceptible differences between contours; and (iii) CC segments can be organized perceptually into higher-order parts. Experiments: In Experiment 1, we displayed contours made from two curvatures and two colors. The transition point for color was near to, but offset from, the transition point for curvature. We then presented the contour again, sometimes shifting the color transition point. When asked whether the coloring was different, participants were much less sensitive to shifts that aligned the color transition with the task-irrelevant curvature transition than to equivalent shifts that increased misalignment. In Experiment 2, we compared participants' ability to discriminate between a contour fragment made of multiple curvatures and one made of one curvature. Sensitivity was considerably higher when multi-curvature contours were predicted to be represented with multiple CC segments than with a single CC segment. In Experiment 3, we tested a hypothesis that CC segments with the same curvature polarity are represented as higher-order "parts" of a contour. Following Palmer (1977), we tested participants' ability to say whether a contour fragment was part of a shape. Participants were significantly faster when the fragment was from a polarity-matched contour region. Performance using polarity-matched fragments was comparable to performance using segments between curvature minima. Conclusion: These experiments suggest that CC segments are obligatorily encoded in contour representation (Exp.1), that contour discrimination depends on encoded CC segments (Exp.2), and that CC segments organize together into higher-order units (Exp.3).

The project received funding through a research award from Duke University's Summer Seminar in Philosophy and Neuroscience.

56.347 THE EFFECT OF RETINAL SIZES ON ENSEMBLE SIZE JUDGMENTS OF OBJECTS IN DIFFERENT DEPTH PLANES

Katrín Fjóla Aspelund[1], Kaleb Ducharme[1], Ömer Dağlar Tanrıkulu[1]; [1]University of New Hampshire

The visual system has been shown to be efficient at summarizing scenes – for instance, capturing the average hue of tree leaves or facial expressions in a crowd, a phenomenon known as ensemble perception. Previous studies have shown that our visual system estimates the average size of multiple objects based on their perceived sizes, rather than their retinal sizes. This means that when representing size ensembles, the averaging is performed after the visual system rescales the retinal size of the objects by their corresponding distances. However, previous studies did not investigate if (and to what extent) retinal size influences this averaging process. In this study, we investigated whether the retinal size of objects at different distances influences average size responses. Observers were presented with multiple spheres rendered across three distinct depth planes within a Virtual Reality environment. They were tasked with reporting the mean size by adjusting a probe sphere positioned in the middle plane. In Experiment 1, the spheres were displayed in an empty setting with a black background. In Experiment 2, perspective cues were added by presenting the spheres within a simulated room with textured walls. Each sphere's size was correlated with its distance, such that the retinal size of each sphere was equal across the different depth planes. In both experiments, object sizes were overestimated in the near plane, but underestimated in the far plane. In contrast to prior research, our findings suggest that the average size is not solely determined by the perceived size of objects, but rather also influenced by their retinal size. These results underscore the complexity of the size-averaging mechanism, and highlight a more intricate interaction between depth and size processing in building ensemble representations of object sizes.

56.348 UNDERESTIMATION OF NUMEROSITY WITH OCCLUSION IS DENSITY DEPENDENT

Hui Men[1], Alexander C. Schütz[1]; [1]University of Marburg

Estimating the number of objects in a crowd can be complicated by the fact that some of the objects might be hidden behind occluders. It has been shown that numerosity is underestimated in partially-occluded displays when the visible objects are sparsely distributed (Men et al., 2023). Since there is a transition from a numerosity to a texture-density mechanism at larger numbers and higher densities (Anobile et al., 2014), it might be that the underestimation of numerosity in partially-occluded displays decreases at larger numbers and densities. Here, we investigated how the number and density of visible objects affects the estimation of numerosity in a partially occluded scene. Two gray boards with different numbers of randomly arranged pieces were displayed simultaneously. Participants had to first discriminate which board contained more pieces, and subsequently report their confidence in discriminating the numerosity. Both of the boards could be fully visible or partially occluded, or one of the boards fully visible and the other partially occluded. We manipulated the size of the board and the number of visible pieces to achieve three combinations of numerosity and density: Large numerosity and high density, small numerosity and low density and small numerosity and high density. Results showed that the underestimation of numerosity was smaller for displays of high densities than for displays of low density. Since these effects were present in numerosity and confidence judgments, they represent genuine perceptual effects rather than response biases. Despite this underestimation, participants were equally confident in their judgments with or without occlusion. Weber fractions were smaller in displays with high density, indicating that participants indeed might have relied on a texture-density mechanism at high densities. Our results show that numerosity is underestimated in partially-occluded scenes, and that the strength of such underestimation depends on the number and density of visible objects.

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101001250).

56.349 PERCEPTUAL GESTALT AT WAR

Zhen Li[1] (psyzjuzl@gmail.com), Ning Tang[1], Siyi Gong[2], Jifan Zhou[1], Mowei Shen[1], Tao Gao[2]; [1]Zhejiang University, [2]UCLA

The tragedy of human warfare is that settlements are difficult to reach even when both sides are suffering. Economic theory has posited that mutually salient rallying points, even when completely irrelevant to human conflicts, can facilitate peace settlement - a speculation that has never been examined with experiments. Here we investigate whether perceptual gestalts, a universal perceptual phenomenon, can function as rallying points influencing war strategies. We devised a 2D multiplayer strategic war game, where participants could freely deploy their troops. Unlike most zero-summed war games, players' task was not annihilating the opponent but accumulating maximum wealth. Just like the real world, while expansion could bring wealth, a war of attrition could be mutually destructive as 1) troop logistics increased with the extension of supply lines, and 2) casualties could be high even for the winner. To examine the role of perceptual gestalt at war, for one group of participants ('Grouping condition') the battlefield grids were perceptually organized into different groups based on war-irrelevant color. Compared to the 'Non-Grouping' condition where all grids shared the same color, task-irrelevant perceptual grouping influenced war strategies in several ways: Firstly, it reduced the number and intensity of wars. Secondly, it constrained strategic decisions: Despite having freedom to develop troops in any direction, participants nevertheless overwhelmingly fought along the direction of perceptual grouping. Thirdly, it altered the dynamics of war: While wars resembled 'blitz' in the 'Non-grouping' condition, with fast troop movements causing high annihilation rate (42%), perceptual grouping led to 'trench warfare' with lower annihilation rate (20%). Moreover, the hierarchical structure of perceptual grouping biased how three players formed alliances, with powers residing in color groups closer in the perceptual hierarchy more frequently allying. Collectively, these findings suggest perceptual gestalt can serve as a mutually acknowledged rallying point constraining human warfare decisions.

56.350 UNDERSTANDING THE TIME COURSE AND SPATIAL BIASES OF NATURAL SCENE SEGMENTATION

Ruben Coen-Cagli[1] (ruben.coen-cagli@einsteinmed.org), Jonathan Vacher[2], Dennis Cregin[1], Tringa Lecaj[1], Sophie Molholm[1], Pascal Mamassian[3]; [1]Albert Einstein College of Medicine, [2]Université Paris Cité, [3]Ecole Normale Supérieure

Image segmentation is central to visual function, yet human's ability to cut natural scenes into individual objects or segments remains largely unexplored because it is notoriously difficult to study experimentally. We present a new experimental paradigm that overcomes this barrier. We flash two dots briefly, before and during the presentation of a natural image, and the observers report whether they perceive that the image regions near the two dots belong to the same or different segments. By repeatedly sampling multiple locations on the image, we then reconstruct a perceptual probabilistic segmentation map, namely the probabilities that each pixel belongs to any segment. Leveraging this method, we addressed two fundamental questions. First, strong spatial biases (a preference to group together items that are close in visual space) have been revealed using synthetic stimuli, but are they part of natural vision? Our data–unsurprisingly, but for the first time–directly shows spatial biases in human perceptual segmentation of natural images. The probability that participants reported two regions as grouped together, decreased with the distance between the two dots, regardless of whether the two regions belonged to the same or different segments in the perceptual segmentation maps. Second, is perceptual segmentation of natural images fast and parallel across the visual field, or a serial, time-consuming process? A prominent theory proposes that judging if two regions are grouped requires a gradual spread of attention between those regions, thus taking longer at larger distances (e.g. Jeurissen et al 2016 eLife). Surprisingly, whereas reaction times in our task increased with distance when the two regions were judged to be in the same segment, consistent with the theory, reaction times decreased with distance otherwise. We show that a dynamic Bayesian ideal observer model unifies these findings, through the interaction between spatial biases and evidence accumulation.

This research was supported by a NIH-ANR CRCNS grant (NIH-EY031166 to R.C.C. and ANR-19-NEUC-0003 to P.M.) and NIH grant

P50 HD105352 (Support for the Rose F. Kennedy IDD Research Center, S.M.).

56.351 PUSHING THE BALL VS. PULLING THE RUBBER BAND: REVERSAL OF CAUSAL AGENT-PATIENT RELATIONSHIP BETWEEN TWO MOVING OBJECTS INDUCED BY A SPEED CHANGE AT THE MOMENT OF SEPARATION

Jimin Ju[1] (psy_jiminju@ewhain.net), Sung-ho Kim; [1]Ewha Womans University

Michottian studies on causal perception have primarily investigated spatiotemporal factors influencing causal relationship between two rigid bodies in billiard-ball-like collisions. Here, we present a novel bistable causal perception phenomenon involving a 'separation' event: A thick horizontal line rooted in a stationary rectangle elongates to the right, as a disc connected to its right end simultaneously moves to the right at an identical speed; halfway through this motion, the line and the disc separate; subsequently, the line returns to its initial length, while the disc continues moving to the right and then comes to a halt. This sequence could be perceived as either "a stick pushed a ball to move" or "a ball pulled a rubber band, stretching it until snapped." In Experiments 1-3, we investigated how the causal interpretation of the line-disc relationship is disambiguated by their kinematic properties before and after the separation, using both free report and forced choice tasks. The results showed that the "pulling" interpretation was dominant over the "pushing" percept when their motion decelerated before separation and when they moved rapidly after separation. In Experiment 4, we employed an indirect, performance-based, pause-detection task, in which to detect whether line motion momentarily paused immediately after the separation. The results indicated higher sensitivity in detecting the pause when the objects' pre-separation speed decreased (i.e., in the "a ball pulling a rubber band" percept-dominant condition), suggesting that the presence of the pause is incompatible with the perceived elasticity of the line returning to its relaxed state. This study demonstrates that subtle changes in low-level kinematic features alone can radically alter perceptions of the agent-patient relationship of moving objects and their material properties (elastic vs. rigid) simultaneously. It further suggests that the visual system is reliably tuned to the physical regularities of various dynamic-kinematic mappings beyond collision events.

56.352 DO VISUAL OBJECTS LOSE THEIR INDIVIDUALITY DUE TO THE PERCEPTION OF COLLECTIVE GOALS?: EVIDENCE FROM NUMERICAL UNDERESTIMATION IN THE WOLFPACK EFFECT

Dawei Bai[1] (dawei.bai@yale.edu), Brian Scholl[1]; [1]Yale University

The perception of multiple shapes depends on how they are arranged: you might see distinct objects, or you might see a single *group*. What cues drive such grouping? Classic answers involve simple physical relationships such as proximity or common motion. Here, in contrast, we ask whether perceptual groups can also be formed by *social* cues — as when the movement of simple shapes triggers an impression of goal-directed agency. To find out, we explored the flip-side of grouping: loss of individuality. Previous work has shown that when observers must quickly estimate the number of discs in a display, for example, estimates will be lower when pairs of discs are grouped by thin lines. Does social grouping due to shared goals lead to a similar loss of individuality? Observers viewed short (e.g. 1.5 s) animations of moving 'dart' shapes. Darts moved randomly, but rotated so as to continuously point at a common target (a moving disc). This display yields a vivid impression of collective goal-directed behavior (known as the Wolfpack effect): despite the random movement, the shapes appear to share the common goal of pursuing the target. In contrast, when darts are always oriented 90° away from the target (thus equating rotary motion correlation), such impressions are destroyed. Observers viewed variable numbers of darts, and simply estimated how many were present on each trial. The results were striking: numerical estimates were reliably lower for Wolfpack displays compared to equated 90°-rotation displays — and this was true for a range of different numerosities, and even for observers who reported not noticing the darts' related orientations. This suggests that shared goals can lead to perceptual grouping and a corresponding loss of individuality. In this way, social groups may be central not only to social cognition, but also to visual processing.

TUESDAY AFTERNOON POSTERS IN PAVILION

Tuesday, May 21, 2:45 – 6:45 pm, Pavilion

Action: Locomotor, flow, steering

56.401 WALKING IMPACTS VISION: SPATIAL AND TEMPORAL FREQUENCY DEPENDENT CHANGES TO VISION DURING LOCOMOTION

Marlon Blencowe[1] (marlon.blencowe@adelaide.edu.au), Steve Wiederman[1], Dominic Thewlis[1], Anna Ma-Wyatt[1]; [1]University of Adelaide

Contrast detection and direction discrimination of motion stimuli has been studied extensively using seated, head-fixed experiments. However, there is increasing evidence from research in non-human animals that activity, including locomotion, might alter processing of visual information. Locomotion dependent changes to visual perception in humans are currently not well understood, and findings have been mixed (Benjamin et al., 2018; Cao & Händel, 2018). To further investigate potential effects of locomotion, we employed a within-subject design, obtaining contrast detection thresholds from 14 participants. Participants completed a two-alternative forced-choice motion direction discrimination task while walking and sitting. Participants walked for 10-minute intervals on a treadmill at a self-selected speed (0.97-1.3 metres per second). The stimulus was a vertical drifting Gabor presented in central vision (σ=0.375 degrees). Four spatial frequencies (0.5, 2, 8 and 16 cycles per degree) and two temporal frequencies (2 and 10 Hz) were tested in combination. A linear mixed model was used to analyse results. We found small but significant spatial and temporal frequency dependent differences in contrast sensitivity between the walking and sitting conditions. While contrast sensitivity for motion direction discrimination was similar for walking and sitting at 2 Hz, a significant difference in sensitivity was found between the conditions at 10 Hz, an effect that was principally driven by differences in the 2 cycles per degree condition. Results suggest there are spatial and temporal frequency dependent changes to contrast sensitivity during locomotion. We provide evidence that similar changes to those observed in non-human animals may exist in humans, and discuss what may be driving these changes.

This research was funded by a Research Training Program Stipend from the University of Adelaide

56.402 WALKING WHILE PERFORMING VISUAL, AUDITORY AND CROSSMODAL TASKS PRODUCES OSCILLATIONS ENTRAINED TO THE GAIT CYCLE

David Alais[1] (david.alais@sydney.edu.au), Gabriel Clouston[1], Matthew Davidson[1]; [1]University of Sydney, Australia

Although we are inherently active beings, most knowledge of perception comes from experiments that restrict natural movement and thus ignore the perception-action loop. Recent evidence from our lab (Davidson, Nature Comms, 2023) has demonstrated that during walking – one of our most common everyday actions – performance on a simple visual detection task cycles through a good and bad phase with every step. Here we extend this line of work to test the influence of walking on more complex visual, auditory and crossmodal tasks. In a series of visual experiments, we observe oscillations of both sensitivity and criterion in a two-alternative forced-choice visual discrimination task, as well as oscillations in perception of (visual) numerosity. In auditory experiments, we demonstrate that performance on an auditory detection task is also entrained to the rhythm of the step-cycle, suggesting stride-cycle oscillations are a domain-general phenomenon. Finally, using an audio-visual temporal synchrony task, we observe that the width of the synchrony window broadens and narrows during the step-cycle. Together, these results demonstrate that gait-related oscillations in performance generalise from simple visual detection tasks to visual discrimination tasks, and to auditory and audio-visual tasks. In most observers, these modulations occur at 1 cycle per step although some participants exhibit a modulation at 2 cycles per step. These findings are clear examples of action modulating perception and show that perception cannot be fully understood without studying the perception-action loop. Studying perception while active provides conditions closer to those for which vision evolved and will likely lead to new insights not possible in static lab environments.

56.403 SPATIO-TEMPORAL COLLISION ENVELOPE IN VIRTUAL REALITY WALKING WITH COLLIDING PEDESTRIANS

Jonathan K. Doyon[1,2] (jkdoyon@gmail.com), Sujin Kim[1,2], Alex D. Hwang[1,2], Jae-Hyun Jung[1,2]; [1]Department of Ophthalmology, Harvard Medical School, [2]Schepens Eye Research Institute of Massachusetts Eye and Ear

Safe locomotion depends on maintaining a collision envelope (i.e., safety margin) that flexibly adapts to dynamic situations. We previously characterized dynamic collision envelopes spatially as the minimum egocentric radial distance maintained while walking with other colliding pedestrians. Spatial collision envelope sizes depended on relative walking speeds between the subject and the colliding pedestrian. From this finding, we hypothesized that collision envelopes may be driven by temporal distances (time-to-collision, TTC) rather than spatial distances. To this end, we used virtual reality (VR) walking scenarios for the Meta Quest 2 head-mounted display (HMD). Normal vision (NV, n=10) and homonymous hemianopia (HH, n=7) subjects physically walked with free gaze in a VR shopping mall presented in the HMD. Subjects naturally avoided (speed/path changes) a colliding pedestrian among 10 non-colliders. Head-on (farther distance, faster relative speed) or rear-end (closer distance, slower relative speed) colliders approached from initial bearing angles of 20°, 40°, and 60°. Collisions had an initial 6-second TTC, which linearly decreased to 0 if the subject failed to avoid. We computed instantaneous TTC (which fluctuates with gait, body volumes, and path changes) in the egocentric domain as the spatial distance between subject and collider divided by the projection of the instantaneous velocity toward the walking direction of the subject. We considered the minimum TTC across relative bearing angles to be the subject's spatio-temporal collision envelope as reaction to the collision. The radius (head-on=0.77 seconds ±0.27, read-end=0.79s ±0.54, p=0.84) and area (head-on=3.26s^2 ±1.20, rear-end=3.87s^2 ±1.79, p=0.26) of spatio-temporal collision envelopes were consistent regardless of distance and relative walking speeds. HH had more conservative spatio-temporal collision envelopes (0.86s ±0.48) than NV (0.63s ±0.23, p<0.01) during avoidance. This finding suggests that collision avoidance may be driven by the spatio-temporal collision envelope based on the minimum TTC that allows enough time (rather than distance) to react.

This work is supported by NIH R01 EY031777 and P30 EY003790

56.404 CAN COVERT AND EXPLICIT "LEADERS" STEER AND SPLIT REAL HUMAN CROWDS?

Kei Yoshida[1] (kei_yoshida@brown.edu), Hector Taylor[1], William H. Warren[1]; [1]Brown University

Collective motion in human crowds is a self-organizing phenomenon that emerges from local visual interactions between neighboring pedestrians. In our previous work, we developed a method of reconstructing spatially-embedded visual influence networks and quantifying local and global influence. We found that some positions in the crowd are more influential than others ("emergent leadership"), as are some participants ("individual leadership"). To test these influential positions, here we experimentally manipulate leadership and network topology in real crowds. In each session, a group of participants (N=16 to 22) was instructed to walk across a field together. One to four confederates, whose presence was either unknown (covert leaders) or visually specified (explicit leaders), turned midway through the trial. In each trial, they received instructions about their initial position (e.g., front or middle of crowd, left/right/center) and turn direction (right, left, no change) via smart watches, and when to initiate the turn via pagers. Head positions were recorded using a Mavic 3 drone with a top-down video camera (60 Hz). To analyze the video recordings, we developed a data processing and analysis pipeline. We trained a convolutional neural network model that uses a multiple-object tracking framework to detect the locations of participants and extract their trajectories (tracklets) and identities. We then used our previous network reconstruction methods to recover visual influence networks and analyze leadership and crowd dynamics. The results show that confederates exerted more influence than non-confederates, particularly when they occupied influential positions. Covert and explicit leaders altered the network topology. We are currently comparing the ability of confederates to steer and split the crowd with predictions from our multiagent simulations of our collective motion model. The results have potential applications to directing emergency evacuations.

Supported by NSF BCS-1849446, NIH 1S10OD025181

56.405 HOW MANY MOVING OBSTACLES DO WE RESPOND TO AT ONCE? A TEMPORAL THRESHOLD MODEL BEST ACCOUNTS FOR COLLISION AVOIDANCE IN A CROWD

Kyra Veprek[1], William Warren[1]; [1]Brown University

On our daily commutes, we seamlessly weave through crowds, avoiding potential collisions with multiple pedestrians. How do we prioritize which obstacles to avoid? It is possible that we respond to the nearest N obstacles (topological threshold), or all obstacles within a temporal range (visual threshold). We previously described a visual model in which the risk of collision is specified by an obstacle's change in bearing direction ($|\psi'|$), and the imminence of collision is specified by its optical expansion (θ'). Here we investigate the number of obstacles avoided by manipulating the visual threshold on ($\theta' \cdot |\psi'|$). In a VR experiment, participants avoided one, two, or three moving avatars (1.1m/s), which crossed their path (±112.5°) while walking toward a goal (11m). We compared models with four different thresholds, measuring error as the mean distance between model and human positions: (1) A visual threshold fit to multiple obstacles had the lowest error ($\theta' \cdot |\psi'|$ = 0.20 deg/s, M = 0.381m, Mdn = 0.267m). (2) A topological threshold for the single next obstacle had the next highest error ($\theta' \cdot |\psi'|$ = 0.03 deg/s, M = 0.422m, Mdn = 0.298m). (3) A topological threshold for the next two obstacles had even higher error ($\theta' \cdot |\psi'|$ = 0.03 deg/s, M = 0.463m, Mdn = 0.323m). (4) Our previous visual threshold fit to collisions with a single obstacle had the worst performance ($\theta' \cdot |\psi'|$ = 0.03 deg/s, M = 0.477m, Mdn = 0.331m). We then simulated previous data on a participant walking through a crowd of criss-crossing avatars (VSS2023) with the same thresholds, and found that the 0.20 deg/s threshold again had the lowest error (M = 0.625m, Mdn = 0.443m); on average, 1-2 obstacles were above threshold. We conclude that a visual threshold that limits the response to moving obstacles provides the most parsimonious model of human collision avoidance.

NIH 1S10OD025181, NIH 5R01EY029745

56.406 CONTROL THEORETICAL MODELS FOR VISUOMOTOR CONTROL EXPLAINS BRAIN ACTIVITY DURING NATURALISTIC DRIVING

Tianjiao Zhang[1], Christopher A Strong[1], Kaylene Stocking[1], Jingqi Li[1], Claire Tomlin[1], Jack L Gallant[1]; [1]UC Berkeley

Visuomotor control is crucial for successfully navigating through the world, as we must continuously adjust our actions to account for the behavior of other agents. Multiple brain regions, including the intraparietal sulcus (IPS), motor cortex, supplementary motor areas (SMA), and the prefrontal cortex (PFC), have been implicated in visuomotor control. However, the control algorithms implemented by these regions during interactions with other agents remain poorly understood. Here, we examined whether the brain may use algorithms similar to control theoretical models for car-following. We used fMRI to record brain activity from six participants performing a taxi-driver task in a large virtual world (110-180 minutes of data per participant). Virtual traffic required participants to constantly monitor other vehicles and adjust their own actions. We implemented three control theoretical car-following algorithms: the optimal velocity model (OVM) and intelligent driver model (IDM), two reactive dynamical systems models, and a model predictive control (MPC) model, a forward predictive model. In preliminary analysis in two participants, we tuned the parameters of the three control models to match the behavior of each participant, and used these tuned control models to create features for modeling brain activity. We used banded ridge regression (Nunez-Elizalde et al., 2019, Dupré la Tour et al., 2022) to estimate voxelwise encoding models for these control models along with 34 other feature spaces for the taxi-driver task. The MPC control model better matches the participants' behavior than IDM or OVM, and the MPC encoding model better predicts brain activity than the IDM and OVM encoding models. Well-predicted regions include parts of the PFC, IPS, SMA, and motor cortex. Encoding model weights reveal multiple timescales of predictive control in the cortex. These results suggest that the human brain may implement a forward predictive algorithm similar to MPC for optimal visuomotor control during driving.

This work is supported by funding from the NEI, ONR, Ford URP, and NSF GRFP

56.407 ANTICIPATORY CONTROL OF STEERING THROUGH MULTIPLE WAYPOINTS

AJ Jansen[1] (jansea@rpi.edu), Brett Fajen[1]; [1]Rensselaer Polytechnic Institute

Effective locomotion often requires the ability to navigate within complex environments at speed, moving smoothly through multiple waypoints while avoiding obstacles. If actors consider only one waypoint at a time, they may be forced to make jerky steering adjustments, collide with obstacles, or miss waypoints altogether. The ability to use information from beyond the most immediate waypoint could in principle allow actors to steer more smoothly and accurately. Recently, Jansen et al. (VSS 2023) and Powell et al. (submitted) found that subjects' heading direction upon reaching a waypoint was systematically related to the relative position of the subsequent waypoint. Such findings provide evidence that humans do in fact anticipate future waypoints. However, neither experiment was designed to reveal how steering trajectories while approaching a waypoint depend on the position and orientation of the next waypoint. Such data would provide the empirical basis for formulating possible control strategies for steering through multiple waypoints, which motivated the present study. Subjects performed a simulated drone-flying task along a straightaway in forest-like virtual environment. The simulation was viewed on a monitor while eye movements were tracked using a Pupil Core headset. On each trial, subjects used a game controller to steer through a series of three gates: two centered on the longitudinal axis and separated by a fixed distance, and the third at a distance, angle, and orientation that was manipulated between trials. After passing through Gate 1, subjects initially turned away from Gate 2 in the direction opposite Gate 3 before turning back, allowing for a smoother (albeit less direct) trajectory through the series of waypoints. Maximum lateral deviation between Gates 1 and 2 increased with the relative angle of Gate 3, but the effects of distance and orientation were weaker. The findings inform the development of models of steering through multiple waypoints.

NSF 2218220

Action: Clinical, neural

56.408 AGE-RELATED PRESERVATION OF PROPRIOCEPTION- AND VISION-GUIDED VIRTUAL HAND MOVEMENTS

Jose Reynoso[1] (jose_reynoso@urmc.rochester.edu), Emily L. Isenstein[1,2], Mariah Steele[1], Khai Du[1,2], Leonardo Benavides[1,2], Ania Busza[1,2], Brian Keane[1,2], Duje Tadin[1,2]; [1]University of Rochester, [2]University of Rochester, Medical Center

Our ability to sense self-movement is a proprioceptive ability that is often overlooked. Without proprioception, a mundane task like reaching for a water bottle can become challenging, especially when out of sight. With a goal of studying interactions between vision and proprioception, we recently developed a virtual reality (VR) approach that can quickly and precisely measure reaching behavior, and empirically isolate proprioceptive function (Isenstein et. al., 2022). In the present work, we aimed to elucidate the effects of age on reaching behavior and proprioception. We compared healthy older adults (ages 58-74; N = 30) to young adults (ages 18-28; N = 37) using the aforementioned VR reaching task. The task was to simply reach out forward to "touch" a virtual ball. Critically, the subject's hand was either visible (rendered in VR) or invisible; a manipulation used to isolate proprioception. Based on previous research, we hypothesized that older adults would make less accurate reaches, especially when proprioception was isolated. However, our results showed no difference in performance between age groups in either visual- or proprioception-guided reaching (main effect of age: $p = 0.058$; main effect of removing vision: $p < 0.001$; interaction: $p = 0.228$). Notably, the marginal effect of age reflected numerically better performance in older adults. This finding cannot be explained by group differences in speed/accuracy (similar reaching times; $p = 0.41$) or differences in data quality (split-half reliability > 0.97 and > 0.93 for older and younger adults). This finding reveals a preservation of upper limb proprioception in older adults. To further investigate what other types of experience may affect proprioception-guided reaching, our ongoing work is now focusing on schizophrenia. Schizophrenia has been associated with impairments in visual perception and motor function, along with accelerated brain aging, motivating our aim to determine their relative contribution to reaching behavior.

56.409 DEVELOPMENT OF GROSS AND FINE VISUO-MOTOR ABILITY: INSIGHTS FROM LATE-SIGHTED CHILDREN

Shlomit Ben-Ami[1,2] (shlomit@mit.edu), Roy Mukamel[2], Chetan Ralekar[1], Sharon Gilad-Gutnick[1], Hannah Kim[1], Dhun Verma[3], Mrinalini Yadav[3], Priti Gupta[3,4], Pragya Shah[3], Suma Ganesh[5], Abhinav Gandhi[6], Pawan Sinha[1]; [1]MIT, [2]Tel-Aviv University, Tel-Aviv, Israel, [3]The Project Prakash Center, Delhi, India, [4]Amarnath and Shashi Khosla School of Information Technology, Indian Institute of Technology, Delhi, India, [5]Department of Ophthalmology, Dr. Shroff's Charity Eye Hospital, Delhi, India, [6]Worcester Polytechnic Institute, Worcester, MA

Grasping an object relies on an integration of gross- and fine-visuo-motor abilities, combining large-scale movements for reaching and hand orienting with precise finger movements ensuring a secure grasp. The acquisition of this intricate ability is influenced by many factors, including sensory processing, motor skill, physiological maturation of neural pathways, and experiential learning. Untangling the individual impact of these factors on visuo-motor development proves challenging given their concurrent progression during typical child development. Here we report an attempt to isolate the impact of visual experience on gross and fine visuo-motor development. We investigated the effects of visual restoration on individuals born with bilateral dense cataracts. These patients had attained neural and motor maturity while remaining visually inexperienced until undergoing cataract-removal surgery in late childhood (as part of Project Prakash). Fourteen patients were tasked with visually guided two-finger pincer grasping of irregularly shaped flat objects (Goodale, 1994). Pre- and repeated post-operative assessments allowed us to track longitudinal changes in visuo-motor measures. Results reveal a developmental progression in gross visuo-motor ability, with all patients demonstrating an increased reach-to-grasp speed, matching the speed of age- and acuity-matched controls within one year after surgery. Most also exhibited enhanced hand orientation variability, indicating adaptation to diverse object orientations. However, challenges persisted in fine visuo-motor ability for most patients, with only three achieving enhanced reliability, and two of these reaching computational optimality in grasping efficiency within a year. Our findings suggest that gross visuo-motor ability can be acquired following late sight restoration, but fine visuo-motor proficiency may suffer permanent compromise, or at least require prolonged experience. Ongoing follow-up beyond one year aims to uncover potential later-stage improvements and explore individual differences in long-term improvement. Our results point to the nuanced interplay between visual experience and visuo-motor development, providing insights for tailored interventions for visuo-motor skill acquisition.

(1) NEI (NIH) grant R01 EY020517 to PS, (2) Global seed funding from the Broshy Brain and Cognitive Sciences Fund for MIT-Israel collaborative studies. (3) Minducate Science of Learning Research and Innovation Center, Tel Aviv University.

56.410 HOW MUCH VISION IMPAIRMENT DOES IT TAKE TO DECREASE PERFORMANCE IN FREESTYLE SWIMMING?

David Mann[1] (d.mann@vu.nl), Rianne Ravensbergen[1], Kai Krabben[1], Daniel Fortin Guichard[1,2]; [1]Vrije Universiteit Amsterdam, [2]University of Toronto

Paralympic sport provides tremendous opportunities for individuals with impairment, including vision impairment, but there is controversy about the fairness of competition. Particularly, it remains unclear how much vision impairment should be necessary to compete. Almost all sports presently require visual acuity of at least 1.0 logMAR to qualify (20/200 or 6/60), though this cut-off is based on the legal definition of low vision and there is no evidence to suggest that this is a level of acuity that decreases sport performance. Moreover, the cut-off is likely to differ by sport. Accordingly, each sport could be including individuals without a disadvantage in the sport, or excluding those who do. The aim of this study was to establish the level of visual acuity loss that decreases performance in freestyle swimming. Twenty-one national level swimmers without vision loss swam 100m freestyle races in each of four different levels of vision impairment simulated using plus lenses (plano, +4.00, +6.00, +8.00). Visual acuities ranged from -0.3 to +1.6 logMAR. ROC analysis was used to establish the level of visual acuity that first brought performance below what would be expected by normal variation. Results revealed that a cut-off of at least 1.1 logMAR provided the optimal classification of performance as below expected performance (82% sensitivity and 68% specificity). In general, performance was not impacted by lesser amounts of vision impairment. The findings suggest that the minimum level of impairment required to compete in swimming for athletes with vision impairment may need to change, because there may be athletes competing whose impairment does not decrease performance in the sport. The results are expected to contribute to a change in the classification rules to be used for swimming in the 2028 Los Angeles Paralympic Games.

This research was supported by a Classification Research Grant from the International Paralympic Committee

56.411 FUNCTIONAL CONNECTIVITY OF ATTENTION NETWORK RELATED TO INDIVIDUAL DIFFERENCES IN VISUAL AND PROPRIOCEPTIVE WEIGHTING

Kess Folco[1], Hu Cheng[1], Sharlene Newman[2], Art Pilacinski[3], Manasi Wali[1], Reshma Babu[1], Hannah Block[1]; [1]Indiana University Bloomington, [2]University of Alabama Tuscaloosa, [3]Ruhr University Bochum

Individual differences in the weighting of visual vs. proprioceptive cues during visuomotor reaching tasks are predictive of functional connectivity between visual and sensorimotor neural regions. When both visual and proprioceptive cues about target position are available (bimodal target), participants weight and integrate these signals to estimate target position. This weighting varies across individuals, with some relying more on vision and less on proprioception, and others vice versa. Neural excitability of the primary motor cortex, or M1, has been related to individual differences in visuo-proprioceptive recalibration, a process related to the individual weighting of proprioception and vision; and the primary somatosensory cortex, or S1, has been related to proprioceptive recalibration. Other preliminary work in our lab has shown increased neural synchrony between cerebellar and visual regions such as the fusiform gyrus and early visual cortex. Together, this evidence suggests that individual variation in unisensory visual and proprioceptive processes are related to variation in neural activity. Despite this evidence, the neural basis of these individual differences in visual and proprioceptive weighting remains unexplored. To measure neural activity related to individual differences in visual or proprioceptive weighting, we collected two resting state functional scans before and after participants performed a visuomotor reaching task. Using a seed-based network analysis, we found that the weight of vision versus proprioception was related to increased synchrony between the dorsal attention network seed in the right frontal eye field (contralateral to the left visual field where visual reach cues appeared) and M1, S1, and the superior parietal lobule, a region well known for its involvement in visual perception. This suggests the neural basis of individual biases for vision or proprioception may be regulated by attentional systems.

56.412 NEUROPHYSIOLOGICAL CROSS-TASK SIMILARITIES BETWEEN METACONTROL STATES

Xi Wang[1] (xi.wang@ukdd.de), Xianzhen Zhou[1], Christian Beste[1,2], Bernhard Hommel[2]; [1]Technische Universität Dresden, Germany, [2]Shandong Normal University, China

Research on cognitive control commonly follows the concept of “willpower”. Accordingly, good control is considered to consist in sticking to one’s goal, and excluding every possible distraction. Yet, it is this style of processing that allows people to perform well in changing situation and adapt behavior. This has, more recently, been referred to as metacontrol which is some abstract ability that may transfer between different tasks and situations. The current study examines the neural underpinnings of this abstract ability using EEG-multivariate pattern analysis and source localization. We (i) examine whether neurophysiological activity pattern in one type of tasks predicts neurophysiological activity pattern in the other type of tasks, and (ii) whether this prediction would be more accurate for task combinations that theoretically can be assumed to share the same metacontrol bias. We show that there are indeed similarities between neurophysiological patterns and functional neuroanatomical structures across different tasks that do not share any obvious characteristics. The similarity is very specific: Metacontrol has the strongest impact during selection processes, be they related to the stimulus or the response. Moreover, similarities only show up under conditions that can be assumed to imply comparable metacontrol states. This is strong evidence for the existence of metacontrol states that are more abstract than concrete task-representations as assumed in cognitive control theories. The source localization analysis suggests that neuroanatomical structures known to form the “multiple demand (MD)” system are associated with the detected dynamics. This provides a novel view on the functional relevance of the MD.

56.413 ORDER-DEPENDENT FUNCTIONAL BRAIN CONNECTIVITY IN A CUE-SEPARATION GRASP TASK

Gaelle Luabeya[1,2] (gaellenl@yorku.ca), Ada Le[1,3], Lina Musa[1,3], Amirhossein Ghaderi[1], Simona Monaco[5], Erez Freud[1,2,3,4], J. Douglas Crawford[1,2,3,4]; [1]Centre for Vision Research, York University, Canada, [2]Departments of Biology, York University, Canada, [3]Department of Psychology, York University, Canada, [4]Department of Kinesiology & Health Sciences, York University, Canada, [5]Center for Mind/Brain Sciences, University of Trento, Italy

Prehension involves location-dependent reach transport and orientation-dependent grasp components. To understand how the brain integrates object location and orientation for grasp, we studied how the order of transport / grasp cues influences whole brain functional connectivity. We collected BOLD signal data from 12 participants in an Event-related fMRI Experiment. Participants were instructed to reach and grasp a cube illuminated to the left or right of midline (Location Cue: L) and a verbal instruction to orient the hand for vertical or horizontal grasp (Orientation Cue: O). The order of these cues (LO vs. OL) varied randomly. fMRI data were analyzed separately based on three predictors: Delay 1 (between the two cues), Delay 2 (between the 2nd cue and go signal), and an Action Phase. Graph Theory Analysis was performed based on 200 regions of interest (nodes) at each phase. Preliminary analysis based on 3 participants: During Delay 1, nodes coalesced into three modules: 1) a central parietofrontal strip approximating primary somatomotor cortex, 2) two more anterior-posterior premotor / visuomotor parietofrontal regions, and 3) a ring of cortex skirting 1+2 but with no occipital/temporal involvement. Occipital involvement increased in Delay 2. Parietofrontal Modules 1 + 2 joined (reducing to two modules) after Delay 2 for LO and the action phase for OL, i.e., always after the location cue. The Global Clustering Coefficient is always reduced in the action phase. We conclude the order of L-O cues influences modularity, such that location information produces more parietofrontal 'binding', presumably in preparation for transport.

The research was funded by the Canadian Institutes for Health Research and Canada Research Chair

56.414 VISUAL CORTEX ENCODES SPATIALLY SPECIFIC REWARD INFORMATION DURING CLOSED-LOOP NATURALISTIC INTERACTION

Royoung Kim[1,2] (royoungkim@skku.edu), Jiwoong Park[1,2], Kendrick Kay[3], Won Mok Shim[1,2]; [1]Center for Neuroscience Imaging Research (CNIR), Institute for Basic Science (IBS), Suwon, Korea, Republic of., [2]Sungkyunkwan University (SKKU), Suwon, Korea, Republic of., [3]Center for Magnetic Resonance Research (CMRR), University of Minnesota, Minneapolis, MN, USA

Our visual system actively gathers information from the environment to facilitate actions aligned with behavioral goals, and reward information plays a significant role in linking sensory inputs to optimal actions. While previous animal studies have demonstrated visual cortex's sensitivity to rewards, the mechanism through which potential rewards modulate visual representations during goal-directed actions in dynamic naturalistic settings remains poorly understood, particularly in humans. In this study, we introduced an innovative naturalistic 3D interactive paradigm and examined the impact of rewards on visual representations. As part of the 7T Naturalistic Perception, Action, and Cognition (NatPAC) dataset, we collected human fMRI and high-precision eye-tracking data during a "shepherding" task in which participants continuously formulated plans and made actions within a Minecraft-based environment to herd as many sheep (i.e. potential rewards) as possible to a specified location while avoiding risks, such as puddles and a fox. We then used high-precision gaze data collected during the shepherding task, in conjunction with results from conventional population receptive field (pRF) mapping, to generate predicted BOLD signals based on gaze-centered visual inputs. Subtracting these predictions from actual BOLD signals removed the retinotopic visual stimulation effect, and the remaining residuals were modeled with reward-related regressors. To examine the spatial specificity of rewards, we segmented the visual field into eight radial bins and explained their variance with the reward-related regressors. The results showed significant spatial specificity for reward loss in V1, V2, and V3. In addition, correlation analysis revealed that higher reward-based spatial specificity in these regions correlated with more successful loss-avoidant behaviors. In summary, we discovered spatially specific reward signals in the visual cortex, which facilitate behaviors aimed at maximizing rewards. This study illustrates the intricate interplay of vision with action and cognition in naturalistic settings, which is rarely studied in conventional laboratory paradigms.

This work was supported by the IBS-R015-D1, NRF-2019M3E5D2A01060299, NRF-2019R1A2C1085566, HI19C1328, and the Fourth Stage of Brain Korea 21 Project in Department of Intelligent Precision Healthcare, Sungkyunkwan University (SKKU).

TUESDAY, MAY 21, 2:45 – 6:45 PM, PAVILION

Eye Movements: Clinical

56.415 ENHANCING EYE MOVEMENT CONTROL IN VIRTUAL REALITY: OCULAR BIOFEEDBACK TRAINING AND ITS POTENTIAL IMPLICATIONS FOR ATTENTION DEFICITS

Tchiya Ben Joseph[1,2], Alexander Rivkind[1], Ofer Karp[1], Ehud Ahissar[1], Yair Zvilchovsky Zvilchovsky[2], Yoram Bonneh[2]; [1]weizmann institute of science, [2]Bar Ilan University

The visual system relies on intricate closed-loop networks, encompassing both conscious and semi-conscious processes. Our research investigates the hypothesis that training semi-conscious eye movement processes in individuals with attention deficits can elicit adaptive changes, promoting novel scanning patterns and potentially ameliorating attention-related deficits. To achieve this, we developed a dynamic virtual reality (VR) setup, incorporating ocular biofeedback training, which responds to subject eye movements. Presented here are preliminary results from healthy individuals engaged in a VR game demanding intense motor activity and focused attention on central stimuli. As a "manipulation assistant" we provided two tools for accurate center-focused viewing: blurring the virtual environment and introducing center-oriented music. The results demonstrate significant improvements in eye movement after manipulation, including a reduction in large saccade frequency (>6 degrees) and differences in gaze dispersion. Furthermore, the transfer of saccade inhibition ability during the retest, without manipulation, holds promising implications for therapeutic interventions targeting attention deficits and visual impairments. These findings suggest the potential efficacy of ocular biofeedback training in enhancing eye movement control and addressing related deficits within this population.

56.416 HOW CONTEXTUAL INFORMATION MODULATES EYE MOVEMENTS DURING NATURAL SEQUENTIAL BEHAVIOR IN A DYNAMIC SCENE.

Antonella Pomè[1], Eckart Zimmermann[1]; [1]Heinrich Heine University Düsseldorf

Our motor system has a remarkable ability to generate precise motor action and the flexibility to adjust in changes in the environment. In conditions where inflexibility in the motor domain has been reported, such as in Autism Spectrum Disorders, motor failures and sensory overload are exacerbating. Here we used a naturalistic interception task to investigate the interplay between autistic traits and the use of contextual information during natural sequential behavior in a dynamic scene. Participants engaged in a modified version of the 'Pong' video game, where they were asked to maintain a continuously moving ball within the game field by striking it with the paddle to prevent it from traversing beyond bounds. By manipulating the speed of the ball (fast vs slow) trial by trial, we demonstrated that participants scored low for autism were able chose the most efficient eye movement strategy, according to the ball physics. However, participants reporting higher autistic symptoms showed an inflexibility in accurately switching eye-movement strategy according to contextual information provided: on one hand, accuracy in adjust prebounce predictions of the ball's post-bounce trajectory were negatively correlated with the autism symptomatology; on the other hand, more frequent pursuit eye movements for the moving object, irrespective of the ball speed, were linked to more pronounced autistic symptoms, which have been interpreted as indices of gaming appraisal. Finally, pupil diameter, a remote indicator of load that is strongly associated to subjective gaming experience, varied with the autistic score of our participants: when predictive saccades were elicited, the control group showed larger pupil dilations compared to the high AQ group, presumably support a higher performance in the game. Overall, our results support atypical sensorymotor inflexibility in high autistic symptomatology that may be influenced by differential use of relevant contextual information in motion prediction.

Marie Skłodowska-Curie (Grant Agreement Number 101029574–APPROVE) to AP; DFG(Grant Agreement Number ZI 1456/6-1) and ERC(Grant Agreement Number 757184–moreSense) to EZ

56.417 DECODING READING CHALLENGES: EYE MOVEMENT PATTERNS IN ITALIAN-SPEAKING POOR READERS

Angela Pasqualotto[1], Francesco Pavani[2], Elvira Fontana[2], Paola Venuti[2], Michele Scaltritti[2]; [1]University of Geneva, [2]University of Trento

Reading is a multifaceted cognitive process involving linguistic abilities, as well as attentional control and eye movements. Poor readers often exhibit deficits in dynamic visual attention (including temporal processing, attention distribution, and sluggish attentional shifting) with an impact on eye movements during reading tasks. In orthographically regular languages like Italian, reading delays in poor readers are manifested in speed rather than accuracy, especially in text reading compared to single words or non-words. This study investigates eye movement patterns in 31 Italian-speaking poor readers (12 females; aged 7 to 14 years) through three carefully designed tasks: two text reading assessments and one pseudo-text reading task — to simulate real-world reading scenarios while accounting for lexicality. Importantly, the stimuli were drawn from standardized and validated assessment batteries, thus ensuring an appropriate control of linguistic parameters that can significantly affect eye movements, such as, such as text difficulty, syntax, word length, and word frequency, can significantly influence eye movements. Eye movements during silent reading revealed distinct patterns among poor readers, characterized by frequent saccades of small amplitude, prolonged fixations, and a high number of leftward saccades. To comprehensively contextualize these observed patterns, participants underwent standardized cognitive and reading tasks, enabling a thorough examination of correlations between eye movement behaviors, reading-related skills (such as rapid automatized naming and phonological awareness), fluid intelligence and age. Notably, the study found the classic lexicality effect (i.e., word advantage over non-words) on saccadic amplitude and fixation duration and count. Additionally, we investigated the impact of linguistic factors on reading, providing nuanced insights into challenges specific to the Italian language for readers with difficulties. These findings have implications for theoretical models of reading and practical interventions targeting reading challenges in children who speak orthographically transparent languages.

This research received funding from the PROMOBILIA FOUNDATION - Sweden and the NCCR Evolving Language, Swiss National Science Foundation Agreement #51NF40_180888.

56.418 MICROSACCADES AND OCULAR DRIFT IN OPHTHALMIC AND NEUROLOGIC DISEASE

Ola Abozid[1], Simrat K Renu[1], Elana Safonova[1], Josey Spiers[1], Evany Guerras-Solares[1], Robert G Alexander[1]; [1]New York Institute of Technology

Clinical disorders can have a significant impact on the quality of our vision, resulting in detectable abnormalities in eye movement patterns and downstream effects on perception. These abnormalities often include impairments in fixational movements that are often not consciously perceived by the patient, but which can be objectively measured via eye tracking. Neurologic and ophthalmic disease can produce specific fixational eye movement patterns with distinct characteristics. As a result, objectively assessing these small eye movements can aid in our understanding of pathologies that impair fixation and can provide insight into the nature and extent of visual impairments. These objective characterizations can also provide a means of early and differential diagnostics, as well as a means of evaluating ongoing treatment by quantifying progression and the response to medical intervention (in terms of the normalization of fixational eye movement dynamics). Methods—We conducted an exhaustive literature search of articles describing efforts to characterize fixational eye movement dynamics in common neurologic and ophthalmological conditions. We selected articles through extensive key word searches and searches of the references of retrieved articles. Results—Relatively few studies address fixational eye movement impairments in patient populations. However, this topic is a growing area of inquiry. We will present an overview summary of recent findings, with a particular focus on studies that have been published in the last five years. Conclusion—Recent discoveries point to some key research areas that may facilitate the translation of fixational eye movement measures into clinical practice. We discuss these discoveries and their implications. We also call for interested researchers to join the International Society for Clinical Eye Tracking (ISCET), which is a recently established community effort to provide guidance and protocols for conducting and analyzing eye tracking tasks in clinical settings.

56.419 SACCADE PROFILES ACROSS TASKS AFTER CHILDHOOD HEMISPHERECTOMY

Maria Z. Chroneos[1,2] (mzchrone@andrew.cmu.edu), J. Patrick Mayo[2], Marlene Behrmann[1,2]; [1]Carnegie Mellon University, [2]University of Pittsburgh

Saccades and smooth pursuit are supported by overlapping, bilateral brain networks including the brainstem and cortex (McDowell et al., 2008, Coiner et al., 2019, Sharpe, 2008, Lencer et al., 2008). Previous studies of patients, following lesions or hemispherectomies, reported both contralesional, as well as task-dependent ipsilesional, saccade deficits, and ipsilesional pursuit deficits (Troost et al., 1972, Herter et al., 2007, Sharpe et al., 1979, Morrow et al., 1993 & 1995, Thurston et al., 1988). Systematic studies of saccade function in hemispherectomy patients across task contexts are relatively lacking. We previously quantified the 'catch-up' saccade main sequence during sinusoidal pursuit in childhood hemispherectomy patients and found increased eye speeds relative to controls, especially ipsilesionally. Here, we compare saccade profiles during sinusoidal pursuit to profiles during visually-guided saccades in an overlapping set of childhood hemispherectomy patients (n = 12; surgery age: < 1 month-10 years, test age: 5-32 years) and similarly-aged controls (n = 14, test age: 6-32 years). We recorded eye movements using an EyeLink 1000 Plus while participants performed visually-guided saccades, with six possible directions containing a horizontal component (0°, 45°, 135°, 180°, 225°, 315°). We compared the latency, amplitude, and peak speed of the first saccade on each trial, and examined the main sequence, in patients and controls. We found that patient saccades had significantly longer latencies, shorter amplitudes, and slower speeds than controls, especially toward their contralesional blind field ($p < 0.05$). The main sequence slope was significantly reduced for patients in both visual fields relative to controls, indicating slower speeds across amplitudes (non-overlapping confidence intervals). This result contrasts with the increased main sequence slopes seen for catch-up saccades in similar patients relative to controls during pursuit. Future analyses will compare saccade and pursuit performance within-participants to elucidate how disrupted oculomotor networks affect saccade function across task contexts.

NIH grant to MB (R01EY027018); MB recognizes support from P30 CORE award EY08098 from NEI, NIH, and unrestricted supporting funds from The Research to Prevent Blindness Inc, NY, and the Eye & Ear Foundation of Pittsburgh. MZC recognizes support from the Fondazione Dompé's Rita Levi-Montalcini award

56.420 DIFFERENCES IN SMOOTH PURSUIT CHARACTERISTICS IN DIFFERENT TYPES OF STRABISMUS.

ELENA SANZ[1] (elena.sanzsanz@estudiante.uam.es), M. PILAR AIVAR[1], JAIME TEJEDOR[1]; [1]Universidad Autonoma de Madrid

Smooth pursuit is a type of eye movement that allows us to maintain a moving target near the fovea. Although the main characteristics of smooth pursuit are well described, even for young people, not many studies have analyzed performance in the case of strabismic patients. Furthermore, it is not clear if there are differences between the different types of strabismus. In this study, we analyzed smooth pursuit eye movements of young strabismic patients suffering from intermittent exotropia or accommodative esotropia, comparing performance with that of controls. Participants were 54 children, of ages between 4 and 13 years old (21 cases of accommodative esotropia, 17 cases of intermittent exotropia and 16 controls). Eye movements were recorded while performing different tasks. To produce smooth pursuit a single image was presented at the center of the screen, which jumped to a new position 14º to the right or to the left and immediately started to move horizontally across the screen with a sinusoidal movement, with a velocity of 11º/s. Smooth pursuit latency and gain, as well as the number of saccades during pursuit, were analyzed. Smooth pursuit showed different characteristics in the three groups, although variability was very high. Gain was significantly lower in intermittent exotropia, compared to the control group. In contrast, for smooth pursuit latency only patients with esotropia differed from controls. Accomodative esotropia patients also showed a significantly higher number of saccades during pursuit, and of a larger amplitude. Main differences with the other groups were also found in anticipatory saccades. When all variables were analyzed according to age, we found that each group progressed differently during development. These results show that smooth pursuit is affected by strabismus, but performance differs between clinical subtypes. More studies are needed to understand these developmental effects and evaluate clinical treatment outcomes.

56.421 FINE SPATIAL VISION IS OPTIMALLY ADAPTED TO THE ABNORMAL FIXATIONAL EYE MOVEMENTS OF PEOPLE WITH AMBLYOPIA

Susana T. L. Chung[1] (s.chung@berkeley.edu), Sunwoo Kwon[1], Dennis M. Levi[1]; [1]University of California, Berkeley

Our eyes are always moving even when we attempt to maintain steady fixation on a visual object. The retinal image motion associated with normal fixational eye movements (FEMs) has been shown to be important, yet not optimal, for seeing fine spatial details (Agaoglu et al., 2018). Because FEMs are more unsteady and are associated with larger amplitudes and higher velocities in eyes with amblyopia, we asked whether the naturally occurring FEMs are optimal for seeing fine spatial details in amblyopic eyes. Using a tracking scanning laser ophthalmoscope that is capable of delivering visual stimulus at precise retinal locations, we asked five observers with amblyopia to discriminate the orientation (±45°) of a sinusoidal grating presented at the fovea, for a spatial frequency 2.5× lower than the equivalent grating acuity of the eye. On each trial, the grating was presented at one of five stabilization gains (ratios of stimulus to eye velocity): -1, 0, 0.5, 1, 2. The contrast of the grating was adjusted such that observer's performance was ~80% correct under the natural viewing condition (gain=0), for the fellow and amblyopic eyes separately. Eye and stimulus positions were extracted from recorded videos using a cross-correlation method. For all observers, task performance was the highest at gain=0 (no stabilization) and dropped at other gains (e.g. performance accuracies for gains of 0, -1 and 2 averaged 78.9%, 49.8% and 49.8%, respectively, in the fellow eyes, and 82.1%, 55.1% and 55.7%, respectively, in the amblyopic eyes). Contrary to our previous finding that in the normal fovea, the best performance was obtained at a gain of 0.43, here we found that amblyopic observers' performance was the best with no stabilization, in both the fellow and amblyopic eyes, suggesting that fine spatial vision is optimally adapted to naturally occurring (abnormal) FEMs in the presence of amblyopia.

Supported by NIH Grants EY030609 (DML) and EY 030253 (STLC) and a postdoc fellowship from CIVO (SK).

56.422 REDUCED VISUAL ACUITY DUE TO DEFOCUS CANNOT FULLY ACCOUNT FOR THE ABNORMAL FIXATIONAL EYE MOVEMENTS OF PERSONS WITH AMBLYOPIA

Sunwoo Kwon[1,2,4,] (kwsunwoo@berkeley.edu), Julien Belen[1], Jessica Lien[1], Artashes Yeritsyan[1], Nam Do[1], Dennis Levi[1,2,3]; [1]Herbert Wertheim School of Optometry and Vision Science, UC Berkeley, [2]Center for Innovation in Vision and Optics, UC Berkeley, [3]Helen Wills Neuroscience Institute, UC Berkeley, [4]Exponent

Amblyopia is associated with impaired visual acuity, abnormal processes in primary visual cortex (V1), and reduced coordination of

activity in binocular neurons. Recent studies have demonstrated that these abnormalities extend beyond V1 with potential impacts on higher order processing and in oculomotor areas that affect visually-guided behaviors. While large fixational instabilities found in amblyopic eyes are highly correlated with worsening acuity in these patients, no such correlation was found in a neurotypical population (Raveendran et al., 2019). It is unclear whether the poor acuity causes fixational instability or vice versa or if the two are correlated as a byproduct of abnormal processes in V1. Therefore, we investigated whether visual acuity is truly a limiting factor in oculomotor and visually-guided behavior. We simulated reduced visual acuity observed in the amblyopic eye in the non-dominant eyes (NDEs) of corrected-to-normal neurotypical observers by applying convex lenses to induce retinal-defocus while the observers viewed naturalistic stimuli – a filtered version of "Where's Waldo" – while performing a visual search task. We tested a set of logMAR acuity differences (0.2, 0.4, 0.6, and 0.8) as observed in amblyopic patients and compared the oculomotor behavior of the normal eye with lens-induced acuity loss in the following conditions: monocular NDE and binocular viewing. We observed lens induced acuity loss had a significant impact on the observer's drift magnitude and fixational instabilities compared to the no-blur condition in the NDE; however, further reductions in visual acuity did not have any significant impact. Binocular viewing significantly improved fixational instability and reduced reaction time when compared with monocular viewing in both amblyopes and neurotypicals; this effect persisted as a function of retinal-defocus induced reduction in visual acuity. Our results demonstrate that the reduced visual acuity of patients with amblyopia cannot be fully accounted for by abnormal fixational eye movements.

Center for Innovation in Vision and Optics (CIVO) Fellowship, NEI R21

56.423 THE EFFECT OF IMAGE SIZE AND DEFOCUS ON CHILDREN'S REFLEX VERGENCE EYE MOVEMENTS TO NATURAL IMAGES

Bhagya L Marella[1], Clara Mestre[2], Stephanie A Biehn[1], T Rowan Candy[1]; [1]School of Optometry, Indiana University, Bloomington, IN, USA, [2]Center for Sensors, Instruments and Systems Development, Universitat Politècnica de Catalunya, Spain

Previously, we have demonstrated anisometropia disrupting oculomotor behavior of 6-8 year-olds viewing small 2deg targets. The purpose of this study was to determine the effect of unilateral and bilateral defocus on vergence behavior using larger natural images. Vergence responses to changes in disparity in a series of 16-degree radius natural images (SYNS database) were recorded using an Eyelink 1000. The images were viewed dichoptically from a 70 cm viewing distance through active circular polarisers. They were displayed on a rear-projection screen (67deg x 41deg) using a Propixx projector. Bilateral defocus of 1.00DS and 2.00DS, and unilateral defocus of 1.00DS, 1.50DS and 2.00DS was induced by convolving these natural images with the PSF calculated for the appropriate pupil size of each participant. After aligned presentation of stimulus, a step change in disparity was presented for 1.5sec. 8 typically developing children (age:6-8 years) participated. For the 16-degree radius natural images, open-loop response amplitudes were greater for convergence than divergence at all disparities (1, 2 and 4 degrees) ($p<0.001$). Induced bilateral and unilateral defocus had no effect on these amplitudes ($p>0.05$), similar to 2-degree image size cartoons (Mestre et al., 2022). Our previous work demonstrated that continuous vergence tracking responses were significantly reduced with simulated anisometropia of more than 2D for the small targets. This needs to be evaluated with larger natural images for children. In combination with our previous study, these data indicate that both adults and 6-8 year-olds are able to make vergence responses to large step changes in disparity in the presence of defocus up to 2D. Permanent developmental disruption in motor binocular function in the presence of anisometropia is most likely initiated by attempts to maintain binocular fixation and track small changes in disparity.

56.424 MEASURING REFRACTIVE ERROR USING CONTINUOUS PSYCHOPHYSICS AND EYE TRACKING

Ethan Pirso[1] (ethan.pirso@mail.mcgill.ca), Jude Mitchell[2], Curtis Baker[1]; [1]McGill University, [2]University of Rochester

For animal subjects, or human patients who have difficulty with conventional measurement methods, finding the best optical correction for refractive errors can be challenging. Refractive corrections are of growing concern in vision research due to the high prevalence of myopia in humans, but also with some colony-reared animals used in research. Traditional methods like psychophysical paradigms require extensive training or retinoscopy, which in animals requires anesthesia. However, tracking a moving target on a blank background is a natural task that is relatively easy for subjects to learn and execute. Here we used continuous psychophysics and eye tracking to measure contrast thresholds and assess refractive errors. Using custom MATLAB software with PsychToolbox-3 and an EyeLink 1000 eye tracker, we evaluated refractive errors by monitoring contrast sensitivity during dynamic visual stimulus tracking with gradually decreasing contrast. Applying this approach to the task of tracking a Gabor stimulus, we evaluated seven spherical lens powers, from -3 to +3 diopters, in successive runs on two human subjects . Each contrast sensitivity measurement necessitated less than a minute of eye tracking data. Contrast thresholds were derived from the positional errors between the target stimulus and the subjects' gaze positions. A plot of contrast threshold vs. lens power showed a clear dependence on positive diopter values and shallow dependence on negative ones, likely due to partial compensation from accommodation. We found that each subject's optimal lens power coincided with their previously measured corrected-to-normal vision. Our findings demonstrate the utility of continuous psychophysics integrated with eye tracking for more ecologically valid measurements of contrast sensitivity and refractive errors. This method could be used for clinically challenging human populations, and might be adapted for non-human primates such as marmosets or macaques, extensively used in vision research, thereby eliminating the need for anesthesia in retinoscopy or prolonged behavioral training.

Funded by Canadian NSERC (RGPIN-2023-03559) and CIHR (MOP-119498) grants to CB.

56.425 BEHAVIORAL AND OCULOMOTOR EFFECTS OF SCOTOMA AWARENESS TRAINING IN PATIENTS WITH CENTRAL VISION LOSS

Marcello Maniglia[1] (mmanig@ucr.edu), Jason Vice[2], Elam Cutts[2], Elliot Maxwell[2], Kristina Visscher[2], Aaron Seitz[3]; [1]University of California Riverside, [2]University of Alabama at Birmingham, [3]Northeastern University Boston

Macular degeneration (MD), a pathology affecting central vision, is the primary cause of visual impairment in the Western world. Late-stage MD leads to blind spots in the central portion of the visual field (scotomas). Loss of central vision significantly hampers patients' daily activities, compromising their quality of life. MD patients spontaneously adapt to this condition by developing compensatory oculomotor strategies, with the most common being the adoption of a Preferred Retinal Locus (PRL), a peripheral region substituting for the fovea. However, time course of PRL development is often suboptimal, taking up to months, with some patients failing at developing one at all. Some visual rehabilitation interventions teach patients to develop and maintain a PRL; however, these protocols encounter challenges. MD patients are often unaware of the size, location, and shape of their scotoma, thereby hindering visual training outcomes. To better understand what training approaches may be efficacious, several studies have employed real-time eyetracking to simulate central vision loss as a model system for the study of compensatory oculomotor strategies in MD. This research showed that eye movement behaviors observed in MD, such as PRL development, can be observed in simulation of central vision loss as well. Importantly, these behaviors emerge after few hours of training, unlike the months usually required in patients, possibly due to the hard-edged, gaze-contingent occluder. Here, we investigated whether using gaze-contingent displays to help patients visualize the retinal scotoma on a computer screen in real-time while performing 10 sessions of a visual discrimination task, would increase scotoma awareness and promote rapid PRL development. We will present and discuss preliminary results from this scotoma awareness group compared to the same training in participants without the visualization of the scotoma, including PRL development and location across training sessions, and behavioral outcomes, such as reading speed and visual acuity.

Tuesday, May 21, 2:45 – 6:45 pm, Pavilion

Binocular Vision: Disparity, stereopsis and suppression

56.426 NO CORRELATION BETWEEN INTEROCULAR DELAY AND STEREOSENSITVITY IN HEALTHY ADULTS

Alexandre Reynaud[1,2] (alexandre.reynaud@mcgill.ca), Paul Colin[3], Daniel Gurman[1], Reza Abbas-Farishta[3]; [1]Dept of Ophthalmology and Visual Sciences, McGill University, [2]BRaIN program, Research Institute of the McGill University Health Center, [3]École d'Optométrie, Université de Montréal

The perception of depth and objects in space hinges on the collaborative information received from both eyes in the visual system. Stereoscopic acuity, defining the smallest perceivable disparity between the 2 eyes, typically ranges from 30 to 60 seconds of arc in healthy individuals. To perceive stereoscopic vision accurately, the information from the 2 eyes must be received synchronously by the visual brain. This study aimed to uncover the correlation between interocular delay and stereoscopic vision quality in healthy vision. We measured and compared stereosensitivity and interocular delay in 29 healthy adults. Interocular delay was measured using a novel continuous psychophysics paradigm (Burge and Cormack, BioRxiv, 2020). Stereosensitivity was measured using a novel stereotest (Tittes et al., Vision Research, 2019) and the standard Randot Stereotest in 29 healthy adults. Our results suggest that there is no correlation between stereoacuity and interocular delay in our healthy population. Although interocular asynchrony is associated with several binocular vision disorders, we did not observe a correlation between interocular delay and stereoacuity in our healthy population. It is worth noting that, in both of our tests, stereosensitivity was measured using an unlimited presentation time. It is possible that interocular delay has more impact when using shorter presentation times.

Startup fund from the Research Institute of the McGill University Health Center

56.427 BINOCULAR COMBINATION UNDER ASYNCHRONOUS VIEWING CONDITIONS

Daniel Gurman[1], Alexandre Reynaud[1]; [1]McGill University

In the standard model, two main processes underlie binocular combination: binocular summation, whereby the input from each eye is added together, and interocular suppression, whereby each eye suppresses the other's input. To function properly, both processes require cooperation between the eyes, both in terms of the spatial information they contribute, and presumably, the temporal synchronization of information processing. Previous research indicates that de-synchronization of the ocular inputs generally impairs binocular vision. However, the effect of varying degrees of de-synchronization, and whether summation and suppression are affected to the same degree, is not known. Our goal was to compare the effect of precisely de-synchronizing the ocular inputs on psychophysical measures of summation and suppression. To measure interocular suppression, we employed a dichoptic masking paradigm using a two-alternative-force-choice task. A Gabor target and noise mask were displayed on a passive 3D screen at the same spatial location but with a variable temporal interval between them. To achieve a high degree of temporal precision, stimulus presentations were brief, lasting 8.3ms. Target detection thresholds were measured for each temporal interval between the target and the mask using a 2-up 1-down staircase. For binocular summation, two targets with identical adjustable contrasts were presented dichoptically at different inter-stimulus intervals. The detection threshold of the 2 targets was measured using a 2-up 1-down staircase. Dichoptically, threshold elevation tended to increase with increasing asynchronicity, particularly when the mask followed the target. Summation index tended to decrease with increasing asynchronicity. At a low spatial frequency, effects were larger and more consistent across tasks. Our findings demonstrate that desynchronization of the ocular inputs impacts both main processes underlying binocular combination and provides insight into the binocular deficits observed in conditions associated with natural interocular asynchronicity, such as amblyopia.

This research was funded by the Projet-Pilot Grant from the VHRN and the startup fund from the RI-MUHC

56.428 TEMPORAL LATENCIES AND POSITION UNCERTAINTY IN STEREOSCOPIC AND LUMINANCE MOTION USING A CONTINUOUS EYE-TRACKING TASK

Ignacio Serrano-Pedraza[1] (iserrano@ucm.es), Ichasus Llamas-Cornejo[1], Omar Bachtoula[1]; [1]Department of Experimental Psychology. Universidad Complutense de Madrid

Stereovision has an unexpected anisotropy: detecting sinusoidal depth corrugations is easier at low spatial frequencies when they are horizontally oriented than when they are vertical. Here, we are interested in the study of the temporal characteristics of this anisotropy. We measured the response of the optomotor system to different suprathreshold stereoscopic stimuli and compared it to the response to luminance stimuli. The stereoscopic stimuli, made with dynamic random dot stereograms, were half-cycle sinusoidal depth corrugations with disparity 2.1 arcsec and the background set to 0 disparity (i.e. 3D bar-shaped on a zero-disparity background). The luminance stimuli were also half-cycle sinusoid with a contrast of 0.8 and with a black background (i.e. white bar on black background). The bar-shaped stimuli moved following a one-dimensional random walk for four seconds and the subject task was to keep the bar centered in the fovea. We tested four spatial frequencies: 0.1, 0.2, 0.4, & 0.8 c/deg and two orientations (e.g. Vertical vs. Horizontal). To analyze the data, we used the eye movement correlogram technique by cross-correlating version velocities and stimulus velocities. The velocity correlogram for luminance stimuli had a positive peak with averaged latency around 140 ms for vertical and 160 ms for horizontal orientation. Interestingly, we found the longer peak latencies for stereo, around 370 ms for horizontal and 425 ms for vertical orientation. We also found higher position uncertainty for stereo than for luminance (using the standard deviation of the positive peak of the correlogram). For luminance stimuli, this uncertainty was higher for horizontal stimuli than vertical ones; and for stereo, the position uncertainty was higher for vertical than horizontal depth corrugations. In summary, our results show longer lags and higher position uncertainty in stereo than in luminance motion perception, and a strong stereoscopic anisotropy with suprathreshold stimuli.

Supported by grant PID2021-122245NB-I00 from Ministerio de Ciencia e Innovación (Spain) to ISP.

56.429 GLOBAL INTERNAL DISPARITY NOISE INCREASES WITH RISING LEVELS OF DISPARITY PEDESTAL

Jian Ding[1] (jian.ding@berkeley.edu), Michelle Y. Ma[1], Hilary H. Lu[1], Dennis M. Levi[1]; [1]University of California, Berkeley

In our previous study (Ding et al., 2023, VSS), we described an equivalent noise model incorporating both global and local internal disparity noises to provide a unified explanation of absolute and relative minimum disparity thresholds. The model predicts smaller thresholds for relative disparity compared to absolute disparity, as global noise present in absolute disparity detection is canceled when detecting relative disparity. However, the model was only tested within the vicinity of the zero-disparity plane. To extend the model beyond the zero-disparity plane, we developed psychophysical procedures involving the introduction of external disparity noise to random-Gabor-patch (RGP) stereograms to measure absolute and relative minimum disparity thresholds at different disparity pedestals. To measure absolute disparity thresholds, two RGP stereograms with external disparity noise were presented in two temporal intervals, one with a mean disparity greater than the pedestal disparity and the other smaller. Observers discerned which interval appeared closer. For relative disparity thresholds, we presented an RGP stereogram with different mean disparities (one larger and one smaller than the pedestal disparity) in the top and bottom halves, along with external disparity noise. Observers identified whether the top or bottom of the stereogram appeared closer. We measured disparity thresholds across 7 pedestals and 6 external noise conditions for both absolute and relative disparity detection, utilizing the constant stimulus method. Consistent with previous studies, our results demonstrated an elevation in both absolute and relative disparity thresholds with increasing disparity pedestals. The proposed model provides a unified explanation for all data sets. Our modeling shows that global disparity noise increases with rising levels of disparity pedestal, contributing to increased absolute disparity thresholds. In contrast, local disparity noise seems to remain unaffected by disparity pedestals. The increase in relative disparity thresholds with disparity pedestals is more likely attributed to a decline in detection efficiency.

NIH RO1EY030544 and NIH RO1EY020976

56.430 THE MECHANISMS OF CROSSED AND UNCROSSED DISPARITIES IN COARSE STEREOPSIS

Penghan Wang[1] (939740232wph@gmail.com), Alexandre Reynaud[1], Robert Hess[1]; [1]McGill University

Stereopsis, our ability to perceive depth, is a fundamental aspect of vision. It allows us to judge whether objects are "in front of" or "behind" each other. High variability in the perception of stereopsis has been reported in the population. Previous studies suggest that this variability may results from different mechanisms subserving the perception of crossed and uncrossed stereopsis. In our previous study, we focused on fine stereopsis. Here we focused on coarse stereopsis. We investigated the difference between crossed and uncrossed stereopsis mechanisms using an identification-at-threshold paradigm. We used a 2-by-2 forced choice procedure where participants had to both report the interval with the stimulus and judge whether the stimulus was crossed or uncrossed. The stimulus consisted of a gaussian bump (size) in a filtered noise texture (0.7 cycles per degree). Preliminary data revealed that typical observers were able to consistently achieve perfect discrimination. However, at extremely large disparities, some individuals could not discern polarity, probably because they could not fuse anymore. As observed before for fine stereopsis, these results suggest that crossed and uncrossed disparities are mediated by 2 different channels for coarse stereopsis.

56.431 UNRAVELING THE IMPACT OF STEREOSCOPIC VISION ON DAILY TASKS IN YOUNGER AND OLDER ADULTS

Adrien Chopin[1,2] (adrien.chopin@gmail.com), Diana Rdeini[1,3], Catherine Agathos[2], Chiara Ciucci[1,4], Yuling Wang[1], Valérie Parmentier[5], Denis Sheynikhovich[1], Angelo Arleo[1,5]; [1]Sorbonne Université, INSERM, CNRS, Institut de la Vision, 17 rue Moreau, F-75012 Paris, France, [2]The Smith-Kettlewell Eye Research Institute, San Francisco, United States, [3]Hôpital National de la Vision, Université Paris Cité, Paris, France, [4]School of Advanced Study Sant'Anna, Pisa University, Pisa, Italy, [5]Center Innovation & Technologies Europe, Essilor International SAS (EssilorLuxottica), Charenton-le-Pont, France

Introduction: Stereoscopic vision is crucial for perceiving depth. While recent research highlights its role in accurate distance estimation and object grasping, its influence in everyday tasks remains elusive, especially during aging. Methods: Sixteen younger and 16 older adults with healthy stereovision completed two daily tasks (making a cup of coffee, setting a table) within a laboratory apartment unit while hand movements were tracked. Each participant performed each task binocularly, allowing for the use of stereovision, or monocularly, preventing it. A novel eye-patching procedure equalized the visual field between monocular and binocular conditions, so that the condition difference reflected the impact of stereovision only. To assess the effect of attentional load, participants repeated each task while performing a concurrent spatial task. The order of conditions was counterbalanced in each group. Completion time and hand movement kinematics served as primary outcomes. We adjusted statistics for multiple comparisons. Results: Stereoscopic vision significantly reduced task completion time, and the reduction was greater in the older compared to the younger group ($t(252) = 3.98$, adjusted $p = 1.8\text{-}4$), with the interaction accounting for 5.2% of the variance. Attentional load increased completion time but without interacting with other factors. Participants moved their hands faster when using stereopsis ($t(508) = 2.35$, adjusted $p = 0.029$) but we found no difference in smoothness ($p>0.05$). Overall, completion time was influenced by translation path length, velocity, smoothness, and idle time of both hands. Conclusion: This study highlights the impact of stereoscopic vision on daily activities, particularly among older adults who benefited more in time savings in task completion. The observed enhancements imply that stereovision contributes to faster actions. Given many daily tasks rely on hand actions, age-specific interventions targeting stereovision might improve autonomy and quality of life in aging.

This research was supported by the Chair SILVERSIGHT ANR-18-CHIN-0002, by the IHU FOReSIGHT ANR-18-IAHU-01, and by the LabEx LIFESENSES (ANR-10-LABX-65).

56.432 PARTITIONING THE EFFECTS OF DISTINCT NATURAL-SCENE PROPERTIES ON VISUAL PERFORMANCE

David White[1] (davey@autistici.org), Johannes Burge[2,3]; [1]Neuroscience Graduate Group, University of Pennsylvania, [2]Bioengineering Graduate Group, University of Pennsylvania, [3]Department of Psychology, University of Pennsylvania

Stimulus variability is the primary source of perceptual uncertainty in everyday natural tasks. How do different properties of natural scenes contribute to this uncertainty? Using binocular disparity as a model system, we report a systematic investigation of how various forms of natural stimulus variability impact performance in a stereo-depth discrimination task. We show, using new analytical methods, the specific impact that two distinct sources of stimulus variability—luminance-pattern variation and local-depth variation—have on discrimination performance. First, we obtained a stimulus set, sampled from a natural stereo-image database with co-registered laser-based range-data at each pixel; this data set contains both natural luminance and natural depth variability. Next, we generated a second data set from the first, by flattening the local-depth profile of each stimulus; this data set contains near-identical luminance variability, but no depth variability. Stimuli were presented using a 2IFC paradigm. Each interval of each trial contained a unique stereo-image patch. The task was to report which patch appeared to be farther away. Each of three human observers collected 20,000 trials across two double-pass stereo-depth-discrimination experiments—one with each stimulus set. The stimuli in each trial were matched across passes, within and between experiments. The resulting stereo-depth discrimination thresholds indicate the absolute limits of discrimination performance at five disparity pedestals and three disparity-contrast levels. Analysis of each double-pass experiment indicates the relative importance of stimulus-variability and internal-noise in determining performance limits. And a novel quasi-quadruple-pass analysis, comparing the data from the two double-pass experiments, allows us to partition how luminance-pattern variability and local-depth variability limit performance in natural scenes. Our results show that luminance-pattern variation and local-depth variation have distinct and, surprisingly, almost completely dissociable effects on performance, and provide a rich picture of the factors contributing to human disparity discrimination performance in natural scenes.

National Eye Institute and the Office of Behavioral and Social Sciences Research, National Institutes of Health Grant R01-EY028571

56.433 PERFORMANCE METRICS OF REAL-WORLD PERCEPTION IN AUGMENTED REALITY

Jacob Duijnhouwer[1] (jduijnhouwer@magicleap.com), Savannah Halow[1,2], Gerrit Maus[1]; [1]Magic Leap, Inc., [2]Universiy of Nevada, Reno

The emerging technology of augmented reality (AR) headsets allows users to interact with virtual content that is stereoscopically superimposed on a view of their surroundings. This is achieved either with transparent displays that allow a direct view of the world ("see-through"), or with cameras that stream video to conventional displays inside the headset ("pass-through"). We sought to benchmark perception of the real world through various commercial AR devices, as well as quantify any discomfort that may result from their prolonged use. To this end, we developed a suite of perception and action tasks that provide performance metrics for complementary aspects of the visual experience. Tasks included color matching, hand-eye coordination, contrast sensitivity and acuity measurements, motion discrimination, visual-field measurements, catching balls, and a demanding locomotion task. The selection and design of these tasks was informed by the anticipated pros and cons of see-through and pass-through systems, such as latency of pass-through devices, and the reduced field of view in see-through headsets. Participants (N=24) performed the tasks while wearing different headsets (Magic Leap 2, Meta Quest Pro/3, Varjo XR3) and in a naked-eye control condition. A single session comprising all tasks took about one hour. To prevent carry-over of potential discomfort, sessions were conducted on separate days. We assessed visual discomfort using a standardized questionnaire that was administered before and after each session, as well as interview style questions. Our results establish the benchmark suite as sensitive and comprehensive, and provide detailed quantifications of the pros and cons of each technological approach in relation to the naked eye baseline.

56.434 THE DICHOPTER

Kenneth Brecher[1] (brecher@bu.edu); [1]Boston University

The Dichopter® is a viewing device designed to aid the viewer in simultaneously or sequentially seeing pairs of images displayed on a smartphone screen. The Dichopter consists of two 100 mm focal length plano-convex lenses mounted in a foldable bracket with an effective septum that blocks cross talk between the eyes. It can be used to view 3D pairs of images displayed in parallel (SBS) format on a smartphone screen. These can be taken from the Web or any other source or be created by the user by shooting two images using the "cha-cha" technique with a single smartphone. The images pairs can then be combined for visual display by using software such as "i3DSteroid." The word "Dichopter" was created to distinguish this device from binoculars, stereoscopes, synopters and other two eyed viewing devices. The current version of the viewer is based on over a decade of development by the author. It is being manufactured by the Korean optics company (MOCOM) by modifying its existing short focal length (5 cm) VR viewer. The image pairs displayed can be of many kinds: they can be of the same image (synoptic view); or they can be two different ones leading to the perception of depth (stereo pairs); or two different images with neither one containing sufficient information to form a single image (e.g., controllable random dot stereograms); or two different color images leading to the percept of another color (e.g., red to one eye, green to the other possibly leading to the percept of yellow); or images presented separately in time; and many other possible pair combinations. A web site has been developed allowing for all these possible uses: https://dichopter.com/.

Tuesday, May 21, 2:45 – 6:45 pm, Pavilion

Object Recognition: Basic features

56.435 INFLUENCE OF BACKGROUND ON THE SPATIAL-FREQUENCY CHANNEL FOR OBJECT RECOGNITION

Ajay Subramanian[1] (as15003@nyu.edu), Najib J. Majaj[1], Denis G. Pelli[1]; [1]New York University

Campbell & Robson (1968) found that grating detection and discrimination are mediated by spatial-frequency-selective filters. Motivated by this, later work found a single “channel”, always an octave-wide, for the recognition of letters, faces, and natural objects. However, it is unclear what properties of the image influence this recognition channel. Can we break the octave-wide channel result by altering image properties? Recognizing an object requires grouping of the relevant features in the presence of possibly irrelevant background features (Wertheimer, 1923). To separate objects from background in this way, an observer might use several available cues. If spatial frequency is one of these cues, i.e., if the observer extracts the object using spatial-frequency features absent in its background, we would expect that simply removing the background would allow the observer to use more spatial frequencies for recognition. We test this hypothesis by measuring the spatial-frequency channel for a 16-way ImageNet object categorization task separately in the presence and absence of image background. We find that removing the background increases the median channel bandwidth from 1 to roughly 1.5 octaves. Thus, without a background, people can use more frequencies to recognize an object.

56.436 IMPROVEMENT OF ACUITY FOLLOWING MOTION ADAPTATION: THE ROLE OF LOW SPATIAL FREQUENCIES

Steven Dakin[1,2] (s.dakin@auckland.ac.nz), Samuel Schwarzkopf[1,2], Selassie Tagoh[1]; [1]University of Auckland, [2]University College London

Motion adaptation can transiently improve visual acuity (VA; Lages et al, Psych. Sci, v28, p1180-8) which is surprising given that adaptation typically reduces contrast sensitivity (CS). Here we consider the role of information at different spatial frequencies (SFs) in this phenomenon. A single interval 4AFC identification task and a 2IFC contrast discrimination task were used to measure VA and CS respectively. We tested 19 normally sighted observers. VA was measured using both conventional and vanishing tumbling T-optotypes (which contain little low SF structure). CS was measured using SF band-pass filtered T’s (at 3.75, 7.5 and 15 c/deg) presented either in isolation, or on a fixed-contrast SF low-pass “pedestal”. We report that adaptation to receding motion improves VA (by -0.051 logMAR change, $p < 0.001$) for conventional but not vanishing optotypes (-0.004 logMAR change, $p = 0.73$), indicating a crucial role for low SFs. Adaptation reduces CS particularly at the lowest SF (by 31%, $p < 0.00003$). In unadapted observers however, CS for higher SF information was enhanced/facilitated by the presence of low SFs (by 192%, $p < 0.00001$). We conclude that low SFs are necessary for adaptation-driven acuity enhancement and adaptation reduces sensitivity to them. However since low SF structure facilitates detection of higher SFs, adaptation cannot be improving acuity indirectly by improving detection. Instead suppression of low SFs must directly improve performance at the recognition stage and (it follows) recognition must operate on information that has been combined across SF (cf MIRAGE, Watt & Morgan, 1985, Vis Res, v23 pp1465-77). We show how a MIRAGE-like model - operating on visual information that has been subject to suppression of low SF structure - predicts improved localization of features within letters that could in turn support superior recognition near the acuity limit.

Supported by the Marsden Fund Council from Government funding, managed by Royal Society Te Apārangi

56.437 MORE PREDICTIVE, EASIER TO DETECT? CONTRAST SENSITIVITIES IN DIFFERENT PREDICTABILITY CONTEXTS

Seyoon Song[1] (songseyoon@daum.net), Minsun Park[1], Chai-Youn Kim[1]; [1]School of Psychology, Korea University

Top-down information such as prediction can aid sensory processing, since bottom-up signals can often be dynamic, ambiguous, and even low in salience (Bar, 2004). Our previous study showed higher contrast sensitivity to targets matching the predictive information, suggesting enhancements in human low-level perceptual performance due to predictive processing (Song et al., 2023). Building on this, we explored whether this enhancement effect can be modulated by the degree of predictability. The task was to report the orientation of the target Gabor patch, whose contrast level was adaptively manipulated with the 1-up-1-down staircase procedure. Three conditions existed: prediction-regular, prediction-irregular, and control. For the prediction-regular and irregular conditions, a stream of three Gabor patches appeared whose orientations sequentially differed to induce the impression of rotation, followed by a target Gabor patch of orientation matching the rotation direction. Compared to the uniform step angles of 30˚ in the regular condition, the irregular condition introduced an irregular 60˚ step of rotation in the middle to weaken the predictability and reliability of the preceding information. The preceding and target stimuli in the control condition were in random angles. Results showed enhanced contrast sensitivities in both prediction conditions compared to the control, replicating the previous study. Of more relevance to the present study, the group difference between the regular and irregular conditions was insignificant, seeming to suggest against the modulatory effect of predictability. However, a subset of participants showed difference in contrast sensitivity between the two conditions, giving hints of potential predictability effects. This inter-individual variability in the results might stem from the tendency discovered from many participants to find their own rules in the preceding stream, possibly leading to additional predictive information of confounding influence. Considering this variance in further investigations, we aim to provide extensive evidence of predictive processing in low-level visual perception performance.

NRF-2023R1A2C2007289

56.438 WHICH SHAPE FEATURES DETERMINE DETECTABILITY OF CAMOUFLAGED RADIAL FREQUENCY PATTERNS?

Wei Hau Lew[1,3] (wlew@central.uh.edu), Samuel P. Smithers[2], Peter J. Bex[2], Daniel R. Coates[1,4]; [1]University of Houston College of Optometry, Houston TX, [2]Department of Psychology, Northeastern University, Boston MA, [3]Schepens Eye Research Institute of Mass Eye and Ear, Department of Ophthalmology Harvard Medical School, [4]University of Washington, Department of Ophthalmology, Seattle WA

When the texture or pattern of stationary targets is closely matched to the background, the detection of these camouflaged targets relies on identifying specific features such as the edge of its outline. Previously, we reported the results of behavioral experiments that tested the detection time of targets defined by radial frequency (RF= 3, 4, 5, 8, 10, 14, and 20) and amplitude (0.1, 0.25, and 0.5) with a pink noise (1/f) texture against a pink noise background. At low amplitude, mid RF targets (i.e. 8 and 10) were among the hardest to detect. For RF≥ 8, the time taken to detect the targets decreased (targets became easier to detect) as amplitude increased. Since it is still poorly understood what aspects of a target shape influence target detectability, here the local and global features of each shape were analyzed. We calculated the global features of perimeter, area, and aspect ratio (width/height of each shape "arm"). The osculating circle method was used to compute the local curvature, its Sum-Squared Difference (SSD), and Mean of Absolute Difference (MAD). We found that target detection time as a function of the perimeter had an inverted U-shape. The area did not systematically affect detection time. The aspect ratio also had an inverted U-shape, whereby targets were easier to detect when their arm's aspect ratio was <2. For target curvature, there was a positive correlation between SSD of the curvature and target detectability (Pearson's r= 0.844, p <0.001), but no correlation for MAD. While the local SSD was directly correlated and most predictive, global features (perimeter and aspect ratio) can also influence detectability.

56.439 FEATURES FOR VISUAL OBJECT RECOGNITION.

Martin Arguin[1,2] (martin.arguin@umontreal.ca), Marie-Audrey Lavoie[1], Mélanie Lévesque[1,2], Gabriela Milanova[1], Pénélope Pelland-Goulet[1], Marc-Antoine Akzam-Ouellette[1,2], Youri Tassé[1]; [1]Universite de Montreal, [2]Centre de recherche, Institut Universitaire de Gériatrie de Montréal

Human visual object recognition largely relies on shape information. However, the nature of the shape features that actually underlie this task remain largely unknown despite the wealth of competing theories aiming to account for the code by which human vision represents shape. Here, we report a series of five object recognition experiments using a spatial sampling paradigm (cf. Bubbles) to calculate classification images (CIs) that demonstrate the efficient features used by human participants. In all experiments, the targets were behind an occluding mask and partially revealed for 100 ms by a collection of 12 circular gaussian apertures of 0.8° in diameter. Participants pressed a keyboard key to indicate the identity of the target. Response accuracy was maintained at 50% correct by manipulating the degree of degradation of the target image. The experiments essentially differ from one another in terms of the class of stimuli and the exposure of instances from various viewpoints or not. The mean CIs in all experiments constitute a fair representation of those of individual participants. Moreover, when only the significantly effective features from these CIs are visible, this image is easily mapped to the target object by any normal human observer. However, the CIs of human participants correlate poorly with those obtained from an ideal observer carrying out the task under the same conditions. There is no particular type of feature such as those proposed by major shape perception theories (e.g. concavities, convexities, edge intersections, object parts, etc.) that dominate in the mean CIs and feature sizes are quite variable. Overall, these features appear most compatible with the 'image fragment' theory proposed by Ullman and collaborators. Remarkably, for all objects that were presented along variable viewpoints, the regions on the object's surface which constituted the effective features were extremely similar across viewpoints.

Supported by the Natural Sciences and Engineering Research Council of Canada

56.440 OBJECT-SUBSTITUTION MASKING: THE ROLE OF LOW-LEVEL CHROMATIC SIMILARITY

Ryan Lange[1] (rlange@uchicago.edu), Michael J. McNicholas[1], Steven K. Shevell[1]; [1]University of Chicago

Synopsis: Object-substitution masking (OSM) with isoluminant chromatic stimuli tested whether flanker-target similarity effects can be selective for low-level (L/M- versus S-cone) chromatic responses, or only higher-level (integrated) representations of color. Tests revealed separable flanker-target similarity effects for the two cone-isolating chromatic axes, and hence a contribution from low-level chromatic representations to OSM. Background: OSM is generally thought to mask targets by substituting or updating one object-level representation to another. Some studies, however, have found feature-level effects in OSM. For example, color-tilt conjunction targets and flankers show greater flanker-target color similarity effects on color masking than tilt masking, and vice-versa (Gellatly et al., 2006; Huang et al., 2018). Whether OSM acts only on object-level representations, or can act on lower-level representations, is relevant to the 'microgenesis' of conscious visual awareness. Methods and results: Subjects reported target color from a 4-probe array (non-forced choice) in an OSM protocol. Within an isoluminant plane, target and distractor chromaticities were on either the L/M- or S-cone axis; chromaticity of the single flanker was either along one axis (50% of trials) or along an intermediate chromatic direction (other 50%). The effect of flanker-target chromatic similarity on masking magnitude was tested using similarity relations at three levels: "same" (flanker and

target identical); “similar” (as in “same” but with added +/- modulation along the non-target axis); or “different” (no shared modulation on the target axis). Planned statistical contrasts showed less masking in the “different” condition than in either the “same” or “similar” conditions, but not between the “same” and “similar” conditions. Conclusions: The observed flanker-target similarity effect for the target axis, but not for added modulation on the non-target axis, is consistent with OSM mediated by low-level chromatic mechanisms. This suggests that OSM may act within and across multiple neural levels.

56.441 COMPOSITE OBJECT REPRESENTATION MAKES TRACKING THROUGH ROTATION DEFICIENT

Qihan Wu[1] (qwu30@jhu.edu), Jonathan Flombaum[1]; [1]Johns Hopkins University

A composite assembles a representation of an object from different kinds of variables. Some variables describe the parts of the object within an internal frame of reference [handle at -4 on the object’s principal axis (OPA)]. Others describe the object’s location in space [hammer at 3,5 on the table]. And some describe the current relationship between the prior two types [OPA aligns to the negative vertical of the table, tilted clockwise by 20]. That brittleness of this last variable type has been cited to explain the specificity of orientation perception deficits in neuropsychological cases, and nonuniform error types in working memory tasks. We predict that the same susceptibility to error should make it uniquely difficult to track the end of a rotating object. We placed two identical discs on opposite sides of an invisible square. A second pair of discs was arranged in the same way on the perimeter of a larger square that shared the same origin. Two discs (one from each pair) were designated as the tracking targets. Participants tracked the targets as the squares rotated around their center point. (The motion included occasional changes in direction). Performance was significantly lower in this rotation condition compared to a highly similar, but nonrotational (instead translating) motion around the perimeters of the squares. In Experiment 2, a visible line connected the discs in each pair, now impairing translational motion. This is consistent with difficulty in the task caused by a reliance on object representations, a reliance that can be caused by visible connections and by the implied connections of rotation. In a third experiment we predicted trial-by-trial errors by identifying configuration changes that demand re-expression of relational variables in an object composite. These experiments support the hypothesis that object representations are built from composites whose formats can constrain visual processing.

56.442 AN OCCIPITOTEMPORAL REGION THAT IDENTIFIES RELEVANT FEATURES

Yuxuan Zeng[1] (zeng.774@osu.edu), Ren E Hentz[1], David E Osher[1]; [1]The Ohio State University, Department of Psychology

Visual stimuli are recognized first and foremost by the relevant features that differentiate them from other stimuli. How does the brain extract the features that are relevant for identifying a specific stimulus or category? While low-level features such as orientation or spatial frequency are represented as early as in V1, they still need to be integrated to form more informative features that are relevant for identifying specific object categories. Participants were trained for 1.5 hours in the lab to categorize 2 sets of novel stimuli; in the first set (featural), category membership was defined strictly by features, while in the second set (control), features were identical and category membership was defined by the locations of features. Afterwards, participants were scanned while performing the same categorization task. We also collected an fMRI localizer for high-level visual categories, including faces, bodies, scenes, and objects. We identified a bilateral region in the ventral visual stream that is specifically recruited while participants are extracting stimulus features. This region lies on the ventral surface just beyond areas V4/V8, and extends to the posterior fusiform gyrus. Interestingly, this region is not selective for any high-level visual category, and is responsive to faces, bodies, and scenes, and strongest to objects and even scrambled objects (object selectivity is typically defined by the contrast of objects > scrambled). In addition, we identified several visual fROIs, and compared featural responses in each fROI. Only the bilateral posterior fusiform sulcus (pFS, an object-selective region) showed a significant difference between conditions. Our findings reveal a region in the intermediary ventral visual stream associated with feature-based processing that does not overlap with other category-specific areas. We propose that this could be a mid-level region that extracts featural information before stimulus identity is determined.

56.443 SPIKY AND STUBBY OBJECTS IN HUMAN VISUAL PERCEPTION

Ashley Au[1] (ashley.au@yale.edu), Marvin Chun, Yaoda Xu; [1]Yale University

Object recognition is an integral part of the visual system that allows humans to extract different physical characteristics of the objects around them. Regions of the macaque inferotemporal (IT) cortex have shown to exhibit differential preferences for objects based on their spikiness/stubbiness (Bao et al., 2020). The same spiky/stubby distinction is also present in convolution neural networks (CNN) trained for object classification, suggesting that this is likely a fundamental feature dimension in visual object representation (Yargholi & Op de Beeck, 2022). However, similar findings have not been observed in human fMRI studies (Yargholi & Op de Beeck, 2022; Coggan & Tong, 2023). The aim of this study is to reexamine this discrepancy by creating a better-controlled inanimate stimulus set and by testing whether the spikiness/stubbiness distinction influences visual search performance. Based on the output from a CNN previously shown to correspond to macaque IT cortex in its representations of spiky and stubby objects, we selected 6 pairs of spiky and stubby objects matched in semantic category, with each object containing 8 different exemplars. We asked human participants to search for the presence of a target object exemplar among the exemplars of a distractor object. Across two experiments using either intact object images ($n = 13$) or images equated in low-level features (e.g., luminance, contrast, spatial frequency; $n = 14$), we observed faster search speeds between the spiky and stubby objects ($M = 631$ and 977 ms for both experiments, respectively) than within spiky or within stubby objects ($M = 724$ and 1148 ms; differences of within and between, $ps < .001$). These results are consistent with spikiness/stubbiness being a salient dimension in visual object representation. Next, we will collect fMRI responses from the human IT cortex for these objects and examine whether a similar saliency is present in neural responses.

Grant support: YX is supported by NIH grant 1R01EY030854.

Object Recognition: Structure of categories

56.444 RESPONSES IN HUMAN EARLY VISUAL CORTEX ARE MORE SENSITIVE TO TASK DIFFICULTY THAN OBJECT CATEGORY.

June Hee Kim[1] (junehee96@gmail.com), Cassie Joynes[2], Grace Edwards[1], Chris I. Baker[1], Elisha P. Merriam[1]; [1]Laboratory of Brain and Cognition, National Institute of Mental Health, NIH, Bethesda, Maryland, United States of America, [2]Center for Alzheimers & Related Dementias, National Institute on Aging, NIH, Bethesda, Maryland, United States of America

Early visual cortex (EVC) exhibits widespread hemodynamic responses in the absence of visual stimulation, which are entrained to the timing of a task and not predicted by local spiking or local field potentials. These task-related responses (TRRs) are associated with aspects of cognitive and perceptual tasks, including reward magnitude and task difficulty, and correlate with markers of arousal, such as pupil size. Here, we tested the role of TRRs in an object discrimination task that has been used previously to study feedback signals in human V1. We measured brain activity (7T BOLD fMRI) and pupil size as subjects performed a two-alternative forced-choice shape discrimination task on peripherally-presented stimuli. On each trial, two stimuli from the Validated Circular Shape Space (Li et al., 2020) appeared simultaneously for 100 ms in diagonally opposite quadrants of the screen at 7 deg eccentricity. Subjects reported whether the stimuli were the same or different. In Expt 1, we partitioned the circular space into 3 equidistant sectors to create object 'categories,' with different categories presented in separate blocks of trials. In Expt 2, we modulated task difficulty by selecting stimuli from either proximate (for difficult trials) or diametrically opposed points (for easy trials) on the circle. We employed multi-variate pattern analysis to identify brain regions sensitive to object category and task difficulty. We were able to decode object category from the lateral occipital complex (LOC), consistent with its established role in object processing. But we were unable to decode category from EVC, inconsistent with the notion that LOC sends object-selective feedback to EVC. However, we could decode task difficulty, both from fMRI responses in EVC, and from trial-locked pupil responses. Our results demonstrate that TRRs are sensitive to important covariates of task performance, and underscores the significance of TRRs in fMRI studies of core visual functions.

56.445 INVESTIGATING THE IMPACT OF GAUSSIAN NOISE ON FACE RECOGNITION PERFORMANCE FOR HUMANS AND CONVOLUTIONAL NEURAL NETWORKS

Ikhwan Jeon[1], Connor Parde[2], Frank Tong[3]; [1]Vanderbilt University

Recent work has shown that convolutional neural networks (CNNs) perform worse than human observers at recognizing objects presented with superimposed Gaussian noise, indicating poor alignment between CNN models and humans (Jang et al., 2021). However, it is unknown whether Gaussian noise might also disproportionately impair CNNs at tasks of face recognition, which more-heavily rely on the processing of lower-spatial-frequency information. We evaluated humans and face-trained CNNs using face images with varying levels of Gaussian noise, where noise intensity was specified by the signal to signal-plus-noise ratio (SSNR). Human participants completed a face-recognition task wherein they viewed face images of 10 different celebrities (80 images/celebrity) in an intermixed, randomized order, and were instructed to select the identity shown in each image. Each face image was presented for 200ms at a randomly-assigned SSNR level (SSNR = 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.8, or 1). Then, the same set of face images was presented to a face-classification trained version of AlexNet. The identity response from the CNN was indicated by the identity node for which each input image elicited the highest response. Performance for both human participants and the CNN was quantified by accuracy as a function of SSNR and was compared to the results of object recognition (Jang et al. 2021). Overall, humans exhibited better performance than the CNNs. But both humans and the CNNs better performed in face recognition than object recognition (SSNR thresholds of human: 0.17 for faces, 0.26 for objects; of CNNs: 0.25 for faces, 0.5 for objects). Moreover, face-trained model performance appeared highly similar to human, suggesting similar degrees of robustness to Gaussian visual noise. The fact that face-trained CNNs provide a reasonable account of human recognition of noisy face images may suggest that these CNNs are fairly aligned with the human face recognition system.

Supported by NIH grants R01EY029278 and R01EY035157 to FT.

56.446 USING TEXTURE SYNTHESIS TO IDENTIFY THE FEATURES SUPPORTING COARSE AND FINE OBJECT CATEGORIZATION

Margaret Henderson[1] (mmhender@cmu.edu), Leila Wehbe[1], Michael Tarr[1]; [1]Carnegie Mellon University

A key question within visual cognition is how we determine an object's category, and which intermediate visual features contribute most strongly to this process. Past work indicates that low-level and mid-level features may be sufficient to support certain types of object category judgments even in the absence of high-level features. However, it is not known how the role of these features differs depending on the granularity of the categories. One possibility is that coarse categories may be detectable on the basis of simpler features, while finer-grained category distinctions might require more complex features to discriminate. Here, we tested this hypothesis behaviorally by leveraging a computational texture synthesis tool (Gatys, Ecker, & Bethge; 2015, NeurIPS), which synthesizes images that maintain the local textural statistics of a target natural image. We generated versions of target object images that matched the statistics of their representations at multiple layers of a deep neural network, resulting in images that continuously vary in their feature complexity. In a series of online experiments human participants judged - at either the coarse (superordinate) or fine (basic) level - the category of the manipulated images presented in either grayscale or color. Our results show that while category discriminability improves for more complex images, some categories – particularly natural categories such as "insect" and "vegetable" – can be discriminated at above chance levels even from the simplest texturized images. Consistent with our hypothesis, performance with the simplest images was higher for coarse versus fine categorization, and varied across individual object categories

within the experiment. These results indicate that features at multiple levels of complexity may contribute to object categorization.

This work was supported by the Carnegie Mellon Neuroscience Institute.

56.447 COMMON REPRESENTATIONAL FORMAT IN OBJECT-SELECTIVE VISUAL CORTEX FOR PHOTOGRAPHS AND DYNAMIC SKETCHES

Sharif Saleki[1], Kevin C. Hartstein[1], Peter U. Tse[1]; [1]Dartmouth college

Object recognition is a central function of the brain's ventral visual pathway. Many experiments have identified putative category-specific processing areas within human ventral temporal cortex. In particular, multivariate pattern analysis has revealed distributed representations of objects in Ventral Temporal (VT) cortex and Lateral Occipital Complex (LOC). Previous object recognition experiments have employed photographic images or photo-realistic line drawings as stimuli. However, humans also readily identify rough, quickly rendered, hand-drawn sketches of objects. Here, we selected stimuli from photographs and sketch drawings (collected from Google's "Quick, Draw!" dataset) in animate (faces and body parts) and inanimate (natural and artificial objects) categories. Participants viewed static images of the photographs and dynamically drawn sketches in separate runs while brain responses were recorded via fMRI. Consistent with past research, multivariate pattern classification methods distinguished representations of different objects in photographic images in VT/LOC. Furthermore, above-chance performance of the classifier in decoding brain activity corresponding to rough sketches indicated the presence of information about their categories in object-processing areas. Cross-decoding between formats also yielded significant positive results, and the representational similarity followed previously reported structure of similarity in these regions. Together, we find shared representations of static photos and dynamically drawn sketches of objects, indicating a high-level representation of object category in the visual hierarchy that is largely feature-independent with respect to the presented stimulus format.

56.448 DREAMON: ENHANCING DEEP LEARNING IN MEDICAL IMAGING WITH REM-DREAM-INSPIRED DATA AUGMENTATION

Luc Lerch[1,2] (luc.lerch@students.unibe.ch), Lukas S. Huber[1,2,4,5], Amith Kamath[2], Walter Senn[1], Mauricio Reyes[2], Alexander Pöllinger[3], Verena Obmann[3], Florian Dammann[3], Aurélie Pahud de Mortanges[3]; [1]Computational Neuroscience Group, Department of Physiology, University of Bern, Bern, Switzerland., [2]Medical Image Analysis Group, ARTORG Centre for Biomedical Research, University of Bern, Bern, Switzerland., [3]Inselspital, University Hospital, University of Bern, Bern, Switzerland., [4]Cognition, Perception and Research Methods, Department of Psychology, University of Bern, Bern, Switzerland., [5]Neural Information Processing Group, Department of Computer Science, University of Tübingen, Tübingen, Germany.

The efficacy of deep learning (DL) models in medical image analysis is significantly influenced by image quality. Image quality variation arises from diverse sources, including differences in imaging equipment, patient-specific factors like movement, and biological variability. Consequently, robustness against noise is a crucial factor for the application of DL models in medical image analysis. Here we investigate the impact of various common data augmentation strategies on the robustness of a ResNet-18 model in classifying breast ultra sound images, and benchmark the performance against trained human radiologists. Additionally, we introduce 'DreamOn', a data augmentation strategy employing a generative adversarial network (GAN) approach to create synthetic images inspired by generation of visual experience during REM sleep. The proposed method involves generating interpolations of training images that mimic the recombination of episodic memories observed in human dreaming. We evaluate the effectiveness of DreamOn in improving model robustness across diverse datasets, each treated with six increasing levels of either Gaussian, speckle, or salt-and-pepper noise, and compare its performance with other off-the-shelf data augmentation methods. Collected data indicates that while standard techniques enhance model robustness, they fall short in high noise environments where radiologists excel. Employing the DreamOn data augmentation, we were able to narrow this robustness gap. The presented study underscores the potential of integrating biologically inspired data augmentation techniques in DL models for medical image analysis. It highlights the importance of considering human-like perceptual and cognitive processes in developing AI tools, particularly in fields where expert human judgment remains the gold standard. The findings are particularly relevant for vision scientists interested in the intersection of artificial intelligence, human cognition, and medical imaging.

56.449 PATTERN OF ASSOCIATIONS ACROSS CATEGORIES IN VISUAL RECOGNITION

Laura Soen[1] (laura.soen@kuleuven.be), Céline Gillebert[1], Hans Op de Beeck[1]; [1]Brain and Cognition, Leuven Brain Institute, Faculty of Psychology & Educational Sciences, KU Leuven

Localized brain damage can result in relatively specific problems in mental faculties, including problems with visual recognition. It is still a matter of debate why certain deficits and combinations of deficits occur in visual recognition. Recent advancements in neuroscience have revealed the complexity of the system that supports visual recognition. In addition to category-selective regions, studies found distinct activation profiles among different categories based on multivoxel selectivity patterns. Our project aims to investigate the interrelation between visual recognition of distinct object categories, faces, and words, focusing upon the categories for which most selectivity has been observed in neuroimaging studies. The objective is to explore potential patterns, associations and dissociations, in visual recognition abilities. Therefore we developed a new test battery which measures the visual recognition of 10 object categories: faces, words, bodies, hands, houses, tools, food, animals (cats & dogs), musical instruments, and cars. Notably, the test battery makes use of 3D stimuli, challenging participants with visual recognition in various orientations. In an initial study involving 250 healthy participants, an exploratory factor analysis conducted on accuracies from different object categories unveiled two factors: one related to inanimate categories and another to animate categories. Moreover, the accuracies on inanimate categories displayed higher correlations with each other, as did the accuracies on animate categories, compared to the correlations between the accuracies on animate and inanimate categories, which were lower. This suggests a pattern in visual recognition depending on the degree of animacy. In subsequent stages, our goal is to demonstrate how the performance associations on our test battery is related to neural overlap in the brain, representational overlap in Deep Neural Networks (DNNs) and brain lesions in posterior stroke patients.

Research Foundation - Flanders

56.450 EXPLORING MENTAL REPRESENTATION OF VISUAL EMOJI SYMBOLS THROUGH HUMAN SIMILARITY JUDGMENTS

Yiling Yun[1] (yiling.yun@g.ucla.edu), Bryor Snefjella[1], Shuhao Fu[1], Hongjing Lu[1]; [1]University of California, Los Angeles

How do people align concepts learned from different modalities, such as visual and linguistic inputs? To address this question, we examined the representations of emojis, which are pictograms commonly used in linguistic contexts and convey distinctive visual characteristics that make them appear engaging. The representational similarity structure of emojis was measured using an odd-one-out paradigm. In Experiment 1, human similarity judgments were measured for 48 emojis from a wide range of emoji categories (faces, animals, objects, signs, etc). We compared human similarity judgments with model predictions from three types of models, including a language model (fastText) that is trained for word prediction in sentences, a vision model (Visual Auto-Encoder) that is trained to reconstruct input images, and a multi-modal neural network (CLIP) that learns visual concepts under language supervision. We found that CLIP correlated with human similarity judgments the highest (rho = .38), followed by fastText (rho = .36), and Visual Auto-Encoder (rho = .17). When controlling for linguistic semantics from fastText, CLIP maintained the significant semipartial-correlation with human judgments (sr = .34). The best performance of CLIP was not simply due to the combination of multimodal inputs since simply concatenating fastText and Visual Auto-Encoder embeddings resulted in a lower correlation (rho = .17). In Experiment 2, we used the 50 most frequently used emojis, which mostly include faces with different expressions and hand gestures. We found that all three models show correlations with human similarity judgments: CLIP (rho = .68), followed by fastText (rho = .52), and Visual Auto-Encoder (rho = .46). These results suggest that models trained with aligned visual and linguistic inputs in a multi-modal way best capture human conceptual representations of visual symbols, such as emojis. However, these models trained with general purposes are inadequate to capture fine-grained social attributes in emojis.

We dedicate this abstract to the memory of our co-author Dr. Bryor Snefjella, whose creativity, brilliance, insight and generosity make this project possible. This abstract was supported by NSF Grant BCS-2142269 awarded to H.L.

56.451 TIME-RESOLVED BRAIN ACTIVATION PATTERNS REVEAL HIERARCHICAL REPRESENTATIONS OF SCENE GRAMMAR WHEN VIEWING ISOLATED OBJECTS

Aylin Kallmayer[1] (aylin.kallmayer@gmail.com), Melissa Vo[1]; [1]Goethe University Frankfurt, Scene Grammar Lab, Germany

At its core, vision is the transformation of sensory input into meaningful representations. Understanding the structure of such representational spaces is crucial for understanding efficient visual processing. Evidence suggests that the visual system encodes statistical relationships between objects and their semantic contexts. Recently, however, a more fine-grained framework of hierarchical relations has been formulated ("scene grammar") according to which scene understanding is driven by real-world object-to-object co-occurrence statistics. More specifically, clusters of frequently co-occurring objects form phrases wherein larger, stationary objects (e.g., sink) anchor predictions towards smaller objects (e.g., toothbrush). Still, we know little about the mechanisms and temporal dynamics of these anchored predictions and whether the processing of individual objects already activates representational spaces characterized by phrasal structures. In the present EEG study, we aimed to quantify shared representations between objects from the same versus a different phrase within the same scene in a MVPA cross-decoding scheme paired with computational modelling to probe the format of shared representations. We presented objects from four different phrases spanning two different scenes (kitchen and bathroom) individually in isolation. Classifiers trained on anchor objects generalized to local objects of the same phrase and reverse, but crucially, not to objects from the same scene, but different phrase. This provides first evidence that phrase-specific object representations are elicited by the perception of individual objects. Computational modelling revealed that high-level semantic features quantified from Resnet50 successfully predicted the classifier's generalization matrix suggesting that late-stage recurrent processes are responsible for the observed generalization rather than low-level visual similarity between the objects. Overall, we provide novel insights into the temporal dynamics of encoded object co-occurrence statistics which seem to reflect a more fine-tuned hierarchical structure than previously assumed. Finally, this also provides a mechanistic account for the hierarchical predictions observed in efficient attention guidance through real-world scenes.

This work was supported by SFB/TRR 26 135 project C7 to Melissa L.-H. Võ and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project 'The Adaptive Mind') and the Main-Campus-Doctus stipend awarded by the Stiftung Polytechnische Gesellschaft to Aylin Kallmayer.

Tuesday, May 21, 2:45 – 6:45 pm, Pavilion

Spatial Vision: Machine learning, neural networks

56.452 SPATIAL SCRAMBLING IN HUMAN VISION: INVESTIGATING EFFICIENCY FOR DISCRIMINATING SCRAMBLED LETTERS USING CONVOLUTIONAL NEURAL NETWORKS AND CONFUSION MATRICES

Xingqi Raffles Zhu[1] (raffles.zhu@mail.mcgill.ca), Robert F. Hess[1], Alex S. Baldwin[1]; [1]McGill University

One limitation in our ability to discriminate different letters would be any spatial disorganization in the projections between different visual areas. This "scrambling" could be a source of a positional noise limiting human performance. In this study, we explored different forms this scrambling could take. Based on the idea that letter identification is supported by an optimal spatial frequency, we used spatially-

bandpass letters. We devised a physiologically-inspired decomposition and resynthesis scheme, to generate letters composed of log Gabor wavelets. The form of these wavelets is similar to that of an oriented "simple cell" receptive field. We then introduced two forms of scrambling. The first was scrambling at the input to the "oriented receptive field" stage (subcortical scrambling of the receptive field). The second was scrambling at the output from that stage (scrambling connections to the higher "cortical" stages). We also performed a bandpass noise control condition. To compare against human performance, we simulated the responses of both a template-matching observer (TMO) and three convolutional neural networks (CNNs). The three CNNs were trained on the letter stimuli to perform each of the three noise conditions. We computed human efficiency relative to CNN performance. We also characterized mistakes using confusion matrices and computed the population stability index (PSI) as a distance measure between mistakes made by human and model observers. We found the CNNs employed distinct strategies for each condition. Human relative efficiency was higher for subcortical than cortical scrambling. In bandpass noise, PSIs for both TMO and CNNs were comparable. For our scrambling conditions however, the PSI of TMO was significantly higher than that of CNNs in all but one comparison. Our results suggest that the human strategy for identifying scrambled letters is better captured by CNNs, which may share more similar strategies for identifying scrambled letters than a simple TMO.

This work has been supported by Réseau de Recherche en Santé de la Vision (RRSV), Fond de Recherche du Québec - Santé (FRQ-S), and Canadian Institutes of Health Research (CIHR).

56.453 TOWARDS DETERMINING THE LOCATION OF THE PREFERRED RETINAL LOCUS OF PATIENTS WITH MACULAR DISEASE: A DEEP LEARNING-BASED SIMULATION

Rijul Saurabh Soans[1], Dongcheng He[1], Susana T L Chung[1]; [1]University of California, Berkeley

In the absence of a healthy fovea, individuals with macular disease often rely on an alternate retinal location, the preferred retinal locus (PRL), for seeing. However, the factors determining the PRL location are still unclear. Previous studies have reported that the PRL location is not driven by exclusively optimizing visual acuity (Bernard & Chung, 2018) or sensitivity (Chung et al., 2023). This study tests the feasibility of applying the semantic segmentation approach to predict the PRL location based on the eccentricity effect and how visual acuity and sensitivity vary across the retina. We first generated 200 scotoma patterns based on the actual scotoma shape and size of 21 eyes with macular disease (by sampling along each of 360° meridians the location of the edge of each generated scotoma). We then placed each scotoma pattern on the retinal images of 79 healthy eyes to generate 15,800 images. Each sample of our simulated data was a 3-channel vector (256×256). The first channel consisted of the image in grayscale pixel values. The second and third channels consisted of visual acuity and sensitivity values, respectively, generated based on values reported in the literature. For training purposes (using a DeepLabV3 network with 13,000 samples), we defined the PRL region of each vector by imposing certain constraints: it is close to the edge of the scotoma (Chung, 2012); its acuity and sensitivity values are near expected thresholds; and it is located closest to the fovea after satisfying the preceding constraints. Testing with the other 2800 samples showed that between the predicted and "true" PRLs, the mean Jaccard index (percent-overlap of area) was 53.7%±25.7%, and the mean Euclidean distance between their centers was 8.6±22.6 pixels. Our results suggest that a model combining eccentricity and visual performance variation across the retina is feasible in predicting the PRL location.

Authors RSS & DH contributed equally. Grant Support: NIH Grant EY030253

56.454 MULTITASK MACHINE LEARNING OF CONTRAST SENSITIVITY FUNCTIONS

Dennis Barbour[1] (dbarbour@wustl.edu), Zhiting Zhou[1], Dom Marticorena[1], Quinn Wai Wong[1], Jake Browning[1], Ken Wilbur[1], Pinakin Davey[2], Aaron Seitz[3], Jacob Gardner[4]; [1]Washington University in St. Louis, [2]Western University of Health Sciences, [3]Northeastern University, [4]University of Pennsylvania

Contrast Sensitivity Functions (CSFs) represent useful diagnostic adjuncts for helping assess both retinal and central visual functionality. Gaussian Process (GP) classifiers have been shown to efficiently estimate individual CSF models by leveraging active machine learning for optimal stimulus selection. Model convergence in these cases can be achieved with between 10 and 50 actively selected stimuli. By assuming model independence, this disjoint process requires sequential estimation to obtain CSF models for multiple eyes or stimulus conditions (e.g., luminance, eccentricity). Conjoint estimators, on the other hand, have now been developed to estimate multiple CSFs simultaneously using an active multitask implementation. In the current study, conjoint CSF estimator performance was compared to disjoint performance on simulated eyes using generative models created from human data. The high degree of expected similarity between CSFs originating from different eyes or conditions allows conjoint learning between the related models. This procedure is designed to enable faster convergence than sequential disjoint model learning. Indeed, conjoint CSF estimation does speed model convergence over disjoint estimation under commonly encountered scenarios. These findings confirm that incorporation of additional information beyond immediate behavioral responses into new machine learning models of vision functions may improve visual system assessment.

R21-EY033553, R01-EY019693

56.455 GRAMSTATTEXNET: USING THE GRAM MATRIX OF MULTI-SCALE PYRAMIDS TO CONTRASTIVELY LEARN TEXTURE MODEL STATISTICS

Vasha DuTell[1,2] (vasha@mit.edu), Christian Kovesdi[1,2], Anne Harrington[1,2], Mark Hamilton[1], William T. Freeman[1], Ruth Rosenholtz[1,2]; [1]MIT CSAIL, [2]MIT Brain and Cognitive Sciences

Visual processing in area V1 is often modeled by multi-scale, oriented filters, both pre-computed such as the steerable pyramid, as well as learned. To understand how these V1 responses are combined in layers V2 and beyond, models for texture synthesis are often employed. These models use pre-defined summary statistics calculated on the correlations of V1 filter outputs (Portilla & Simoncelli,1999), and combined with spatial pooling, are used to understand peripheral visual processing (Balas, 2009, Freeman & Simoncelli, 2011). More recently, learned features in deep neural networks (DNNs) have been used to represent texture, employing the Gram matrix to encode the texture-like representation seen in mid-level peripheral vision (Wallis et al, 2017). While models with hand-picked statistics are well-validated, they cannot faithfully represent some texture families (Brown et al, 2023); DNN approaches on the other hand are based on biologically-implausible and over-parameterized representations. To address this, we propose a new framework to learn texture statistics. We employ the contrastive learning approach from StatTexNet (Koevsdi et al, 2023), modifying this model to learn elements from the Gram matrix of pyramid images, rather than learning pre-defined statistics. The learned component consists of a single-layer, fully-connected network, which reduces the set of statistics from the full Gram matrix, down to a reduced set of meta-statistics. It is trained through contrastive learning to push similar textures together, and pull dissimilar textures apart in representation space. As an indicator of successful learning, we show that this network successfully clusters both same- and similar- texture samples. We find that the learned weight matrix is innately sparse, with pyramid image auto correlations most highly weighted, and low-pass pyramids least utilized. This work demonstrates a both biologically-inspired and learned approach to the texture and peripheral vision models of V2, giving further insight into the complex transformations of mid-level vision.

CSAIL METEOR Fellowship, US National Science Foundation under grant number 1955219, National Science Foundation Grant BCS-1826757 to PI Rosenholtz, MIT SuperCloud and Lincoln Laboratory Supercomputing Center

56.456 RECURRENCE IS NEEDED TO ACCOUNT FOR THE SHARPER ORIENTATION-TUNED SURROUND SUPPRESSION FOR OBLIQUE VERSUS CARDINAL ORIENTATIONS

Huiyuan Miao[1] (huiyuan.miao@vanderbilt.edu), Frank Tong[1]; [1]Psychology department, Vanderbilt University

Surround suppression can reduce both the perceived contrast and the strength of V1 neuronal responses to stimuli within the classical receptive field when the neighboring surround is visually similar due to inhibitory neural interactions. Our previous study measured the orientation tuning curves of surround suppression and found that surround suppression was more broadly tuned around cardinal than oblique orientations, revealing an inverted oblique effect when compared to human orientation discrimination (Miao, Coggan, & Tong, VSS2023). Recent studies have demonstrated surround suppression in feedforward deep neural networks (DNN) optimized for object recognition (Pan et al., 2023). Here, we asked: Do DNNs show a similar inverted oblique effect in surround suppression orientation-tuning? We compared two types of models: a feedforward network (AlexNet) and a recurrent network (CORnet-RT). We presented these networks with both cardinal and oblique bandpass-filtered patterns to evaluate the magnitude of surround suppression across varying center-surround orientation differences (0-90˚). Stimuli with three different spatial frequency ranges and eight different target sizes were evaluated. While we found surround suppression in feedforward AlexNet's intermediate layers, no stable inverted oblique effect was detected. Is recurrence necessary for the inverted oblique effect? We evaluated CORnet-RT, which contains four computational blocks with recurrent processing analogous to V1, V2, V4, and IT. In each block, the model implemented convolution, group normalization (GN), and ReLU during feedforward processing, followed by recurrent processing in which the output of an area's last time point was used for processing in the next cycle. We found that while the network revealed a general effect of orientation-tuned surround suppression in first processing step of the V1 block, the inverted oblique effect appeared only after recurrent processing. Taken together, our results reveal that surround suppression leads to a counterintuitive oblique effect that can be attributed to recurrent processing in the human visual system.

Supported by NIH grants R01EY029278 and R01EY035157 to FT.

56.457 MAPPING MODELS OF V1 AND V2 SELECTIVITY WITH LOCAL SPECTRAL REVERSE CORRELATION

Timothy D. Oleskiw[1,2] (oleskiw@nyu.edu), R.T. Raghavan[1], Justin D. Lieber[1], Eero P. Simoncelli[1,2], J. Anthony Movshon[1]; [1]New York University, [2]Flatiron Institute

Neurons in macaque area V2 respond selectively to higher-order visual features, such as the quasi-periodic structure of natural texture, but it is unknown how selectivity for these features is built from V1 inputs tuned more simply for orientation and spatial frequency. We have recently developed an image-computable two-layer linear-nonlinear network that captures higher-order tuning from a sparse combination of subunits tuned in orientation and scale. This model can be independently trained to predict data from single-unit recordings of V1 and V2 neurons from the awake macaque, responding to a stimulus set comprised of multiple superimposed grating patches that localize oriented contrast energy. These optimized models accurately predict neural responses to other stimuli, including gratings and synthetic textures with higher-order features common to natural images. However, the high-dimensional parameter space of these models makes them difficult to interpret. A systematic comparison of V1 and V2 selectivity is therefore elusive. To address this limitation, we investigate neural tuning across our population of models in silico using local spectral reverse correlation (LSRC; Nishimoto et al., 2006). Here, we present thousands of ternary white noise stimuli to fitted models, computing a response-weighted windowed frequency spectrum across image coordinates. LSRC can effectively characterize the tuning properties of our model neurons, thereby estimating the linear and nonlinear components of model receptive fields. These estimates are qualitatively similar to our direct LSRC measurements made using identical methods from single-unit V1 and V2 recordings collected in separate experiments.

Simons Foundation, NIH EY022428-10

56.458 POSSIBLE OPTIMAL STRATEGIES FOR ORIENTATION CODING IN MACAQUE V1 REVEALED WITH A SELF-ATTENTION DEEP NEURAL NETWORK (SA-DNN) MODEL

Xin Wang[1], Cai-Xia Chen[2], Sheng-Hui Zhang[3], Dan-Qing Jiang[4], Shu-Chen Guan[5], Shi-Ming Tang[6], Cong Yu[7]; [1]Peking University, [2]Peking University, [3]Peking University, [4]Peking University, [5]Justus-Liebig Universität, [6]Peking University, [7]Peking University

The orientation tuning bandwidths of individual V1 neurons are not sufficiently narrow to support fine psychophysical orientation discrimination thresholds. Here we explore the possibility that V1 neurons as a population may apply optimal orientation coding strategies to achieve superb orientation tuning. We trained a self-attention deep neural network (SA-DNN) model to reconstruct a Gabor stimulus image from neuronal responses obtained through two-photon calcium imaging in five awake macaques. Each response field of view (FOV) contains 1400-1700 neurons, and their responses to a Gabor stimulus are used as the model inputs. The SA-DNN model consists of a self-attention mechanism followed by a feedforward layer. The self-attention mechanism can reveal cooperative coding by neurons activated by the Gabor stimulus, yielding attention maps that display two-way connections among neurons. The results suggest: (1) Neurons tuned to the stimulus orientation tend to have higher attention scores with all other neurons. The top 25% of orientation-tuned neurons with the highest mean attention scores can best reconstruct the stimulus images, while the bottom 50% neurons are unable to do so. (2) The responses of the top 25% neurons, after self-attention transformation, generate significantly sharpened population orientation tuning functions, with the amplitude increased by 3-5 times and bandwidth narrowed by approximately 30%. (3) After excluding the self-attention component, the forward propagation of the model would only reconstruct very coarse stimulus images. (4) The tuning sharpening displays an oblique effect: attention maps have higher variabilities at cardinal than at oblique orientations, producing more sharpened orientation tuning functions at cardinal orientations. These modeling results suggest that the self-attention mechanisms optimize orientation coding in macaque V1, reweighting responses to accentuate neurons based on attention scores. The results provide new insights into V1 neuronal connectivity, elaborating how self-attention refines neuronal interactions and reweights responses to process orientation information.

This work was supported by the National Science and Technology Innovation 2030 Major Program (2022ZD0204600)

56.459 VISUAL INPUTS RECONSTRUCTING THROUGH ENHANCED 3T FMRI DATA FROM OPTIMAL TRANSPORT GUIDED GENERATIVE ADVERSARIAL NETWORK

Yujian Xiong[1] (yxiong42@asu.edu), Wenhui Zhu[1], Yalin Wang[1], Zhong-Lin Lu[2,3]; [1]Arizona State University, [2]NYU Shanghai, [3]New York University

Unraveling the intricacies of the human visual system via the reconstruction of visual inputs from functional Magnetic Resonance Imaging (fMRI) has seen significant strides with deep learning. However, the persistent demand for high-quality, subject-specific 7-Tesla (7T) fMRI experiments poses challenges. Integrating smaller 3-Tesla (3T) datasets or accommodating subjects with short, low-quality scans remains a hurdle. Here we propose a novel framework employing an Optimal Transportation Guided Generative Adversarial Network (GAN) to enhance 3T fMRI, surmount limitations in scarce 7T data and challenges associated with short, low-quality 3T scans which have less burden for subjects. Our model, the OT Guided GAN, comprises a six-layered U-Net designed to enhance 3T fMRI scans to a quality comparable to the original 7T scans. Training is conducted across 17 subjects in two datasets with distinct experimental conditions: the 7T Natural Scenes Dataset and the 3T Natural Object Dataset. Shared input images between these datasets consist of a common set viewed by both 3T and 7T subjects, enabling an unsupervised training scenario. Subsequently, two linear regression models transform the combined set of original 7T and enhanced 3T fMRI for input into the pre-trained Stable Diffusion model, facilitating the reconstruction of visual input images. We test the framework's ability to reconstruct visual input images of natural scenes from an untrained 3T subject. The capabilities of the enhanced 3T fMRI data are demonstrated through the Fréchet Inception Distance (FID) score and human judgment, underscoring its proficiency in generating superior input visual images compared to recent methods that demand extensive 7T data. Once the framework is adequately trained, it can enhance any new subject with only 3T fMRI beyond the training set, utilizing the improved results to excel in demanding data tasks with superior performance.

The work was partially supported by NSF (DMS-1413417 \& DMS-1412722) and NIH (R01EY032125 \& R01DE030286).

WEDNESDAY MORNING POSTERS IN BANYAN BREEZEWAY

Wednesday, May 22, 8:30 am – 12:30 pm, Banyan Breezeway

Perceptual Organization: Segmentation, shapes, objects

63.301 EVERYONE KNOWS WHAT CONVEXITY IS, YET NO SINGLE CONVEXITY MEASURE CAPTURES ITS EFFECT ON FIGURE/GROUND PERCEPTION

Ömer Dağlar Tanrıkulu[1] (daglar83@gmail.com), Manish Singh[2]; [1]University of New Hampshire, [2]Rutgers University - New Brunswick

Among geometric cues to figure/ground, "convexity" is one of the strongest, with a long history in vision science. However, a literature review reveals that researchers have been using the term "convex" for decades without a unified formal definition. This results in the concept being used inconsistently across studies. Our goal here was to 1) document the various definitions used in the literature to measure convexity (contour-based/region-based; local/global); 2) run computational experiments applying these measures to stimuli from the literature (and their variants) to see how well they track the effects of "convexity" on figure-ground judgments. We started with creating graded versions of commonly-used highly "convex" stimuli and incrementally turned them into unbiased stimuli by manipulating the relative strengths of their part boundaries. For each, we computed the values of various convexity measures on each side of the contour. Finally, we compared the results of these convexity computations with figure-ground judgments from existing studies. Our results showed that certain versions of region-based definitions (i.e., percentage of pairs of dots that can "see" each other) yielded convexity values that contradict what is generally judged as "convex" in the literature. Whereas different versions of curvature-based measurements (i.e., arc-length ratio of positive-to-negative curvature, or the integral of signed curvature), as well as global region-based measures (i.e., ratio of region's area to its convex hull), yielded values consistent with the side judged as more "convex"; however, the computed magnitudes of convexity did not parallel their effects on figure-ground judgments. Our results demonstrate that no single formal measure of convexity consistently captures the strong effects of "convexity" on figure/ground perception. We suggest that these effects are closely related to the organization of visual shape in terms of parts and axes—specifically to the visual salience of the parts that result from viewing each side as potentially figural.

63.302 BORDER OWNERSHIP SIGNALS EMERGE IN AN ARTIFICIAL NEURAL NETWORK TRAINED TO PREDICT FUTURE VISUAL INPUT

Zeyuan Ye[1] (y.zeyuan@wustl.edu), Ralf Wessel[1], Tom P. Franken[1]; [1]Washington University in St. Louis

To identify the objects that surround us, the brain needs to segment visual scenes into organized collections of objects. A dominant segmentation signal in the visual cortex of the primate brain is border ownership: border ownership neurons signal which side of a border is owned by a foreground surface. Neural border ownership is known to display hysteresis with scene changes or object motion, suggesting that these neurons aid in processing dynamic visual scenes. Here we explore whether border ownership signals emerge in a neural network trained to predict future visual input. We evaluated whether units in PredNet, an artificial neural network trained to predict the next frame in natural videos, are selective for border ownership. We measured the response of units in PredNet to static scenes of an isoluminant square on an isoluminant background. We focused our analysis on units for which the classical receptive field only contained a luminance contrast border which could – in different trials – be owned by a square on one or the other side (border ownership). Scenes also varied in border orientation and luminance contrast polarity. We found that R and E units in PredNet are often selective for border ownership, irrespective of the luminance contrast polarity. The preferred side of ownership was remarkably tolerant to border orientation, similar as in our data from the non-human primate brain. The proportion of border ownership units was higher in deeper layers (L1,2,3>L0). The magnitude of border ownership signals increased with depth (L3>L2>L1>L0). Our data show that units selective for border ownership emerge in PredNet even though this network was not explicitly trained to segment visual scenes, but instead to predict the next frame in natural videos. This suggests that border ownership units in neural networks aid in efficiently predicting future input in natural videos.

This work was supported by NIH grant R00EY031795 (TPF) and the "Incubator for Transdisciplinary Futures: Toward a Synergy Between Artificial Intelligence and Neuroscience" (RW).

63.303 IMPLIED OCCLUSION AND SUBSET UNDERESTIMATION EXPLAIN THE WEAK-OUTNUMBER-STRONG NUMEROSITY ILLUSION

Eliana Dellinger[1] (dellinger.eli@gmail.com), Katelyn Becker[1], Frank Durgin[1]; [1]Swarthmore College

When equal numbers of white and gray dots are scattered against a dark gray background without overlapping, the number of white dots appears, subjectively, to be fewer than the number of gray dots (Lei & Reeves, 2018, 2022). This has been dubbed the weak-outnumber-strong numerosity illusion. Whereas previous reports of this illusion have argued that it is due to the gray dots creating a false pedestal in computing the total contrast of the white dots (thus lowering their perceived number), we report here evidence that two other principles seem to determine the effect. Four experiments will be reported. In Experiment 1, we show that subset estimation (as for the white dots) occurs equally for most subsets of different-color dots, but not for gray dots, suggesting that it is the gray dots that are primarily the cause of the illusion. In Experiments 2 and 3, we used partial occlusion or stereopsis to separate the white and gray dots into two depth planes. In both cases, the illusion is eliminated when the gray dots are in front, suggesting that it may be caused as a result of perceiving the dimmer gray dots as "filled in" behind the white dots. In Experiment 4, we re-evaluated the false floor illusion by varying the contrast of the gray dots. Our results suggest a U-shaped function, consistent with the idea that the subset underestimation of white is reduced when the gray dots are sufficiently dim, and is applied more equally to the gray dots as they are made much brighter, resulting in less relative difference at low and high gray levels. The weak-outnumber-strong illusion is visually compelling, but appears to be due to both an issue with underestimating subsets and with implied occlusion of the dimmer dots.

Swarthmore College Faculty Research Grant and Summer Grants for Research

63.304 CAN ADOLESCENTS IMITATE THE EVENT SEGMENTATION BEHAVIOR OF ADULTS BY UTILIZING COGNITIVE CONTROL?

Xianzhen Zhou[1] (xianzhen.zhou@ukdd.de), Astrid Prochnow[1], Foroogh Ghorbani[1], Christian Beste[1,2]; [1]Cognitive Neurophysiology, Department of Child and Adolescent Psychiatry, Faculty of Medicine, TU Dresden, Schubertstrasse 42, 03107 Dresden, Germany, [2]School of Psychology, Shandong Normal University, Jinan, China

In our daily lives, the continuous stream of information is segmented into distinct events, a concept outlined in Event Segmentation Theory (EST). How individuals perform event segmentation and the granularity of the resulting segmented events are likely influenced by cognitive control. Compared with adults, adolescents prefer to segment event at a coarser granularity, lacking a hierarchical gradient. In this study, our aim is to explore whether adolescents can exhibit behaviour similar to adults when instructed to employ increased cognitive control during the segmentation of a movie. We provided different instructions to two groups of healthy adolescents when they were performing event segmentation task: one group was instructed to segment a movie as finely as possible (Fine Grained Adolescent), while the other group received no specific segmentation instructions (Free Segmentation Adolescent). Additionally, one group of healthy adults segmented a movie with no specific segmentation instructions (Free Segmentation Adult). We observed that, compared to Free Segmentation (FS) Adolescent, Fine Grained (FG) Adolescent exhibit greater similarity with Free Segmentation (FS) Adults. This similarity is prominently evident in two main aspects: Regarding the mean segment length, there is no distinction between FG Adolescents and FS Adults, but both groups exhibit a significant shorter length compared to FS Adolescents. As the number of situational changes in the movie increases, the estimated probability of segmentation is similar between FG Adolescents and FS Adults. However, FS Adolescents show a significantly lower probability compared to both of them. This suggests that while adolescents may not be fully developed, following instructions allows them to employ more cognitive control and behave similarly to adults during event segmentation process.

The work was supported by a grant from the Else-Kröner-Fresenius-Stiftung [2020_EKSE.105]. We thank all participants for taking part in the study.

63.305 CONJUNCTIVE REPRESENTATION OF COLORS AND SHAPES IN HUMAN OCCIPITOTEMPORAL AND POSTERIOR PARIETAL CORTICES

Benjamin Swinchoski[1] (ben.swinchoski@yale.edu), JohnMark Taylor[2], Yaoda Xu[1]; [1]Yale University, [2]Columbia University

How does the human brain jointly represent color and shape? In contrast with the traditional view that color and form are represented by separate visual areas and bound together via selective attention, a recent study using simple artificial shape stimuli and an orthogonal luminance change task found that color and form were largely jointly encoded in the same brain regions (including regions defined by their univariate response to color or shape), albeit in an independent manner, such that a classifier trained to discriminate shapes in one color could cross-decode the same shapes in a different color. The present study aims to understand how attention may impact feature representation when complex, real-world object shapes are encoded. We used three shapes (generated from side-view silhouettes of cars, helicopters, and ships) and three colors (red, green, and blue, equated in luminance and saturation). We obtained fMRI response patterns from 12 human participants as they viewed blocks of images, with each block containing exemplars of the same object and color with slight variations in shape and hue. In different fMRI runs, participants either attended to shape, color, or both features and had to respond to repetitions in the attended feature dimension(s). Unlike in the earlier study examining simple shape features with an orthogonal task, regardless of the feature attended, we found a drop in cross-color shape decoding compared to within-color shape decoding across occipitotemporal and posterior parietal cortices. These results indicate that nonlinear conjunctive coding of shape and color exists across the human ventral and dorsal visual regions when attention is directed towards real-world object features.

This project is supported by NIH grant 1R01EY030854 to Y.X. J.T. is supported by NIH grant 1F32EY033654.

63.306 CURVATURE FORMATION IN THE VISUAL CORTEX: HOW DO WE SAMPLE?

Irfa Nisar[1] (inisar@eecs.yorku.ca), James Elder[2]; [1]Center for Vision Research,York University, [2]Department of Electrical Engineering and Computer Science, York University

Background. The circle-polygon illusion produces a polygon percept when a static dark outline circle was pulsed at 2 Hz with a luminance gradient around the inner border. We modify the method of Sakurai (2014) and display arclengths that are 1/8, ¼ , 3/8, ½ , 5/8 , ¾ , 7/8 and 1 (whole) of a circle. We test if different arc lengths, that are fractions of the circumference of the same circle, change the edge length reported. If the edge length does not change with arc length, this implies neurophysiological design that codes for curvature at specific eccentricities in the same manner. Method. Arc lengths (1/8, ¼ , 3/8, ½ , 5/8 , ¾ , 7/8 and 1 (whole)) of a circle of 4 and 8 deg and eccentricity (0,1,2,4 and 8 deg) were varied in a cross design. Observers (30 online participants; experiment was hosted on Pavlovia) indicated the edge length formed as part of the percept by selecting a edge length from an array of edge lengths shown between 0.105 deg and 2.107 deg. The perceived edge lengths were displayed as a series of twenty line segments built from the equation log(a*b c) such that a=1 and b=0.9. c belonged to the range [1,20] and each value of c gave rise to a unique line segment. Observers reported strength on a scale 1-10. Result. Edge length remained approximately uniform for all arc lengths for a fixed circle size. Strength was strongest for the shortest arc length, gradually decreasing as the arc length increased. Lowest strength was reported for whole circles. This may indicate that the early visual cortex has an affinity for closure and prefers to see shapes as composited closed curves or circles, reporting smaller edge lengths and lower strength for polygonal percepts in the process.

VISTA; York Doctoral Fellowship

63.307 NUMBER: STILL A PRIMARY VISUAL FEATURE

Caroline Myers[1] (cmyers60@jhu.edu), Chaz Firestone[1], Justin Halberda[1]; [1]Johns Hopkins University

Some of the strongest evidence that number is a primary visual feature (like color or contrast) comes from experiments demonstrating visual adaptation to number (e.g., Burr & Ross, 2008), wherein staring at a large number of dots decreases numerosity estimates of subsequent probe displays. Recently, these findings have been challenged by a deflationary account on which these effects reflect spatiotopic attenuation to unchanging information (Yousif et al., 2023). Here, we conduct a crucial comparison of these accounts by testing number adaptation for arrays whose spatial properties constantly change. Centrally fixating observers viewed two large discs subtending 9° that independently and randomly translated on either side (left or right) of the display. During the adaptation phase, varying numbers of dots appeared and faded at changing locations within each continuously-moving disc. After 12 seconds, the dots disappeared and the discs continued moving for an additional 1000ms. Following this delay, a tone signaled the appearance of a new number of probe dots, appearing for 500ms in new locations within each disc; observers judged which disc contained more dots. On critical trials, probe dots were equal in number at the time of the tone. If number is a primary visual feature that can be bound to an object, then subjects should show an adaptation effect (thereby judging the disc previously containing the smaller adapter number as greater). If number adaptation effects are really spatiotopic attenuation to unchanging information, then subjects should not show an adaptation effect. Subjects showed the predicted adaptation effect. This result suggests that number adaptation persists despite drastic changes to spatiotopic and retinotopic position, contra an explanation in which adaptation results from filtering out information that remains unchanged between adapter and probe displays. Our findings re-open the case for genuine number adaptation and numerosity as a primary visual feature more broadly.

NSF BCS #2021053 awarded to C.F.

63.308 OBJECT-BASED WARPING IN OBJECTS AND HOLES

Yujin LEE[1] (susu18@ewhain.net), Sung-Ho Kim; [1]EWHA WOMANS UNIVERSITY

While contours bounding holes (cutouts in a surface) are perceptually assigned to the surrounding surface, people can perceive the shapes of holes as well as figural objects. Reflecting this paradox, recently there has been an ongoing debate on holes' figural status. Previous studies on attention have indicated that the object-based attentional effect diminishes when an object is replaced by a hole, suggesting that holes are perceptually organized as the background. To further investigate whether the representation of a hole corresponds to the background region seen through the hole or its immaterial surface enclosed by a material object, we utilized "object-based warping" (Vickery & Chun, 2010)—an illusion where two dots within an object appear farther apart than two equidistant dots in empty space. Participants viewed 3D stereoscopic displays where reference dots, surrounded by a rectangular region—a solid object (Experiment 1) or an empty hole (Experiment 2)—were shown along with adjustment dots. They manipulated adjustment dot spacing to match that of the reference dots. We compared judgments when the dots were placed in different depth planes (i.e., on, in front of, or behind the rectangle). In Experiment 1, the warping effect was more pronounced when the dots were on the object's surface than when they were in front of it. In Experiment 2, where a rectangular hole surrounded reference dots, the warping effect was greater when reference dots were at the same depth plane as the hole (i.e., on the immaterial surface) compared to when they were both in front of the hole and on the background visible through it. These findings suggest (1) that the warping effect is surface-based and (2) that in terms of object-based attention, the representation of a hole corresponds to the immaterial surface enclosed by the contour rather than the background seen through it.

63.309 OPTIMAL NOISE CHOICE FOR REVERSE CORRELATION EXPERIMENTS DEPENDS ON THE BASE IMAGE FEATURES

Ethan Duwell[1] (eduwell@mcw.edu), Gennadiy Gurariy[1], Adam Greenberg[1]; [1]Medical College of Wisconsin

Psychophysical reverse correlation (PRC) is a versatile, data-driven method that can be used to study a wide array of sensory, perceptual, and cognitive tasks. In visual PRC experiments, subjects typically perform discriminations on a base image with white noise overlaid. After many trials, noise frames from correct and incorrect trials are combined systematically to generate classification and significance images (CIs, SIs) which highlight locations and features critical for performing a perceptual task. However, the large (5-10k) number of trials required to generate quality CIs/SIs greatly limits the practicality of PRC. We explored the possibility of improving PRC efficiency by optimizing the noise used in a series of simulations. In these experiments, a simulated observer detected the orientation of angled Gabors with a variety of different noise types overlaid including white noise, sinusoidal noise, and Gaussian blob noise. We also systematically varied trial number (1-10k) and the spatial frequency of the base image Gabor (2-18 Hz) to fully explore the optimization space. For each simulation, CI quality was assessed by its correlation with the base image, correlation with a reference CI (generated via 20,000 trials), and the number of significant pixels. PRC efficiency was assessed as the number of trials required for the resulting CIs to meet a given correlation value or number of significant pixels. Our results suggest that noise selection greatly impacts PRC efficiency, and that the optimal noise choice depends on specific features of the base image used. Highest efficiency seems to be attained when the feature profile of the noise most closely matches that of the base image. We believe that similar, tailored simulation methods could be used to improve PRC experiment efficiency by informing and optimizing the noise used prior to collecting human subject data.

Wednesday, May 22, 8:30 am – 12:30 pm, Banyan Breezeway

Spatial Vision: Image statistics, neural mechanisms

63.310 SACCADE TRANSIENTS ENABLE EFFICIENT POPULATION ENCODING OF NATURAL SCENES

Soma Mizobuchi[1,2] (soma.mizo@gmail.com), Jonathan D. Victor[3], Michele Rucci[1,2]; [1]Department of Brain and Cognitive Sciences, University of Rochester, USA, [2]Center for Visual Science, University of Rochester, USA, [3]Weill Cornell Medical College, USA

Humans use rapid eye movements (saccades) to center the high-acuity fovea on objects of interest. Saccades cause the image to shift on the retina, delivering temporal transients rich in spatial information to neurons highly sensitive to input changes. Previous research has shown that the luminance modulations resulting from saccades are stereotypically structured (Mostofi et al, 2018). The strength of luminance modulations increases with the spatial frequency of the stimulus in a way that counterbalances the power spectrum of natural scenes in a low spatial frequency band (whitening regime) and remains constant at higher frequencies (saturation regime). The bandwidth of whitening critically depends on saccade amplitude, increasing for smaller saccades. Here we examine how the space-time reformatting resulting from saccades interacts with the response sensitivity of retinal ganglion cells at a population level. We modeled spatial and temporal response properties of parvocellular (P) and magnocellular (M) neurons at various eccentricity in the macaque's retina following published neurophysiological data. Modeled neurons were exposed to the visual input signals from saccades by moving their receptive fields over natural scenes following sequences of eye movements recorded from human observers. We measured the average pairwise correlation for neurons at systematically spaced distances following saccades. As expected from the spatial characteristics of cell receptive fields, broad correlations occurred in both M and P cell populations at all eccentricities when images were presented without saccades. Saccades greatly attenuated correlations. This effect was particularly pronounced for small saccades (amplitudes less than 2 degrees) which resulted in an almost complete decorrelation of natural scenes. These results indicate that luminance modulations from saccades contribute to an efficient neural code at the retinal output.

NIH EY18363, NIH EY07977 and P30 EY001319

63.311 NEURONAL INTERACTIVITY SCALES WITH STIMULUS COMPLEXITY, NOT HIERARCHICAL STRUCTURE IN MOUSE VISUAL CORTEX

Peter G. Poggi[1,3] (gregorio.p.poggi@vanderbilt.edu), Brock M. Carlson[1,2], Blake Mitchell[1,2], Alexander Maier[1,2]; [1]Vanderbilt University, [2]Vanderbilt Vision Research Center, [3]The SyBBURE Searle Undergraduate Research Program

We evaluated how neuronal interactivity changes with visual stimulus complexity in mouse subcortical and cortical visual areas. To do this, we leveraged Integrated Information Theory (IIT4.0; Mayner et al., 2023) to derive a measure for neuronal interactivity termed PHI hat ($\hat{\varphi}$). This novel measure, which slightly deviates from IIT's measure PHI, allows us to go beyond pairwise analysis and quantify higher-order causal interactions in population spiking data to better understand neural complexity. We used an open-source dataset collected by the Allen Institute (Siegle et al., 2022). From this dataset, we used 52 Neuropixel recordings spanning six visual areas. We filtered the local field potential (LFP) data for gamma-range activity and binarized our data around the median. We then created a state-by-state transition matrix that quantified the probability of transitioning from a given state at time point t to any other possible state at time t + 1 (Leung et al., 2021). Using the open-source PyPhi toolbox (Mayner et al., 2018), we calculated 4 values of $\hat{\varphi}$ for each presentation in a system of three randomly selected channels 3 times per probe. We found that $\hat{\varphi}$ increases with visual stimulus complexity, from static gratings to natural movies across all visual areas ($p < .01$, ANOVA). However, we found no statistically significant difference in $\hat{\varphi}$ values between brain regions ($p > 0.227$, ANOVA). Even early visual areas such as the LGd and higher areas beyond V1 exhibited similar $\hat{\varphi}$ values for each type of stimulus. Both the increase in neuronal interactions as a function of stimulus complexity and the lack of clear contrast between cortical areas match a previous finding on neural differentiation (Mayner et al., 2022). Taken together, these findings suggest that $\hat{\varphi}$ is a promising measure of complexity.

The SyBBURE Searle Undergraduate Research Program

63.312 THE SPECTRAL IDEAL OBSERVER: FOCUS ERROR AND PUPIL SIZE ESTIMATION

Johannes Burge[1] (jburge@psych.upenn.edu), Wilson S. Geisler[2]; [1]University of Pennsylvania, [2]University of Texas at Austin

Perception science has long appreciated that amplitude and phase relationships carry critically important information about behaviorally-relevant properties of proximal stimuli and the distal environment, ranging from edge structure, to focus error, to fixation error (binocular disparity), to face identity. But it has not been clear how to make optimal use of phase information. Here, we report a new class of image-computable ideal observers—spectral ideal observers—that make optimal use of both amplitude and phase. They operate on the real and imaginary coefficients of a stimulus' Fourier transform—rather than on amplitude and phase, or on pixels—, by characterizing the joint probability distributions of these coefficients across frequency and the latent variable of interest. They show, quantitatively, the information phase provides over-and-above amplitude alone. Significant computational advantages are obtained. For noise images (e.g. 1/f noise), the coefficients are conditionally Gaussian; for natural images, appropriate normalization Gaussianizes them. From these distributions, the posterior over the latent variable, or the optimal Bayesian-theoretic decision variable, can be computed. The results provide principled predictions of human performance, and of the supporting neural computations. Spectral ideal observers are well-suited for problems in which the transformation from scene- to image-space is naturally modeled as a (shift-variant) convolutional operation. To demonstrate the effectiveness of spectral ideal observers, we develop one for the joint task of estimating focus error and pupil size from individual images blurred by the optics of human eyes. We show that focus error and pupil size can be accurately estimated with exquisite precision from L- and S-cone responses to individual images; performance differences can be predicted on an eye-by-eye basis. In addition to this application in optics, the spectral ideal observer should have broad application to other estimation and discrimination tasks, including binocular disparity in stereo-depth perception and binaural sound localization in audition.

This work was supported by NIH grant R01-EY028571 from the National Eye Institute & Office of Behavioral and Social Science Research to J.B..

63.313 NOISE REVERSES THE OBLIQUE EFFECT: A HORIZONTAL EFFECT IN ORIENTATION ESTIMATION AND SUBJECTIVE UNCERTAINTY

Jason Samaha[1]; [1]UC Santa Cruz

The oblique effect refers to the classic finding that sensitivity to orientation information is worse for oblique (i.e., diagonal) compared to cardinal (i.e., horizontal and vertical) orientations. This textbook phenomenon is often taken to demonstrate our brain's adaptive coding principles in response to environmental regularities. The oblique effect, however, is typically studied with stimuli that contain very narrow-band orientation content (such as lines or grating stimuli) which contrasts with most natural scene statistics, which are more broadband in nature (i.e., they contain a variety of orientation content). Surprisingly, several papers have found a reversal of the oblique effect using detection tasks and stimuli filtered to include more broadband orientation content, a finding dubbed the horizontal effect (since detection is usually worse for horizontal, followed by vertical, then oblique stimuli). Here we show that simply adding pixel noise to a grating stimulus is sufficient to induce the horizontal effect. In a large sample, using an orientation estimation task combined with subjective uncertainty judgments, we found a clear increase in both errors and uncertainty ratings for horizontal stimuli and a relative reduction of error and uncertainty for oblique stimuli, in direct contrast to the standard oblique effect. We also found that increasing the contrast of the noise in the stimulus seems to induce a larger orientation-related modulation of errors. Pixel noise is a common manipulation used to degrade stimuli yet our results suggest that doing so also fundamentally alters the sensitivity to orientation content in the image. Moreover, the dependency of the horizontal effect on noise contrast could suggest an explanation of the phenomenon in terms of divisive normalization.

63.314 INTERNAL NOISE AND EFFICIENCY UNDERLIE PERFORMANCE ASYMMETRIES THROUGHOUT THE VISUAL FIELD

Shutian Xue[1] (shutian.xue@nyu.edu), Antoine Barbot[1], Qingyuan Chen[1], Marisa Carrasco[1]; [1]New York University

[Background] Visual performance peaks at the fovea and decreases with eccentricity and differs around polar angle: it is better along the horizontal meridian (HM) than the vertical meridian (VM) and at the lower than upper VM. Whether performance differences throughout the visual field are due to differential ability to extract task-relevant information in a noisy environment remains unknown. Here, we investigate how two factors limiting performance –internal noise (amount of internal variability in the system) and efficiency (ability to extract information from the target)– underlie performance differences throughout the visual field. [Method] We used an equivalent noise method, which maps contrast threshold as a function of noise contrast, to estimate and disentangle internal noise and efficiency at the fovea and around polar angle (left & right HM and upper & lower VM) in parafovea (4°) and perifovea (8°) . At each location, observers discriminated the orientation of a 5 cpd Gabor (±45º off the vertical axis) embedded in dynamic white noise. We obtained contrast thresholds at 7 noise contrast levels at each location. [Results] We found that (1) Internal noise increased with eccentricity and was lower at HM than VM in the parafovea, but did not differ around polar angles in perifovea; (2) Efficiency (a) was higher at the fovea and parafovea than perifovea; (b) did not differ around polar angle in the parafovea, and (c) was higher on the HM than VM, and at the lower- than the upper-VM in the perifovea. [Conclusion] Distinct computations limit performance throughout the visual field: Internal noise primarily underlies eccentricity differences, consistent with variation in the distribution of the cortical tissue. Efficiency primarily underlies polar angle differences and more so at perifovea than parafovea, consistent with variation in tuning properties.

Funding: NIH R01-EY027401 to M.C.

63.315 A HIGH-RESOLUTION STUDY OF POSITIVE AND NEGATIVE RETINOTOPIC CODES IN THE HIPPOCAMPUS

Peter A. Angeli[1] (peter.a.angeli@dartmouth.edu), Adam Steel[1], Edward H. Silson[2], Caroline E. Robertson[1]; [1]Dartmouth College, [2]University of Edinburgh

Position-dependent activity in response to stimulation of the retina, or retinotopic coding, has long been considered a unique feature of the visual system. However, recent work has shown that higher-order cortical areas, including the default network, show retinotopic sensitivity in the form of decreasing signal during stimulation of their visual receptive field (Szinte and Knapen 2020; Klink et al., 2021). Motivated by this negative coding's possible relevance to perception-memory interactions (Steel*, Silson* et al. 2023), here we characterize positive and negative retinotopic responses to visual stimuli in the hippocampus. Specifically, building on recent observations of retinotopic sensitivity in the hippocampus (Silson et al., 2021) using population receptive field (pRF) mapping, we explored this region's positive and negative retinotopic responses using high-resolution (7T, 1.8mm isotropic) pRF mapping data from the Natural Scenes Dataset (Allen et al. 2022). We found that the hippocampus contained a relatively high proportion of negative pRFs, consistent with our prediction. On average across participants, 42% of hippocampal voxels exhibited a significant ($R2 \geq 0.08$) retinotopic response, and 43% of these voxels were negative in valence. Negative voxels had similar pRF sizes to their positive counterparts. Consistent with prior work, positive pRFs showed a significant contralateral visual field bias (i.e., left hippocampus tended to represent the right visual field) ($p < 0.05$), while negative pRFs were more foveal and did not evidence the contralateral bias. Interestingly, resting state functional connectivity analyses suggest positive and negative hippocampal pRFs are differentially coupled to cerebral cortex, co-fluctuating more strongly with congruently signed cerebral pRFs. This work suggests the importance of visual coding in structuring the interaction between the hippocampus and cerebral cortex, and adds support for the view that negative pRFs may play an important role in hippocampally-dependent cognitive processes like episodic memory and scene construction.

Supported by NIH R01 MH130529

63.316 MACULAR PIGMENT AND CONTRAST SENSITIVITY IN HEALTHY ADULTS

Jeffrey Nightingale[1], Colin Gardner[1], Lisa Renzi-Hammond[2], Billy R. Hammond[1]; [1]Vision Sciences Laboratory, University of Georgia; Athens, GA, USA, [2]Institute of Gerontology, College of Public Health, University of Georgia; Athens, GA, USA

Introduction: Some past data has shown that macular pigment (MP) is related to improved contrast sensitivity (CS) under certain conditions. The mechanism for this improvement, however, is unclear. The question of whether higher MP improves CS under carefully controlled optical conditions (e.g., similar to refractive correction) is not known. Methods: 22 participants (M=28.45 +/-10.36 years; 68% female; BCVA 20:30 or better) provided photopic contrast sensitivity functions (from measurements at 1.6, 3.2, 8, 16, and 24 cpd, with spatial frequency presentation randomized) using a simple optical device. The optical system used 520 nm lasers (not absorbed by MP) that were made Lambertian using two integrating spheres with a 3.5-deg circular exit port. These beams were combined with a beam splitter that allowed constant measurement of light output and contrast modulation using sine-wave gratings on glass. MP optical density (OD) was measured using customized heterochromatic flicker photometery (peak 460 nm, 1-degree test stimulus). Results: MPOD was not significantly related to CS at any measured frequency (range in correlation p-values, 0.92-0.41). Conclusions: The general finding that MP increases CS under glare conditions is reasonable: by reducing a veiling glare source (with significant SW energy), the visibility of a target would be increased. A pure filtering mechanism makes less sense in the absence of a glare source. If MP reduces the luminance of a grating target, CS would also be expected to be reduced (CS is positively related to luminance). The results of this study suggest that, under carefully controlled optical conditions, MP is not related to CSF in young healthy subjects. Interventions that have found a relation, therefore, are likely achieving this change through some other mechanism, either direct and physiological or through some health benefit (e.g., reduced oxidative/inflammatory stress).

63.317 ARE VISUAL DEFICITS IN AMBLYOPIA DRIVEN BY SPATIAL FREQUENCY TUNING?

Emily Wiecek[1,2,3] (emily.wiecek@childrens.harvard.edu), Luis D. Ramirez[3], Michaela Klimova[4], Sam Ling[3]; [1]Boston Children's Hospital, [2]Harvard Medical School, [3]Boston University, [4]Northeastern University

Amblyopia is characterized by visual deficits at high spatial frequencies. Prior work has demonstrated individuals with amblyopia have larger population receptive field (pRF) sizes in early visual areas, but how this increased RF size is linked to changes in spatial frequency tuning has yet to be determined. Model-based functional neuroimaging provides a method to map and interrogate how population spatial frequency tuning (pSFT) may differ in individuals with amblyopia. Here, we compared pSFT estimates to pRF maps in early visual cortex (V1–V3) in individuals with amblyopia and normally-sighted control participants. Monocular voxel-wise pRF and pSFT estimates were acquired in separate scanning sessions. We then compared pRF and pSFT estimates between eyes (amblyopic/non-dominant vs fellow/dominant eye) and between subject groups (amblyopia vs control). Across all individuals, RF size increased with eccentricity, and amblyopes had larger RF sizes compared to controls ($p < 0.001$), supporting previous findings. Moreover, spatial frequency preferences were higher and broader in amblyopes across V1–V3, and the amblyopic eye had broader bandwidths compared to the fellow eye. To compare pSFT and pRF estimates, we calculated a Scale Invariance metric, taking a ratio of preferred spatial frequency over pRF size (cycles per RF). The amblyopic eye showed fewer cycles per RF in central vision compared to the fellow eye and normally-sighted controls, consistent with deficits in the sampling of SF information in amblyopia. Together, our data suggest that differences in both RF size and spatial frequency tuning contribute to visual deficits observed in individuals with amblyopia.

This work was funded by the National Institutes of Health (NIH) Grant K23EY034212 to E.Wiecek, R01EY028163 to S.Ling, and F99NS124144 to L.Ramirez

63.318 HOW DOES EMOTIONAL AROUSAL MODULATE POPULATION SPATIAL FREQUENCY TUNING?

Luis D. Ramirez[1], Jasmine Pan[1], Sam Ling[1]; [1]Boston University

Emotional arousal guides cognition and visual perception to improve the detection of threats and opportunities, which can be critical in poor visual conditions. One way emotional arousal can influence vision is by altering spatial frequency (SF) processing, altering the scale of visual detail encoded by the visual system. Previous studies investigating the interaction between emotion and SF processing report an impairment in the processing of higher spatial frequencies (HSFs), while prioritizing lower spatial frequencies (LSFs) under fearful conditions (relative to a baseline or neutral condition). However, there is scant neural evidence supporting these changes in SF perception with emotion. To address this, we leveraged a population spatial frequency tuning mapping (pSFT) approach with fMRI to test how emotionally arousing auditory stimuli (Neutral vs. Negative) modulate SF processing in early visual cortex (V1–V3). If emotional arousal modulates SF processing, we expect negative/fearful stimuli to trigger the prioritization of LSFs at the cost of HSFs, which would be reflected in SF preference-specific changes in pSFT. Indeed, we found that emotional auditory stimuli profoundly altered SF processing: HSF-preferring populations experienced a decrease in preferred SF and increased tuning bandwidth, while LSF-preferring populations experienced an increase in preferred SF and decreased bandwidth. While our physiological markers of arousal reveal greater pupil diameter and skin conductance with Negative sounds vs. Neutral sounds, the observed changes in pSFT were found in both the Neutral and Negative conditions, suggesting that the processing of emotional auditory stimuli differentially impacts SF processing in early visual cortex.

This research was funded by National Institutes of Health Grant EY028163 to S. Ling and supported by F99NS124144 to L.D. Ramirez.

63.319 NONLINEAR TRANSFORMATION FROM THRESHOLD TO SUPRATHRESHOLD TEXTURE SIMILARITY JUDGMENTS

Jonathan Victor[1] (jdvicto@med.cornell.edu), Suniyya Waraich[1], Mary Conte[1]; [1]Weill Cornell Medical College

Suprathreshold similarity judgments can depart substantially from predictions based on thresholds. For example, judgment of suprathreshold grating contrast is more nearly veridical than predicted by the inverted-U threshold contrast sensitivity function. However, it is unclear whether the shift from threshold to suprathreshold judgments is merely one of channel-specific gain changes. To address this, we compared suprathreshold similarity judgments and thresholds in a well-characterized multidimensional domain of visual textures (Victor and Conte 2015). Suprathreshold similarity judgments were obtained with the paradigm of Waraich and Victor (2022). On each trial, subjects (N=4) ranked eight comparison stimuli in order of similarity to a central reference. Via a variant of multidimensional scaling, rankings across 1000 unique trials and 25 stimuli were used to build a geometric model of a perceptual space that accounted for the similarity judgments, which we compared to the perceptual space inferred from threshold sensitivities for the same stimuli. We found two kinds of differences. First, high-order image statistics contributed disproportionately to suprathreshold distances. That is, texture pairs that were predicted to be equally distant based on threshold sensitivities were perceived as more dissimilar if they differed in higher-order statistics (crossings and corners), vs. lower-order ones (luminance and edges). This finding implies selective gain increases for higher-order statistics. Second, coordinate axes in the threshold perceptual space became sharply bent in the suprathreshold space. Gain changes alone could not do this: with just gain changes, straight lines in the threshold space would map to straight lines in the suprathreshold space. Instead, the observed distortion implies that positive and negative values of the same image statistic, which are treated similarly in the threshold space, undergo separate transformations when used for similarity judgments. In sum, channel-specific gain changes and rectification-like nonlinearities are needed to account for the shift from threshold sensitivities to suprathreshold similarity.

NIH EY07977, Fred Plum Fellowship in Systems Neurology and Neuroscience

63.320 INFLUENCE OF LUMINANCE POLARITY CONGRUENCE ON NUMEROSITY ADAPTATION

Tomoki J Takahashi[1,2] (uytrewq4223@gmail.com), Masamichi J Hayashi[1,2]; [1]Graduate School of Frontier Biosciences, Osaka University, Suita, Japan, [2]Center for Information and Neural Networks (CiNet), Advanced ICT Research Institute, National Institute of Information and Communications Technology, Suita, Japan

Humans and animals have the ability to estimate approximate numerosity without explicitly counting individual elements. One psychophysical evidence supporting the existence of numerosity-tuned neural mechanisms is the susceptibility of numerosity perception to adaptation. Although conventional studies often use dots of opposite luminance polarity (e.g., half black and half white, presented on a gray background) to control the overall luminance of the display, it remains unclear whether numerosity perception is bound to such low-level visual features. To address this question, we conducted a series of numerosity adaptation experiments. In our first experiment, observers were adapted to either 20 or 80 dots and responded whether the subsequent test numerosity, ranging from 24 to 72 dots, was larger or smaller than the reference numerosity of 40 dots. The congruence of luminance polarity between the adapting and test dot arrays was manipulated. The results showed significant negative aftereffects in both congruent and incongruent conditions, although weaker in the incongruent condition. To minimize the possible effect of retinal afterimages, our second experiment introduced white noise masking stimuli after the adaptation and test stimuli, yielding results consistent with the first experiment. Subjective reports of the frequency and intensity of afterimages did not correlate with the individual differences in the aftereffect size, confirming the little influence of the afterimage. Finally, we found that randomly switching the luminance polarity of the test stimuli across trials did not eliminate aftereffects, suggesting that the predictability of luminance polarity is not relevant for the presence of aftereffects in the incongruent condition. Taken together, our results show that the negative aftereffects in numerosity adaptation are robust to the changes in luminance polarity. Nevertheless, aftereffect magnitude is still affected by the changes in luminance polarity, suggesting that numerosity perception may not be completely independent of the processing of such low-level visual features.

Wednesday, May 22, 8:30 am – 12:30 pm, Banyan Breezeway

Motion: Neural mechanisms

63.321 IDENTIFYING FUNCTIONAL SUBDIVISIONS OF THE HUMAN MT COMPLEX

Puti Wen[1] (pw1246@nyu.edu), Rania Ezzo[1], Michael Landy[2], Bas Rokers[1,2]; [1]Psychology, New York University Abu Dhabi, [2]Psychology and Center for Neural Science, New York University

The motion-sensitive MT complex (MT+) in macaque monkey contains several functional subdivisions, including areas MT, MST and FST. Previous neuroimaging work has established functional localizers for the human analogue of areas MT and MST (Huk, Dougherty, & Heeger, 2002), but not FST. We employed a multifaceted approach, incorporating functional localizers, retinotopic organization, and myeloarchitecture to separate FST from MT and MST. We first used a traditional localizer to identify MT+ (radial and tangential 2D moving versus static dots). Considering FST's sensitivity to 3D motion and its robustness to opponent motion in monkey, we then explored disparity-defined stereomotion (coherent versus temporally scrambled) and opponent motion (paired versus unpaired counter-phase oscillating dots) localizers. We also compared our results to myelin maps and retinotopic organization within the same participants. We consistently localized putative FST (pFST) within the ventral section of MT+. We found weaker responses to 2D motion, but stronger responses to stereomotion in pFST relative to MT/MST. Our results contrast with a previous report that identified an area outside and anterior to MT+ (Likova & Tyler, 2007). 2D motion and opponent motion localizers generally identified the same areas, suggesting a more suppressed response to locally opponent motion in MT/MST than in pFST. Retinotopic mapping aids in distinguishing MT and MST but is less effective for pFST. The MT+ foveal confluence was identifiable in all subjects, but the central visual field was overrepresented in pFST. Myelin mapping showed significant individual and hemispheric variability, though on average, converged at the border of MT and MST. Finally, in all participants, we observed mismatches between functionally and atlas-defined MT, MST, and pFST. Our results demonstrate the importance of functional localization for delineating subdivisions of human MT+.

NYUAD Center for Brain and Health, funded by Tamkeen under grant CG012, ASPIRE Precision Medicine Research Institute Abu Dhabi (ASPIREPMRIAD) award grant number VRI-20-10, NIH EY08266 (MSL)

63.322 INVESTIGATING THE DEVELOPMENT OF THE VISUAL MOTION PATHWAY AT THE NETWORK LEVEL

Brandon R. Nanfito[1,2,3] (bnanfit1@jh.edu), Kristina J. Nielsen[1,2,3]; [1]Johns Hopkins School of Medicine, [2]Zanvyl Krieger Mind/Brain Institute, [3]Kavli Neuroscience Discovery Institute

Interactions between visual areas clearly play an important role in the processing of visual information in adult animals. At the same time, we know little about how these interactions shape development. This question is especially important for the visual motion pathway, as anatomical data from primates show a very early development of higher motion area MT. Here, we use the ferret's motion pathway – in particular V1 and area PSS, an MT homolog – to investigate this issue. In a first set of experiments, we quantify the contribution of either area on the other at different ages by combining simultaneous extracellular recordings in V1 and PSS with reversible inactivation of either area by cortical cooling. These data so far show that PSS in visually naïve animals, but not in adults, retains responsiveness and tuning despite V1 inactivation. In a second set of experiments, we then directly test how development of one area depends on input from the other. To this end, we use an established training paradigm, in which visually naïve animals are exposed to drifting gratings for 8 hr, in conjunction with cooling. Normally, this paradigm induces rapid maturation of direction selectivity in V1 and PSS. Our preliminary data show that if V1 is cooled during the training, PSS still develops direction selectivity, but V1 does not. Finally, using retrograde tracer injections we demonstrate that the lateral geniculate nucleus in the thalamus sends a much stronger projection to PSS in visually naïve than in mature animals. All these findings contradict the basic assumption of a feedforward, sequential development that propagates forward from V1, and highlight the need to study development at the network level. This research is essential for understanding the etiology of perceptual symptoms associated with disorders of development or abnormal visual experience in early life.

NIH, Kavli NDI

63.323 EXPLORING THE IMPACT OF APPARENT MOTION IN V1 RETINOTOPIC REPRESENTATIONS AND BEHAVIORAL CORRELATES IN PRIMATES

Salvatore Giancani[1] (salvatore.giancani@univ-amu.fr), Mohit Srivastava[1], Kevin Blaize[1], Sandrine Chemla[1], Matteo Di Volo[2], Anna

Montagnini[1], Frederic Chavane[1]; [1]Institut de Neurosciences de la Timone, Unitée Mixte de Recherche 7289 Centre National de la Recherche Scientifique and Aix-Marseille Université, Faculty of Medicine 27, Boulevard Jean Moulin, 13385 Marseille Cedex 05, France, [2]Université Claude Bernard Lyon 1, Institut National de la Santé et de la Recherche Medicale, Stem Cell and Brain Research Institute U1208, Bron, France

It is still poorly understood how the visual system processes a simple stimulus moving along a trajectory. Any local, static, stimulus generates waves propagating in V1 retinotopic maps (Muller et al 2014, 2018). As a consequence, when these local stimuli are presented in sequence, in different positions in space and/or time, they trigger a complex cascade of embedded intra-cortical propagation waves that can shape the representation of the individual stimulus (Reynaud 2012, Chemla 2019). In this study, we employed a spatio-temporal sequence of three strokes eliciting long-range apparent motion (AM) and we measured the dynamic response of the V1 neural population in behaving macaques using voltage-sensitive dye imaging (VSDI). Expanding upon the findings of Chemla (2019), our results demonstrate that, after the first two dots appear, the spatial profile of the response to the third dot is significantly modified compared to the single-stroke control. This modification involves the facilitatory activation of the cortex ahead of motion direction and suppression leading to the displacement of the peak of activity in the opposite direction. Complex non-linear dynamical interactions are therefore changing the representation of the individual strokes suggesting at the same time motion extrapolation and motion repulsion in the retinotopic maps of V1. To explore this paradoxical phenomenon more comprehensively, we complemented our observation with (i) computational modeling approach, to investigate whether intra-cortical propagation of excitatory and inhibitory activity can explain this dual effect, and with (ii) psychophysics in humans, to test how such stimulus affect the perceived position of the last stroke of the same apparent motion stimulus.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N° 956669

63.324 DIRECTIONAL ASYMMETRIES FOR VISUAL MOTION MEASURED WITH FMRI

Rania Ezzo[1,2] (rje257@nyu.edu), Marisa Carrasco[1,3], Jonathan Winawer[1,3], Bas Rokers[1,2,3]; [1]Department of Psychology, New York University, [2]Psychology, New York University Abu Dhabi, [3]Center for Neural Science, New York University

Background. Humans are more sensitive to certain motion directions than others. We previously reported three asymmetries that combine to predict motion direction discrimination around the visual field: a large advantage for cardinal over oblique directions; a moderate advantage for radial and tangential directions over other directions; and a small advantage for radial over tangential directions. Here, we assessed corresponding neural asymmetries by comparing BOLD amplitude for various directions of motion across the visual hierarchy. Methods. In an fMRI protocol, observers (n=10) viewed large-field gratings within a stationary circular aperture (12.2° radius, centered at fixation). The gratings were either static (horizontal, vertical, oblique orientations) or drifting (1 of 8 motion directions along the cardinal and oblique meridians). Regions-of-interest (ROIs) were defined based on retinotopic mapping procedures and a motion localizer, including: V1, V2, V3, hV4, hMT+. Response amplitudes were computed by a GLM, and motion-selective responses were quantified as the difference in responses for each drift direction and its corresponding static condition. Each neural asymmetry was then quantified for each polar angle bin (45° width) and for each ROI. Results. Areas V1, V2, V3, hV4, and hMT+ had greater motion-related BOLD amplitude for cardinal than oblique directions. This cardinal advantage occurred across all polar angle bins. Evidence for an effect of radial and tangential motion over other directions was less clear. Finally, on average, we found greater responses to radial than tangential motion directions in V2 and V3v, but this effect was less consistent across polar angles and across individuals. Conclusion. The neural asymmetry between cardinal and oblique motion directions is consistent across individuals, several ROIs and polar angles, while differences between radial and tangential motion directions occur in select regions, and is more variable across polar angles. This pattern of results broadly aligns with our prior behavioral results.

(1) US NIH National Eye Institute R01-EY027401 to MC and JW, (2) NYUAD Center for Brain and Health, funded by Tamkeen NYU Abu Dhabi Research Institute #CG012 to BR, (3) ASPIRE Precision Medicine Research Institute Abu Dhabi (ASPIREPMRIAD) #VRI-20-10 to BR and (4) Global PhD Fellowship, NYUAD to RE.

63.325 MOTION ADAPTATION INDUCED OBJECT POSITION BIAS IN MACAQUE IT AND SLOWFAST VIDEO RECOGNITION MODELS

Elizaveta Yakubovskaya[1] (eyakub@my.yorku.ca), Hamidreza Ramezanpour[1], Sara Djambazovska[1], Kohitij Kar[1]; [1]Department of Biology, York University

To efficiently interact with their environment, primates excel in not just recognizing objects ('what') but also discerning their spatial attributes ('where'). This dual capacity, traditionally attributed to the functional segregation of ventral and dorsal visual processing pathways, is currently being reexamined in light of emerging evidence. Recent work of Hong et al. (2016) revealed the macaque inferior temporal (IT) cortex's role in encoding object positions. Our study further ventures into this relatively uncharted territory with three main objectives: firstly, to extend the findings of Hong et al., assessing how scaling of neural recording sites in IT influences object-position estimates; secondly, to investigate the impact of motion adaptation (a phenomenon typically associated with dorsal stream) on these estimates; and thirdly, to evaluate whether existing ventral stream models align with our observations. We performed large-scale recordings across IT cortex of 3 monkeys (~500 sites). Monkeys passively fixated Test images (640; 1 of 8 objects, varying latent parameters, embedded in naturalistic backgrounds). Indeed, we observed highly accurate (Pearson R >0.7) IT-population based linear decodes of object positions. Next, to test whether motion-direction adaptation biases position estimates, we preceded the Test image presentation by prolonged (3000 ms) oriented gratings moving in one of four directions. Remarkably, IT-based (192 sites) position decodes showed a significant bias ($p<0.0001$; permutation test) in the direction opposite to the preceding motion. These biases align with perceptual reports, suggesting that the IT cortex represents perceptual rather than ground-truth positions. Interestingly, simulating the experiments in-

silico on SlowFast networks (video recognition model with ResNet-50 backbone) demonstrated a similar bias (absent in vanilla ResNet-50 with scaled activation mimicking neural fatigue). Our findings introduce a framework for probing how dorsal-ventral interactions could generate adaptation after-effects and a model-based hypotheses space to guide the exploration of computational mechanisms critical for dynamic scene perception.

Canada Research Chair Program, Google Research, CFREF, Brain Canada, SFARI

63.326 THE ALPHA-BAND OF THE EEG MODULATES THE PERCEIVED LOCATION OF MOVING TARGETS

Andrew Zavala[1] (azavala2@uoregon.edu), Steven Shofner[1], Paul Dassonville[1]; [1]University of Oregon

The alpha rhythm of the EEG has been reported to gate the flow of sensory information in the visual system by causing fluctuating states of excitation and inhibition, with processing occurring more quickly during periods of excitation. This fluctuation in excitability presumably modulates perceptual outcomes, such as stimulus detection and perceived timing (VanRullen, 2016). If this is true, one would also expect that the perceived location of moving targets would be similarly modulated, since faster rates of processing during periods of excitation would cause the target to be perceived at an earlier location in its trajectory. In this study, we tested this hypothesis by asking participants to report the perceived starting locations of moving targets, with target onsets timed to occur at different phases of the alpha rhythm after entrainment to 10 Hz by a rhythmic auditory click. The behavioral time series of localization accuracies was analyzed with a Fast-Fourier Transform (FFT) across subjects after bootstrapping. A significant 10Hz frequency in the FFT confirmed the entrainment's modulation of participants' perceptual processing of the target's location. These findings reveal a pivotal role of alpha oscillations on the localization of moving targets, enhancing our understanding of the role played by neural rhythms in visual perception and providing a possible explanation for several mislocalization phenomena, such as the Frohlich effect.

63.327 VISUAL MOTION EXTRAPOLATION OF MOVING OBJECTS DRIVES REAL-TIME TEMPORAL RE-ALIGNMENT ACROSS HIERARCHICAL NEURAL POSITION REPRESENTATIONS

William Turner[1,2] (w6.turner@qut.edu.au), Charlie Sexton[2], Philippa Johnson[2,3], Ella Wilson[2], Hinze Hogendoorn[1,2]; [1]Queensland University of Technology, [2]The University of Melbourne, [3]Leiden University

There is mounting evidence that the brain uses predictive mechanisms to encode the position of moving objects. Recent theoretical work has shown that for hierarchical networks (e.g., the visual system) to accurately represent the position of moving objects under neural delays, predictions must be generated along both forwards and backwards pathways. An important consequence of this is the alignment of position representations across network layers. However, empirical evidence of this occurring for position representations in the human brain is lacking. In this study, we investigated the temporal dynamics of object-position representations over human visual cortex during smooth motion processing. Participants (N = 18, 2 sessions) viewed a stimulus moving along a circular trajectory, while EEG was recorded. Using multi-class LDA classification, we constructed high-resolution probabilistic maps of the stimulus' location over time. These revealed clear evidence of 'representational overshoot' following the unexpected disappearance or reversal of the stimulus, indicative of predictive position encoding. Importantly, examining neural dynamics immediately following motion onset we found evidence of rapid temporal re-alignment occurring between distinct object-position representations. We show that a similar temporal shift emerges spontaneously at all layers of a simulated neural network via spike-timing-dependent plasticity – providing a simple account of this predictive effect. Ultimately, this study sheds light on the neural encoding of moving-object position, and constitutes the first empirical evidence of predictive temporal re-alignment occurring in the human brain.

This work was supported by Australian Research Council Grants FT200100246, DP220101166, and DP180102268 awarded to HH, as well as a QUT ECRIS grant and Unimelb DSH Research Support Scheme grant awarded to WT.

Wednesday, May 22, 8:30 am – 12:30 pm, Banyan Breezeway

Eye Movements: Perception, cognition and memory

63.329 SIMILARITY OF MEMORY REPRESENTATIONS MODULATE SACCADE CURVATURES

Golnaz Forouzandehfar[1] (gforouz@vols.utk.edu), Chengrui Li[2], Aaron T. Buss[1], A. Caglar Tas[1]; [1]University of Tennessee, Knoxville, [2]School of Computational Science & Engineering, Georgia Institute of Technology

The present study investigated the effects of spatial proximity and memory similarity on saccade curvatures. Participants were asked to learn a feature-space association to guide their saccades (Wifall et al., 2017) (e.g., for red, saccade to the leftmost object). In the learning block, participants were shown four colored circles 8dva above the fixation. The fixation object changed color, and participants were to execute a saccade to the circle that matched that color. After the learning block, they were asked to complete the same task with four black circles. That is, the task required participants to make a memory-based target selection. We manipulated both spatial proximity of the circles and similarity of the colors associated with the circles. In the close location condition (CL), the separation between adjacent squares was ~2.5dva while it was ~7dva in the far condition (FL). In the similar color condition (SC), the colors for adjacent squares were 30° apart while in the dissimilar color condition (DC), they were 90° apart in HSV color space. Participants were randomly assigned to one of the four conditions: SC-CL, SC-FL, DC-CL, DC-FL. Based on spatial proximity effects in previous studies, we expected to find greater curvature in the CL conditions than in the FL conditions. If memory representations can affect saccade curvatures similarly as perceptual representations, then we also predicted larger curvatures in the SC conditions compared to the DC conditions (Mulckhuyse et al., 2009). We found a significant effect of color (p=.006) with similar colors resulting in significantly larger curvatures than dissimilar colors, but this effect was only present when squares were spatially close. The effect of spatial proximity was not significant (p>.05). Importantly, our results show that oculomotor control is not only influenced by similarity in perceptual representations, but also similarity in memory representations.

63.330 TRANSSACCADIC FEATURE MEMORY PERFORMANCE ASYMMETRIES IN VISUAL FIELD AND SACCADE DIRECTION

Jessica L. Parker[1] (jparke87@vols.utk.edu), A. Caglar Tas[1]; [1]University of Tennessee-Knoxville

Previous work examining transsaccadic memory for spatial information has shown better memory for stimuli in the right visual field compared to the left visual field, for leftward compared to rightward saccades, and information that was remapped within the same visual field rather than between visual fields (Moussaoui et al., 2023). These findings may be explained by asymmetrical activation, showing left hemisphere activation for right visual field and right hemisphere activation for both visual fields (Sheremata et al., 2010). Because feature information is consulted for transsaccadic updating of spatial information and the establishment of visual stability (Tas & Parker, 2023), the current study examines whether feature and spatial information are similarly represented in transsaccadic memory. In the present study, participants fixated on a dot in the center of the screen surrounded by two squares in the right visual field and two in the left visual field. Three squares were then filled with different colors and then masked. Next, participants executed either a leftward or a rightward saccade which manipulated whether color memory information remained in the same visual field or switched visual fields. At the end of each trial, participants reported the color of a cued memory location on a color wheel. Expanding previous work, we found better color memory when the color was presented in the right visual field compared to the left visual field and for leftward saccades compared to rightward saccades. Contradicting previous work, color memory was better for items that are remapped between rather than within the same visual field, likely due to differences in eccentricities. These findings suggest feature information, like spatial information, shows stronger representation in the right visual field and for leftward saccades, indicating similar representation in transsaccadic memory.

63.331 AUDITORY AND VISUAL WORKING MEMORY LOADS REDUCE FIXATION SPREAD DURING FREE SCENE-VIEWING

Kaito Makishima[1] (kmakipsy@gmail.com), Jun-Ichiro Kawahara[1]; [1]Hokkaido University

Retaining information in working memory (WM) interferes with eye movements. Specifically, the presence of visual (color patches) and verbal (alphabet letters) WM loads reduced fixation spread during the free viewing of scenes (Cronin et al., 2020). In the present study, we examined whether auditorily or visually presented verbal WM loads similarly interfere with the free viewing of scenes. If the visual components of WM are critical to the reduction of fixation spread, no such reduction would occur with auditorily presented WM load. A sequence of seven letters was presented visually in the center of the monitor (Experiment 1, N=48) or auditorily through headphones (Experiment 2, N=48) as memory items, followed by a scene image during which eye movements were recorded. Under the load condition, participants maintained the letter sequence during scene-viewing and were tested by a memory probe at the end of each trial. In contrast, under the no-load condition, participants merely viewed the scene without any probe. The results of Experiment 1 indicated that the scan path length and the dispersion of fixations from the center under the load condition were smaller than those under the no-load condition. The same pattern of results was found when the memory items were presented auditorily in Experiment 2. These findings demonstrate that verbal WM load reduced fixation spread during the free viewing of scenes, regardless of the presentation modality. The reduction of fixation spread found in Cronin et al. (2020) was not specific to the presence of the visual component in the memory material. Rather, the present results suggest that modality-independent WM load restricts eye movement control, resulting in the shrinkage of the fixation spread.

63.332 A FUNCTIONAL ROLE OF VISUAL WORKING MEMORY-RELATED SACCADE BIASES?

Patrik Polgári[1] (patrik.polgari@uni-marburg.de), Alexander C. Schütz[1]; [1]University of Marburg, Germany

Visual working memory (VWM) can influence oculomotor control. Hollingworth & Luck (2009) showed that corrective saccades can be biased by a memorized feature: attention and gaze shift away from the original target and towards a distractor matching the content of VWM. This may simply reflect a poor separation of different mechanisms. Alternatively, could VWM-related ocular biases have a functional role? The attentional shift coupled with saccade biases may refresh the content of VWM by updating it to the last fixated item. In this case, the magnitude of saccade correction biases also might depend on VWM task demands. 20 observers had to memorize a color hue that was tested against another hue from the same (Difficult condition) or a different color category (Easy condition). Prior to the test phase, they performed a saccade task where the position of the target changed in a gaze-contingent manner. Concurrently, the color of a neighboring distractor changed (or not) to a color (dis)similar to the one maintained in VWM. VWM performance was higher in the Easy than in the Difficult condition, indicating a successful manipulation of task difficulty. Independently of task difficulty, saccade corrections were biased more towards the distractor when it matched VWM content compared to when it did not, replicating previous findings. Most importantly for our research question, saccade corrections towards the distractor were more frequent in the Difficult compared to the Easy condition, indicating an influence of task difficulty on the attentional bias. Furthermore, VWM performance in the Easy condition was higher when saccades were biased towards the distractor rather than towards the target, suggesting a role of corrective saccades in the updating of VWM. Overall, our results indicate a potential functional role of VWM-related saccade biases since they increase with task difficulty and can facilitate VWM performance.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project Number 222641018 - SFB/TRR 135 TP B2.

63.333 INHIBITION OF RETURN DEPENDS ON IMAGE CATEGORY

Tobias Thomas[1,2] (tobias.thomas@tu-darmstadt.de), David Pietras[1], Constantin A. Rothkopf[1,2]; [1]Technische Universität Darmstadt, [2]Hessian Center for Artificial Intelligence (hessian.AI)

Humans move their eyes multiple times every second and behind every movement is a decision, where to direct gaze. Past research has identified factors influencing this selection, including the stimulus, the task, the scene context, but also oculomotor factors such as the "inhibition of return" (IOR), which describes the effect that saccades to previously visited locations are subsequently inhibited. IOR has also been prominent in many computational models of attentional selection. However, while some studies have found spatial biases away from recent gaze targets consistent with IOR, other studies have found the bias only in some task conditions, and some studies instead even found the opposite bias. Here we conducted an eye tracking experiment using the free-viewing paradigm with 60 natural images, 20 of which were from a previous study (Bays & Husain, 2012), while the other 40 were chosen from 5 distinctly different image categories, including faces, landscapes and fractals. We have analyzed resulting eye movement data with respect to their (relative) amplitude, directions, but most importantly their change in direction. The results show significant differences with respect to the previously mentioned eye movement features between the different image categories. These differences are so pronounced, that using only the proportion of forward and return saccades as features for clustering is sufficient to assign most of the images back to the image category. To quantify this effect and add predictability for new images and categories, we related our empirical results to information theoretic measures of spatial properties of the images' saliency maps. Taken together, our experiment reveals that IOR is not solely an oculomotor bias but instead actively depends on the stimulus itself as a key factor influencing the presence and strength of the IOR and show potential underlying causes using an analysis of the spatial distribution of image features.

The authors acknowledge the support of the cluster project "The Third Wave of AI" as part of the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art

63.334 IDIOSYNCRATIC FIXATION PATTERNS GENERALIZE ACROSS DYNAMIC AND STATIC FACIAL EXPRESSION RECOGNITION

Anita Paparelli[1], Nayla Sokhn[1], Lisa Stacchi[1], Antoine Coutrot[2], Anne-Raphaëlle Richoz[1], Roberto Caldara[1]; [1]Eye and Brain Mapping Laboratory (iBMLab), Department of Psychology, University of Fribourg, Fribourg, Switzerland, [2]Laboratoire d'Informatique en Image et Systèmes d'information, French Centre National de la Recherche Scientifique, University of Lyon, Lyon, France

Facial expression recognition (FER) is crucial for understanding the emotional state of others during social interactions. It has been assumed that all humans share universal visual sampling strategies to achieve this feat. While several recent studies have revealed striking idiosyncratic fixation patterns during face identification, very little is yet known about whether such idiosyncrasies extend to the recognition of static and more ecologically valid dynamic facial expressions of emotion (FEE). To this aim, we tracked observers' eye movements categorizing static and dynamic faces displaying the six basic FEE, all normalized for time presentation (1s), contrast, luminance and the overall sampled energy. We used robust data-driven analyses combining statistical fixation maps (iMap) with hidden Markov Models (EMHMM). Then, by dividing our subjects' fixations into 12 conditions (2 visual modality x 6 basic expressions) we assessed the generalizability from their grouping with EMHMM. Incorporating both spatial and temporal dimensions of eye-movements provides powerful and well-suited measures to assess the presence of reliable individual differences in face scanning strategies. With the use of these comprehensive statistical computational tools, our data revealed the presence of marked idiosyncratic fixation patterns. Interestingly, these individual visual sampling strategies generalized for the decoding of both static and dynamic modalities of FEEs. Moreover, the fixation patterns varied with the expression at hand. Importantly, altogether our data show that spatiotemporal idiosyncratic gaze strategies also occur for the biologically relevant recognition of emotions, further questioning the universality of this process.

This work was supported with funding from the Swiss National Science Foundation awarded to RC (10001C_201145).

63.335 SUCCESSFUL BIASING OF ATTENTION REVEALS COVERT INFORMATION EXTRACTION CONTRIBUTES TO SACCADE TIMING VARIABILITY WITHIN AND ACROSS INDIVIDUALS.

Ryan V Ringer[1] (ryan.ringer@ucdenver.edu), Clayton Ridder[1], Tamar Japaridze[1,2], Carly J Leonard[1]; [1]University of Colorado, Denver, [2]University of Pennsylvania

The decision of where to saccade during search depends on information acquired via covert processes, which takes time. Previously, we found that longer first saccade latency led to fewer eye movements per trial, possibly due to participants using more perceptual input to guide saccades. In this experiment, the probability of stimulus difficulty was varied in an attempt to increase covert processing before saccade initiation to benefit search. Task difficulty was manipulated by jittering the elements in a contour target (high, medium, and no-jitter). This varied across three blocked conditions: a neutral condition containing equal amounts of each jitter, an easy-bias condition containing medium and no-jitter, and a hard-bias condition containing medium and high jitter. Jitter bias was hypothesized to influence attentional strategies with the predominance of no-jitter trials in the easy-bias condition encouraging greater covert processing of the same medium-jitter targets. Consistent with this, the results showed longer first saccade latencies in the easy-bias condition, and these were associated with fewer total trial fixations as in previous work. Moreover, latencies declined significantly with increasing target eccentricity and this slope was steeper in the easy-bias condition, suggesting a change in covert peripheral processing. Lastly, individual differences were analyzed using random effects correlations. Positive correlations were found between bias conditions for both saccade latencies and fixation counts, indicating participants had general search strategies and outcomes. Also, mean first saccade latencies negatively correlated with eccentricity slope in the neutral and easy-bias conditions, meaning the drop in saccade latency with target eccentricity was steeper for participants who tended to wait longer. Impetuous participants showed little systematic variation in terms of target eccentricity. In conclusion, these results indicate that saccadic behavior is influenced by stable but malleable attentional strategies that determine the degree to which extrafoveal information is used to guide saccadic eye movements.

Funding from NEI-R15EY030370 to CJL and the CU Denver Center for Faculty Development & Achievement

63.336 HEALTHY OLDER ADULTS HAVE INEFFICIENT STRATEGIES FOR TASKS REQUIRING SEQUENTIAL MOVEMENTS

Anna Ma-Wyatt[1] (anna.mawyatt@adelaide.edu.au), Adam Georgiou; [1]University of Adelaide, [2]University of Adelaide

There are notable changes to perception and motor performance with healthy ageing. While there is a large literature on the completion of single movements, there is less evidence on how healthy ageing impacts the completion of tasks requiring sequences of coordinated movements with real objects. These tasks are generally analogous to those required for activities of daily living. We investigated how eye and hand movements differed between older (n=14; mean age 70 years) and younger (n=20, mean age 22 years) participants during the completion of a block copying task with real Lego. Participants were asked to complete the task as quickly and accurately as possible. We measured hand movements and eye movements throughout the task. Data analyses showed that older participants had longer movement latencies for the first hand movement, lower peak velocity for reaches and initial saccades, and significantly longer task duration compared to younger participants. Younger participants tended to use eye movement strategies that involved looking between the model and workspace and blocks in a strategic way, while older participants showed less efficient eye movement strategies. Analysis of the sequences of eye and hand movements showed that older participants also made significantly more eye and hand movements during task completion. This pattern of results demonstrates that changes to saccade and movement characteristics affect older participants' planning and execution of eye and hand movements. The results suggest that the ability of older participants to use visual information effectively across multiple saccades also seems to decline, increasing the time needed to complete the task and necessitating more overall saccades and longer fixations.

63.337 THE IMPACT OF FANDOM ON IMAGE VIEWING

Stevie Ann Hill[1], Josey M Spiers[1], Lekhana Ramachandra[1], Ayesha Mulla[1], Jessica Hautsch[1], Robert G Alexander[1]; [1]New York Institute of Technology

Fan studies scholars theorize that interpretations of fandom-related images are reinforced by different ways of "seeing" character interactions. Fans themselves recognize that their interpretations are self-reinforcing: once they begin to see characters in particular ways (e.g., interpreting characters as being in romantic relationships even when those relationships are not explicitly part of the narrative), they report being unable to understand how other viewers could interpret things differently. However, there has not yet been any empirical research assessing how fans actually look at character interactions. Our interdisciplinary study fills that gap, taking a novel approach by tracking the ocular movements of fans and non-fans as they look at images of characters from the television series Supernatural. We tracked participants' eye movements non-invasively with an EyeLink1000+ while they viewed images of characters from the show. We hypothesized that fans would exhibit different patterns of gaze behavior than non-fans during the viewing of character interactions. Our results show that, relative to non-fans, fans exhibited longer fixation durations (gazing longer at each position on average before shifting their gaze to another location) and increased saccadic amplitudes (making larger eye movements while gaze was shifted). Thus, fans and non-fans actively view images in objectively different ways. We theorize that, through repeated exposure to images, fans learn to 'read' images in measurable ways that reinforce their interpretations and that drive specific gaze dynamics. Future work will directly compare the gaze differences for images from the show that fans consider either particularly meaningful or not meaningful and will assess how fan viewing behavior changes over time.

63.338 DID YOU LOOK AT THE MOOSE? DRIVER GAZE BEHAVIOUR WHILE SEARCHING FOR HAZARDS IN DYNAMIC ROAD SCENES

Jiali Song[1] (jiali.song@utoronto.ca), Benjamin Wolfe[1]; [1]Unersity of Toronto Mississauga

Safely responding to a road hazard requires knowing when and where it is in the scene. Conventional eye-tracking studies in traffic safety use direct fixation as a proxy for awareness of the hazard. However, fixation location is a poor indicator of awareness because peripheral vision enables drivers to gather information without direct fixation. Conversely, fixating an object does not necessarily guarantee awareness (e.g. inattentional blindness, looked-but-failed-to-see errors). Given that road scenes change dynamically over time, little is known about when and where drivers look when searching for hazards, particularly when they miss hazards. We investigated these questions by asking drivers (n=8) to watch 262 dashcam videos of near-collisions from the Road Hazard Stimulus dataset. Each video (2-8 sec) contained a hazard that was eventually involved in a near-collision in the left or right half of the video, and participants reported its location. Hazard onset was annotated as the first frame in the video that indicates that a near-collision would occur. We grouped the trials based on whether participants correctly reported the hazard's location. During the time window immediately prior to hazard onset, drivers' gaze distance from the hazard decreased over time on correctly localized trials. On incorrect trials, gaze position was significantly further away for 185ms after hazard onset compared to correct trials ($t(7)>2.38$, $p<0.05$). Furthermore, participants were more likely to look within the annotated hazard region for correct trials than for incorrect trials, particularly from 480ms before hazard onset and onwards ($t(7)>2.37$, $p<0.05$), and intermittently in the 1000ms leading up to the time of response ($t(7)>2.37$, $p<0.05$). These results suggest that, on average, drivers look at hazards when they correctly detect them and that looked-but-failed-to-see errors are relatively rare. Rather, incorrect trials are more likely to be "looked too late" or "failed to look" errors.

We thank Dr. Anna Kosovicheva for her guidance and advice throughout the design and analysis of this project. This work was funded by a National Science and Engineering Research Council grant awarded to Benjamin Wolfe (RGPIN-2021-02730).

63.339 OPENNESS TO EXPERIENCE PREDICTS EYE MOVEMENT BEHAVIOR DURING SCENE VIEWING

Nicholas Wyche[1] (nicholas.wyche@anu.edu.au), Mark Edwards[1], Stephanie Goodhew[1]; [1]Australian National University

Introduction: Individuals' abilities to perform on-demand, goal-directed spatial deployments of attention are distinguishable from their broader preferences for how they deploy their attention when circumstances do not compel a specific deployment strategy. Although these preferences are likely to play a major role in dictating how we interact with the visual world during daily life, they remain relatively understudied. The present study investigated two key questions about preferences for the deployment of spatial attention: firstly, are individuals consistent in their preferences for how they deploy their spatial attention when making shifts of attention versus adopting an attentional breadth? Secondly, which other factors are associated with these preferences? Methods: Across two experiments, we measured how participants preferred to deploy both attentional breadth (using an adapted Navon task) and exploratory eye-movement behaviour (using a free-viewing recall task). We also measured participants' working memory capacities (Experiment 1), as well as their personalities and world beliefs (Experiment 2). Results: Across both experiments, individual differences in preference for attentional breadth and eye movement characteristics were consistently observed, but these two kinds of preference were unrelated to each other. Working memory capacity was not linked to these preferences. However, the personality trait of Openness to Experience robustly predicted two aspects of eye movement behaviour preference (saccade amplitude and scan path length). Higher levels of Openness predicted smaller saccades and shorter scan paths. Conclusions: These findings suggest that personality dimensions may predict preferences for more absorbed engagement with visual information. However, it appears that individuals' preferences for shifts of attention during scene viewing do not necessarily relate to the breadth of attention they choose to adopt.

63.340 WHAT ARE YOU LOOKING AT? THE NEURAL BASIS OF EYE MOVEMENT GUIDANCE TO OBJECTS

Avniel Ghuman[1] (ghumana@upmc.edu), Matthew Boring[1], Michael Ward[1,3], R Mark Richardson[1,4], Taylor Abel[1], Arish Alreja[1,2]; [1]University of Pittsburgh, [2]Carnegie Mellon University, [3]UCLA, [4]Massachusetts General Hospital

When you look, you look at something: your daughter's face, a computer mouse, a road sign, etc. Yet most of our understanding of the neural guidance of eye movements is about how we look at locations in space. Consequently, little is known regarding the neural basis of how we guide eye movements to particular objects, faces, etc. in our environment. To take a first step towards addressing this gap, we recorded simultaneous intracranial electroencephalography with eye tracking in individuals with epilepsy during both free viewing of scenes and natural social interactions. In both paradigms, we could predict the category of object participants were going to look at next based on their brain activity from prior to the saccade onset even after controlling for the spatial location of the eye movement. Neural activity that allowed us to predict what a person was going to look at next came from parietal cortex regions traditionally associated with eye movement guidance, as well as category selective ventral temporal cortex regions. To assess whether this result was due to para-foveal responses to the object of the next fixation, we compared saccades traversing greater than 5 degrees of visual angle to smaller saccades and found that prediction accuracy did not diminish for the larger saccades. Preliminary results from one of our participants shows that next fixation prediction may involve theta frequency activity in category selective ventral visual cortex, putatively related to interactions with theta frequency activity in hippocampal regions involved in visual navigation. While many open questions remain, these results suggest that ventral temporal cortex may be involved in guiding what to look at next, with interactions between ventral temporal cortex, hippocampus, and parietal cortex then guiding where to look for that object.

R01MH132225, R21EY030297

WEDNESDAY, MAY 22, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

Eye Movements: Accuracy, pursuit and eccentricity

63.341 MOTION-CORRECTED EYE-TRACKING (MOCET) IMPROVES GAZE ACCURACY DURING VISUAL FMRI EXPERIMENTS

Jiwoong Park[1,2,3] (jiwoongpark@skku.edu), JaeYoung Jeon[1,3], Royoung Kim[1,3], Kendrick Kay[4], Won Mok Shim[1,2,3]; [1]Center for Neuroscience Imaging Research, Institute of Basic Science (IBS), Republic of Korea, [2]Department of Biomedical Engineering, Sungkyunkwan University (SKKU), Republic of Korea, [3]Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University (SKKU), Republic of Korea, [4]Center for Magnetic Resonance Research (CMRR), Department of Radiology, University of Minnesota

Human eye movements are deeply connected to perception, attention, and memory (Hayhoe and Ballard, 2005), and are worthy of study from both neural and behavioral perspectives. In visual fMRI experiments, acquiring eye-tracking data could enable ecologically valid experiments in which eye movements are allowed. However, it is notoriously challenging to perform traditional camera-based eye-tracking in fMRI with high accuracy. Hence, most visual fMRI experiments are conducted under central fixation. In this study, we aim to improve eye-tracking methodology: specifically, we tackle the challenge that when the participant's head deviates from the initial calibration position, a significant drop in gaze accuracy is incurred (Morimoto and Mimica, 2005). First, we performed simulations using a computational geometry-based eyeball model in order to confirm that small head shifts on the order of what is typically observed in fMRI can lead to substantial inaccuracies in eye-tracking results (0.5-mm head shift can lead to 2-3° gaze error). Next, we quantify the effects of subtle head motions on gaze accuracy during actual fMRI scans, and propose a novel method that leverages head motion parameters derived from standard neuroimaging data preprocessing to compensate for head shifts. This approach, termed Motion-Corrected Eye-Tracking (MoCET), does not require any additional hardware and can even be retrospectively applied to existing data. Our results, based on 3T and 7T fMRI datasets encompassing a diverse range of structured and naturalistic tasks (e.g. interactive 3D video gameplay and retinotopic mapping experiments), reveal that MoCET effectively compensates for head motion-driven drifts, leading to a significant enhancement in gaze accuracy. Specifically, MoCET reduces the error to 1.29 visual degrees compared to traditional eye-tracking methods (e.g. polynomial detrending; 3.24°, uncorrected; 4.4°). Our findings provide a feasible and efficient approach to addressing a major challenge of integrating eye-tracking with fMRI, contributing substantially to the field of cognitive neuroscience research.

This work was supported by the IBS-R015-D1, NRF-2019M3E5D2A01060299, NRF-2019R1A2C1085566 and the Fourth Stage of Brain Korea 21 Project in Department of Intelligent Precision Healthcare, Sungkyunkwan University (SKKU)

63.342 FOLLOW THE DOT: DO WE HAVE IMPLICIT AWARENESS OF OUR OWN EYE MOVEMENTS?

Avery H. Chua[1] (avery.chua@mail.utoronto.ca), Anna Kosovicheva[1]; [1]University of Toronto Mississauga

When asked where they have previously looked, people rarely report their visual behavior correctly. Similarly, people can seldom identify their own eye movements in recognition tasks, despite the large individual differences in gaze metrics previously established in literature. However, most tasks generally probe explicit awareness or memory of one's own eye movements. It is unclear whether poor awareness extends to implicit awareness as well, which may speak to individual differences in gaze behaviour. To investigate this question, we designed a tracking paradigm that involved two tasks. First, participants (n=7) completed a classic visual search task while their eye movements were recorded. Next, participants completed a tracking task in which they were instructed to follow a moving red dot on the screen that replayed either their own previously recorded gaze position or that of another participant. During the replay portion, the dot was visible for 50% of the time, randomly disappearing for brief segments of time, and participants were instructed to move their eyes to where they believed the dot would appear next. Furthermore, replayed eye movements were either superimposed on the same stimulus array displayed during the search task, or on a plain grey background. Tracking accuracy was measured by calculating the cross-correlation between the previously recorded gaze position and the tracked positions. Our results show that, on average, people were not significantly better at tracking their own eye movements versus others ($p = .20$) and that people were not better at tracking the replayed eye movements when it was superimposed on the same stimulus array from the search task ($p = .15$). While tracking accuracy across conditions was overall very high (Fisher Z range: 0.85-1.11), our results suggest that poor awareness of one's eye movements may extend to both explicit and implicit measures.

This work was supported by an NSERC Discovery Grant to A.K.

63.343 OBLIQUE ECCENTRIC GAZE HOLDING EVOKES LARGER HORIZONTAL THAN VERTICAL REBOUND NYSTAGMUS

Terence L. Tyson[1,2] (terencetyson@berkeley.edu), Jorge Otero-Millan[1,3]; [1]University of California, Berkeley, [2]NASA Ames Research Center, [3]Johns Hopkins University

Rebound nystagmus (RN) has been used to probe the adaptation mechanisms of the gaze holding system. This type of nystagmus is invoked upon returning to central gaze after prolonged fixation at an eccentric point and it is thought to reflect a set-point adaptation of the neural integrator circuit. Previous studies have only observed horizontal RN after horizontal eccentric gaze holding, reflecting adaptation of the horizontal neural integrator. The current study seeks to elucidate how eccentric gaze at oblique target locations, where there are equal horizontal and vertical components, affects the characteristics of the resulting RN. For a given trial, subjects were instructed to fixate a central flashing target for 5 seconds, then fixate an eccentric continuous target for 30 seconds at one of six possible locations (4 oblique and 2 horizontal targets), and finally fixate centrally at a flashing target for 15 seconds where RN is assessed. Binocular eye-tracking data were collected at 250 Hz. Slow-phase velocity (SPV) or the slow drift part of the nystagmus was used to quantify the magnitude of RN. Subjects showed elevated SPV in the horizontal component (2.3 deg/s), but less in the vertical component (1.4 deg/s) after oblique eccentric gaze holding, reflecting an oblique RN. The smaller vertical RN may reflect a different behavior of the horizontal and vertical neural integrators which are anatomically separated. Under the current paradigm, oblique gaze holding would appear to have a smaller effect on adaptation of the vertical integrator than in the horizontal integrator. Further investigation is needed to establish the reference frame of this distinction between the horizontal and vertical integrator circuits.

This work was supported by the NASA Ames Research Center Graduate Fellowship Program.

63.344 PRE-SACCADIC SHIFTS OF ATTENTION ARE INDEPENDENT OF STIMULUS ECCENTRICITY

Luan Zimmermann Bortoluzzi[1] (luanbortoluzzi@usp.br), Estevao Carlos-Lima[1], Melissa Hongjin Song Zhu[1], Gabriela Mueller de Melo[1], Gustavo Rohenkohl[1,2]; [1]Institute of Biosciences, University of São Paulo, São Paulo, Brazil, [2]Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Frankfurt, Germany

Humans often explore the visual environment through saccades, bringing relevant visual stimuli to the center of gaze. During saccade programming, visual attention is directed to the location where the eyes will move to, enhancing the perception of visual stimuli presented at the saccade target. Recent studies have suggested that rather than being uniformly distributed in space, there are some important asymmetries in how these pre-saccadic shifts of attention modulate perception across the visual field. This study investigates the effects of pre-saccadic attentional shifts to stimuli at different eccentricities during a psychophysical discrimination task. Sixteen human participants were instructed to fixate at the center of the screen while dynamic 1/f noise stimuli (50 ms refresh rate) were presented at 4, 5.76, 8.32, and 12 degrees of visual angle (dva), in separate blocks of trials. The stimuli sizes were scaled by the magnification factor (1.7, 2.09, 2.65, 3.47 dva respectively). After a random fixation period, a central cue (line with 0.4 dva) was presented for 75 ms indicating the saccade direction. A visual target (1/f noise filtered at 45º or 135º) was randomly presented on either side of the screen 140 ms after the cue onset (i.e., during saccade preparation). Participants had to discriminate the target's tilt (clockwise or counterclockwise), while its contrasts were adjusted using a staircase procedure. We found that saccade preparation improved contrast sensitivity in all tested eccentricities. Because eccentricity and saccade amplitude are always coupled in our study, our results show that the amount of pre-saccadic perceptual modulation is independent of the size of the saccade being prepared.

This work was supported by São Paulo Research Foundation - FAPESP (process number: 2017/10429-5 and 23/08440-1), and National Council for Scientific and Technological Development (CNPq)

63.345 TRANSFER OF REAFFERENCE-SIGNAL RECALIBRATION IN HUMAN SMOOTH PURSUIT EYE MOVEMENTS

Rozana Ovsepian[1] (rosanaovs@gmail.com), David Souto, Alexander C. Schütz; [1]University of Marburg

Smooth pursuit eye movements decouple motion on the retina and motion in the environment. To recover motion in external coordinates, the retinal motion may be compensated by the predicted sensory consequences of pursuit. Previous research showed that the size of the pursuit-related reafference signal is continuously updated, based on an interaction between the motor command of the pursuit, and the retinal motion during the eye movement (Haarmeier et al., 2001). Here, we tested if this recalibration is specific to the speed of pursuit or if it generalizes to other speeds. First, we asked subjects to execute horizontal smooth pursuit at a velocity of 8.5°/s. In the middle of the target trajectory, a background pattern moving at 5°/s in a horizontal direction was presented for 200ms. In these exposure trials, the background was moving either in the same direction (Reafference High) or opposite to the direction of the eye movement (Reafference Low). These exposure trials were interleaved with test trials, in which subjects were asked to report the perceived direction of the background motion, the velocity of which was varied with an adaptive staircase procedure to estimate the point of subjective stationarity (PSS). Replicating previous results, the PSS was clearly shifted in the direction of the background speed observed in the exposure trials. Secondly, we investigated how the exposure to background motion during pursuit affected the perceived motion when different pursuit speeds (5.5°/s, 11.5°/s) were used for test and exposure trials (8.5°/s). We found that the recalibration of the reafference signal transferred robustly to lower and higher pursuit speeds and was not restricted to the exposure speed. We propose that the recalibration of the reafference signal might be based on a gain-control mechanism that allows to generalize recalibration to different pursuit speeds.

This work was supported by the “The Adaptive Mind,” funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art.

63.346 BAYESIAN SPECTRAL ANALYSIS OF CONTINUOUS SMOOTH PURSUIT

Todd Hudson[1,2], John-Ross Rizzo[1,2], Janet Rucker[1]; [1]NYU Grossman School of Medicine, [2]NYU Tandon School of Engineering (BME)

Continuous smooth pursuit, implemented via sinusoidal visual target motion, is a task well-known in both neuroscience and clinical neurology. Typical measures extracted from smooth pursuit include velocity gain, and the presence/absence of saccadic pursuit. Here, pursuit was performed in two 12s epochs, starting at screen center and moving sinusoidally first in the righthand and then in the lefthand side of the screen (3 cycles each). We performed both a conventional and Bayesian spectral analysis of the pursuit response. A conventional measurement of smooth pursuit gain (velocity gain) is carried out by computing the sample-by-sample ratio of estimated eye speed along the axis of target motion, to the speed of the pursuit target. In contrast, Bayesian spectral analysis takes advantage of the built-in structure of the data, and measures the amplitude gain of the sinusoidal component of the response at the fundamental (.25 Hz) frequency. Further, we demonstrate the possibility of measuring response components that might occur at the harmonics, and along dimensions other than the (horizontal) axis of pursuit. Pursuit gain, as obtained in one example participant under these conditions, is measured with two orders of magnitude greater precision using the Bayesian relative to the conventional method. We can also see that the conventional method substantially underestimates pursuit gain [0.9 vs. 1.0] in our example. This method in addition allows us to subtract out the measured pursuit response, greatly simplifying the task of saccade detection. In our example data, we find microsaccades throughout the two pursuit epochs, but no saccades large enough to be reliably detected in a bedside neurological exam. This approach offers new opportunities for basic scientific understanding of the pursuit system during sustained tracking, and clinical discovery in, for example, Parkinson's and MS research, as these additional metrics and sensitivity relate to disease progression and prodromal detection.

63.347 IMPRECISION IN SMOOTH PURSUIT FACILITATES VISUAL PROCESSING

Bin Yang[1] (yangbin1004@gmail.com), Janis Intoy[1], Jonathan Victor[2], Michele Rucci[1]; [1]University of Rochester, [2]Weill Cornell Medical College

Humans use smooth pursuit eye movements to inspect non-stationary objects of interest. These movements fail to perfectly stabilize the object at the center of gaze and are often interspersed by small saccades, resulting in considerable motion of the stimulus on the retina. This behavior is commonly regarded as the unavoidable consequence of limits in oculomotor control. However, it may be perceptually beneficial, as it yields luminance transients on the retina that are likely to strongly stimulate neurons in the visual system. A similar oculomotor behavior during fixation on stationary objects has been shown to play a fundamental role in visual processing by reformatting spatial patterns into an efficient temporal code. Here we examine whether “errors” in smooth pursuit could also serve similar functions. Emmetropic observers (N=9) were asked to discriminate the orientation (±45°) of a Gabor (1 or 10 cycles per degree) embedded in a circular uniform patch (4° diameter), which either remained stationary on the display or translated at a constant speed (4°/s). The Gabor contrast increased gradually (1.5 s ramp) to minimize transients not caused by eye movements. Eye movements were continually measured at high resolution via a dDPI eye-tracker (Wu et al, 2023) and used to reconstruct the spatiotemporal stimulus on the retina. Our results show that, like during fixation, luminance modulations from imperfect tracking counterbalance (whiten) the power spectra of natural scenes over a broad range of spatial frequencies, an operation that attenuates redundant correlations and improves coding efficiency. The frequency band of whitening is narrower than for fixation, but contains more power. Perceptual performance closely followed the structure of temporal modulations, with better discrimination at low and high spatial frequency during pursuit and fixation, respectively. These results suggest that imperfect tracking, which maintains retinal image motion of a tracked object, facilitates visual encoding.

NIH EY18363, NIH EY07977

63.348 KNOWING THE MOMENT OF TARGET OCCLUSION INFLUENCES TIME-TO-CONTACT ESTIMATION STRATEGIES

Pamela Villavicencio[1] (pvillavicencio@ub.edu), Joan López-Moliner[1], Cristina de la Malla[1]; [1]Vision and Control of Action Group, Department of Cognition, Development, and Psychology of Education, Institute of Neurosciences, Universitat de Barcelona, Barcelona, Catalonia, Spain

Estimating the time-to-contact (TTC) of moving objects allows us to properly interact with them even if they are momentarily hidden from view. Sometimes, occlusions are unexpected, disrupting the ongoing accumulation of sensory evidence. Other times, the occlusion onset is predictable. The effect that this may have in estimating TTC has never been addressed. To do so, we asked participants (N=12) to press a key when they thought a 1 cm diameter white target moving rightwards would be aligned with the end of a rectangular occluder. In different, randomly interleaved conditions, the occluder could be located either above or below the target's path. If it was above, it reliably indicated the part of the trajectory during which the target would be occluded. If it was below, the occlusion position could coincide with the occluder or not. Eye movements were recorded to discern whether differences in reliability led to different TTC estimation strategies. In all cases, targets moved at 5, 10 or 15 cm/s, were visible for 150, 300, 600 or 1200 ms, and occluded for 100, 200, 400, 800 or 1600 ms. Participants got feedback about their performance after each trial. TTC estimation errors were more accurate when participants could predict the targets' disappearance (1.73 ms ± 16.18 ms vs 20.08 ms ± 14.42 ms, mean ± SEM). In both conditions, precision clearly improved as the ratio of target visibility to occlusion time increased. The eye movement patterns differed depending on occlusion cue reliability: the predictable condition prompted fixations on the occluder's left edge, while in the unpredictable condition, participants tended to smoothly pursue the target until its disappearance. These results suggest that individuals use different strategies to estimate TTC when confronted with varying reliability in occlusion cues.

This work was funded by grants PID2020-116400GA-I00 and PID2020-114713GB-I00 funded by MCIN/AEI/10.13039/501100011033 to CM and JLM, respectively. PV was supported by grant FPI PRE2021-097890 from the Spanish Ministry of Science and Innovation.

63.349 A COVERED EYE FOLLOWS A TARGET ON A TANGENT SCREEN BUT DOESN'T POINT TO IT

Stephen Heinen[1] (heinen@ski.org), Arvind Chandna[1], Devashish Singh[1], Scott Watamaniuk[1,2]; [1]Smith-Kettlewell Eye Research Institute, [2]Wright State University

We previously showed asynchronous eye movements of an occluded eye during midline smooth pursuit (Chandna et al., 2021) which we attributed to a failure of yoking (ala Hering's law) during vergence. Here we investigate the integrity of yoking during "conjugate" pursuit. Neurotypical observers pursued a small (.4 deg) spot on a tangent screen that moved horizontally with either a predictable sinusoidal profile, or an unpredictable combination of sinusoids. Peak target velocity was either 25 or 5 deg/sec. Binocular eye movements were recorded during binocular and monocular viewing with an EyeLink at 1000 Hz, and occlusion was implemented with an infra-red pass filter to allow recording. We found that during pursuit, the covered eye followed the trajectory of the target, but was horizontally displaced from it. This displacement is typically called phoria by clinicians, and can be a biomarker for strabismus, nevertheless it occurs in neurotypicals. However, it is unknown whether phoria is caused by the extraocular muscles "relaxing" under cover and returning to primary position or by a neural signal. If phoria is due to relaxed muscles, its magnitude should vary systematically with eye position, as the muscles should be tenser as the eyes stray further from primary position. If neural, it should remain constant or be modulated by task demands. Consistent with a neural signal driving phoria, its magnitude varied across conditions (range .8 - 4.8 deg), but usually remained relatively constant throughout a trial. Furthermore, systematic differences in the phoria between conditions occurred; the phoria was smaller during high-speed pursuit, and lower yet in unpredictable conditions. The results suggest the eyes are driven conjugately in a yoked fashion, but with added independent monocular signals that control phoria and are modulated by cognitive factors.

Supported by NIH R01EY034626 and Smith-Kettlewell Eye Research Institute

WEDNESDAY MORNING POSTERS IN PAVILION

Wednesday, May 22, 8:30 am – 12:30 pm, Pavilion

Object Recognition: Models

63.401 HOW TO ESTIMATE NOISE CEILINGS FOR COMPUTATIONAL MODELS OF VISUAL CORTEX

Zirui Chen[1] (zchen160@jhu.edu), Michael Bonner[1]; [1]Johns Hopkins University

A pivotal goal in neuroscience is to develop computational models that can account for the explainable variance in cortical responses to sensory stimuli. It is widely recognized that when evaluating the similarity between brain and model representations, it is necessary to estimate the noise ceiling in measurements of cortical activity. Traditional approaches have focused on factors such as reliability across trials or subjects, with the goal of establishing a benchmark for the maximum predictive accuracy that any model could theoretically achieve. However, one important source of noise that has been largely overlooked in the literature is the reliability of the computational models themselves. In the case of deep learning models, a natural measure of reliability is the consistency of learned representations across different random initializations. Using such a metric of model reliability, we demonstrate how an aggregate noise ceiling can be estimated that accounts for the reliability of trials, subjects, and computational models. Our approach provides a more comprehensive assessment of the limitations in representational models of sensory systems. Our results unveil a striking impact of model reliability as a key constraint in explaining variance in cortical representations. More broadly, our findings highlight the importance of identifying and mitigating model variability, and they open new avenues for refining computational models of cortical sensory representations.

63.402 3D SHAPE RECOGNITION IN HUMANS AND DEEP NEURAL NETWORKS

Shuhao Fu[1] (fushuhao@g.ucla.edu), Daniel Tjan[2], Philip Kellman[1], Hongjing Lu[1]; [1]University of California, Los Angeles, [2]Cerritos High School

Both humans and deep neural networks can recognize objects from 3D shapes depicted with sparse visual information, such as a set of points randomly sampled on the surfaces of 3D objects (termed point cloud). Although networks achieve human-like performance for recognizing objects from 3D shapes, it is unclear whether network models acquire similar 3D shape representations to human vision for object recognition. We hypothesize that training neural networks enable the model to gain access to some local 3D shape features and distinctive parts associated with objects, which are adequate to provide good object recognition performance. However, the networks lack representations of the global 3D shapes of objects. We conducted two experiments to test this hypothesis. In Experiment 1, we created Lego-style point clouds to mimic object shapes constructed by Legos. Lego-style 3D objects disrupt local shape features but preserve the global 3D shape of objects. Point clouds of Lego-style objects were shown to both human participants and a dynamic graph convolutional neural network (DGCNN) trained to recognize 3D objects from point cloud displays. Humans maintained high recognition performance when the disruption of local shape was moderate (e.g., the size of Lego pieces was small) (recognition performance for intact 3D shapes: 90% vs. Lego shapes 89%). In contrast, the DGCNN performance dropped significantly, from 90% to 54%. In Experiment 2, we spatially scrambled object parts to disrupt the global 3D shape. We found the opposite result: human recognition performance for part-scrambled displays significantly worsened, but the neural network showed similar recognition performance for the part-scrambled objects when recognizing objects from intact 3D shapes. Hence, the two experiments provide double-dissociation results to show that human object recognition relies on global 3D shapes, but neural networks learn to recognize 3D objects from local shape features.

We gratefully acknowledge the generous funding provided by NSF BCS-2142269.

63.403 CHARACTERISTICS OF THE EMERGENCE OF CATEGORY SELECTIVITY IN CONVOLUTIONAL NEURAL NETWORKS

Niels J. Verosky[1] (niels.verosky@nyu.edu), Olivia S. Cheung[1,2]; [1]New York University Abu Dhabi, [2]Center for Brain & Health, NYUAD Research Institute

Convolutional neural networks (CNNs) have attained impressive performance on visual categorization. Are CNNs appropriate working models of the human visual system? We investigated how CNN performance might resemble human categorization of animacy in four critical aspects: 1) successful categorization of animals vs. objects independent of image statistics, 2) continuum of perceptual to conceptual processes, 3) early emergence of animal compared with object representations, and 4) stable performance across altered images, such as images filtered to contain only high or low spatial frequencies. We tested ResNet-50 with ImageNet pretraining or Contrastive Language-Image Pretraining (CLIP) to categorize grayscale images of animals and objects that were either of round or elongated overall shapes, where all images of the same overall shapes shared comparable image statistics. Each category contained 12-16 items (e.g., squirrel, dolphin), with 16 exemplars from each item. Low-level visual properties were controlled using the SHINE toolbox. We examined categorization accuracy for animals vs. objects of the CNNs, and used representational similarity analysis (RSA) to examine their internal representations. For RSA, each layer of the CNN representations of all items was compared with theoretical category-selective, shape-selective, animal-selective, and object-selective models. We found that, consistent with human performance, 1) both CNNs categorized the images at high accuracy (92-98%) in the absence of image statistics differences across categories, and formed category-selective representations towards the final layers, 2) the shape-selective representations arose prior to the category representations across the layers, 3) the animal-selective representations emerged from early layers and were stable across layers, whereas the object-selective representations appeared late. However, 4) CNN performance was dramatically impacted by spatial frequency changes: categorization accuracy dropped substantially (53-80%) and the internal representations became highly shape-selective throughout the layers. These results suggest that CNNs reflect strong similarities to human categorization, but are limited in generalization across spatial frequencies.

This work was funded by a New York University Abu Dhabi faculty grant (AD174) and a Tamkeen New York University Abu Dhabi Research Institute grant (CG012).

63.404 DIFFERENTIAL SENSITIVITY OF HUMANS AND DEEP NETWORKS TO THE AMPLITUDE AND PHASE OF SHAPE FEATURES

Nicholas Baker[1] (nbaker1@luc.edu), John Wilder[2], James Elder[3]; [1]Loyola University of Chicago, [2]Northeastern University, [3]York University

Background: While humans are highly sensitive to global shape information, deep neural networks models (DNNs) trained on ImageNet seem to favor local shape features. In the Fourier descriptor (shape frequency) domain, this manifests as much higher human sensitivity to low shape frequencies. Here we ask how this differential sensitivity depends upon the amplitude vs phase structure of these Fourier shape components. Methods: Human observers (n=68) classified animal silhouettes into nine categories. The shapes were lowpass filtered in the shape frequency domain, over a range of frequency cutoffs, using two filtering methods. In method 1, Fourier components beyond the cutoff were zeroed. In method 2, phases were randomized but amplitudes were preserved. We compared human performance against three representative networks: a convolutional model (ResNet-50) and two transformer models (ViT, SWIN). Results: While switching from filtering method 1 to method 2 resulted in a slight decline in human performance, it led to a significant improvement for the networks. What could explain this improvement? One possibility is that networks were simply confused by the smooth shapes produced by method 1. To assess this possibility, we retested the networks using a third filtering method in which phases were randomized and

amplitudes set to normative, uninformative values. While performance improved for these more realistic shape stimuli, for the two transformer models (ViT and SWIN), performance remained below levels seen with method 2, indicating that these networks, unlike humans, are able to make effective use of the amplitude structure of low shape frequency components, even when phases are randomized. Conclusions: While humans use low-frequency shape information more effectively than DNNs, they depend critically on the phase structure of these low-frequency shape components. In contrast, transformer networks exploit the texture-like amplitude structure of these components even when phase is randomized.

The York University Research Chair program, VISTA, NSERC

63.405 QUANTIFYING THE QUALITY OF SHAPE AND TEXTURE REPRESENTATIONS IN DEEP NEURAL NETWORK MODELS

Fenil R. Doshi[1,2] (fenil_doshi@fas.harvard.edu), Talia Konkle[1,2], George A. Alvarez[1,2]; [1]Harvard University, [2]Kempner Institute for the Study of Natural and Artificial Intelligence

Deep Neural Networks (DNNs) have emerged as leading models of high-level visual processing, but there exist key disparities between DNNs and human vision, in particular the models' substantially greater reliance on texture over shape in object recognition. This bias has been quantified through a shape-bias score (Geirhos et al., 2019) where models are presented with images containing conflicting shape and texture cues, and the number of correct shape decisions is divided by the total number of correct decisions (shape correct + texture correct). This shape-bias metric shows significant variance across a broad range of vision models, with more-recent models showing more human-like shape-bias. However, this metric lends itself towards slightly misleading interpretations by not taking the models' absolute performance into account; for example, a model making just a single correct shape decision, and no correct texture decisions, would have a 100% shape bias, which may incorrectly suggest the model having a strong shape representation. To address this limitation, we propose a revised metric, the accuracy-corrected shape-bias, as the square root of the product of both the original shape-bias score and the shape-dependent accuracy (which reflects the total proportion of correct decisions guided by shape cues). We show that a randomly initialized AlexNet model (5.75% shape-dependent accuracy) shows high original shape-bias (0.466) but low accuracy-corrected shape-bias score (0.164), better capturing the fact that these models have impoverished shape representations. Moreover, across over 100 trained models, we find that increases in shape-bias are due to shape-enhancement and equal or greater texture-suppression, and that none of the models examined have "strong" shape representations (none exceed 52.5% shape-dependent accuracy). Overall, we find that the gulf between human and DNN shape representations remains much larger than suggested by bias-scores alone, and that there has been little improvement in shape quality beyond early AlexNet models.

NSF PAC COMP-COG 1946308 to GAA, NSF CAREER BCS-1942438 to TK, Kempner Institute Graduate Fellowship to FRD

63.406 GEOMETRIC PROPERTIES OF OBJECT MANIFOLDS IN NEURAL NETWORK MODELS OF VISUAL CORTEX

Michael Bonner[1] (mfbonner@jhu.edu), Eric Elmoznino, Colin Conwell; [1]Johns Hopkins University

A classic signature of high-performing neural network models of visual cortex is their strong accuracy on object classification tasks. However, recent work suggests that classification accuracy alone is an impoverished metric that fails to capture the complexity of biologically relevant visual representations. Here we sought to gain a richer understanding of the representational structure in a diverse set of deep neural networks (DNNs, N=492) by examining multiple geometric properties of their object manifolds (e.g., dimensionality, radius, between-category separation). We also examined the encoding performance of these networks for predicting scene-evoked fMRI responses in human visual cortex using the Natural Scenes Dataset. Our findings show that geometric properties of object manifolds are in some cases robust predictors of encoding performance, and they reveal the specific ways in which the representational spaces of these networks are structured. Rather than compressing images onto low-dimensional concept manifolds, the best models appear to rely on high-dimensional manifolds that are, nonetheless, well separated in representational space. This representational strategy allows models to separate categories in meaningful ways while still maintaining rich information about image-to-image variation within each category. Together, these findings suggest that the geometry of object manifolds offers a promising lens on the key characteristics of biologically relevant visual representations

63.407 A BIOLOGICALLY INSPIRED FRAMEWORK FOR CONTRASTIVE LEARNING OF VISUAL REPRESENTATIONS: BIOCLR

Zhixian Han[1] (han594@purdue.edu), Vlada Volyanskaya[1], Anne B. Sereno[1,2]; [1]Purdue University, [2]Indiana University School of Medicine

Self-supervised learning is a machine learning paradigm where a model is trained by supervisory signals generated by itself. Self-supervised learning is interesting because it does not require human-labeled data and can train artificial neural networks to capture essential features that are useful for downstream tasks. SimCLR (a Simple framework for Contrastive Learning of visual Representations) is a contrastive self-supervised learning framework that has been shown to outperform many other models. However, the data augmentation procedure (generating additional images by randomly modifying the original images) used by SimCLR may not be biologically plausible. Therefore, we propose BioCLR (a Biologically inspired framework for Contrastive Learning of visual Representations) to better align with how a brain might implement contrastive self-supervised learning models. We used the CIFAR-10 dataset to train our model. Much research supports the idea that primate cortical visual processing is segregated into two streams that generate similar but different visual representations. Using the CIFAR-10 dataset, we trained one artificial pathway to recognize the average color of each image (average vector value of all pixels in the image) and another pathway to recognize the orientation of the object in each image. Then we used the internal neural representations generated by the two pathways to replace the data augmentation procedure in SimCLR and trained our BioCLR model with contrastive self-

supervised learning. We used the supervised object recognition task as the downstream task to test the models. We found that in the downstream task, our BioCLR model achieved significantly higher testing accuracy than a baseline model with the same neural network architecture but was not trained with self-supervised learning. Our results suggest that the different internal representations produced by segregated visual pathways may be used to implement contrastive self-supervised learning and improve the object recognition performance.

Funding to Anne B. Sereno: Purdue University and NIH CTSI (Indiana State Department of Health #20000703). We extend our gratitude to Pankaj Meghani and the Viasat Corporate Partner Mentors offered by the Data Mine at Purdue University for their comments on the earlier versions of the model.

63.408 EVALUATING THE ALIGNMENT OF MACHINE AND HUMAN EXPLANATIONS IN VISUAL OBJECT RECOGNITION THROUGH A NOVEL BEHAVIORAL APPROACH

Yousif Kashef Alghetaa[1] (yousifalghetaa@gmail.com), Simon Kornblith[2], Kohitij Kar[1]; [1]Department of Biology, York University, [2]Anthropic PBC

Understanding how computer vision models make decisions is paramount, particularly with increasing scrutiny from various institutions. The field of Explainable Artificial Intelligence (XAI) provides tools to interpret these model decisions, but the explanations are often at odds. Kar et al. (2022) suggested evaluating the goodness of machine explanations based on their alignment with human cognitive processes. This study builds on that concept, addressing the challenge of reliably approximating human explanations, a task complicated by the limitations of existing psychophysical tools like 'bubbles' and classification-images. Our study introduces a novel method to assess the alignment between human and machine explanations in object discrimination tasks. We establish a two-model framework: a target (ResNet-50, whose explanations are under scrutiny) and a reference model (a fully differentiable model, AlexNet, as a stand-in for humans). The objective is to eventually compare the target model's explanations with human explanations. We begin by analyzing feature attribution maps (heat maps showing how image features influence model outputs) from both models. We compare these maps using various metrics to create a baseline ranking of explanation similarity between ResNet-50 and AlexNet. Following this, we create explanation-masked images (EMIs) by retaining only the most informative pixels based on ResNet-50's (Target) feature attributions. We hypothesize that the impact of these EMIs on both model behaviors could reflect the similarity of their underlying explanations. We then estimate the object discrimination accuracy of both ResNet-50 and AlexNet on these EMIs. The correlation between their performances provides a ranking of explanation similarity. Our results showed a significant correlation (Spearman R=0.65, p=0.003), indicating a strong alignment between the two models' explanations. This finding sets the stage for extending our method to human subjects, using their behavioral responses to EMIs to evaluate the accuracy of ResNet-50's explanations, offering a new direction for comparing machine and human explanations.

Google Research, CFREF, Brain Canada, SFARI

63.409 INTERPRETING DISTRIBUTED POPULATION CODES WITH FEATURE-ACCENTUATED VISUAL ENCODING MODELS

Jacob S. Prince[1] (jacob.samuel.prince@gmail.com), Christopher Hamblin[1], George A. Alvarez[1], Talia Konkle[1,2]; [1]Harvard University, [2]Kempner Institute for the Study of Natural and Artificial Intelligence

A typical view of a natural scene will often contain many different people and objects in a broader surrounding environment, generating an expansive pattern of activity along the ventral visual stream. Longstanding highly productive paradigms in visual neuroscience have focused on understanding different regions of the ventral stream in isolation, by identifying the kinds of stimuli that activate each highly (e.g. in an fMRI localizer study). As such, these approaches do not directly assess how different parts of the broader population code operate in parallel to encode a single complex natural image. Here we introduce a new analytical paradigm aimed at this goal. First, we fit voxel-wise encoding models using the Natural Scenes Dataset and focus our analysis on voxels whose responses are accurately predicted for new images. Then, we apply a new interpretability method called “feature accentuation”, which identifies the features of an image that are critical for driving a voxel’s response, by synthesizing a new version of the image with the relevant features emphasized. As a proof of concept, we show that in everyday images of people in different scene contexts—where both face- and scene-selective voxels are moderately active—we can attribute the activation of face-selective voxels to the people within the scene, and the scene-selective voxels to the surrounding scene context, all within the same image. These initial demonstrations offer a roadmap for subsequent analyses along high-level visual cortex, especially targeting voxels with less-well-understood tuning properties. Critically, this method is general, effective for any voxel or neuron over any image, without presupposing specific content distinctions or tuning ahead of time. As such, this analytical approach enables the dissection of the joint operation of a distributed activation profile, which may provide new insight into how the ventral stream encodes a glance of the rich, complex visual world.

This research was supported by NSF CAREER BCS-1942438 (TK), and an NDSEG research fellowship (JSP).

63.410 INVESTIGATING POWER LAWS IN NEURAL NETWORK MODELS OF VISUAL CORTEX

Keaton Townley[1] (wktownley@gmail.com), Mick Bonner[1]; [1]Johns Hopkins University

Recent work has shown that both the visual cortex and deep neural networks exhibit a power law in the covariance spectra of their representations, suggesting that optimal visual representations have a high-dimensional structure. Of particular interest is the power-law exponent that defines how variance scales across latent dimensions. Here we extracted layer-wise activations from convolutional neural networks (CNNs) pre-trained on a variety of tasks and characterized the power-law exponent in their covariance spectra. We found that CNNs with lower power-law exponents were better models of the visual cortex. We also observed that the covariance spectra differed when spatial information was taken into account. The covariance spectra between convolution channels or within a single spatial location in a convolution filter exhibit a minimum power-law exponent

of 1—similar to what has been observed in the visual cortex. However, the covariance spectrum for the full tensor of layer activations could be significantly lower. To investigate further, we developed a method for initializing untrained CNNs with a power-law in the covariance spectrum of their weights. Here we found that the power-law exponent of the weights significantly modulated the exponent in their activations, and a lower power-law exponent improved their ability to model neural activity in the visual cortex. Interestingly, these networks continue to exhibit a minimum power-law exponent of 1 even when the power-law exponent in their weights is far smaller. In sum, our work suggests that the power-law exponent of channel covariance spectra in CNNs is a key factor underlying model-brain correspondence and that there may be fundamental constraints on the power law exponent of deep neural network representations.

63.411 SPARSE COMPONENTS DISTINGUISH VISUAL PATHWAYS AND THEIR ALIGNMENT TO NEURAL NETWORKS

Ammar Marvi[1] (amarvi@mit.edu), Nancy Kanwisher[1], Meenakshi Khosla[2]; [1]MIT, [2]UCSD

What distinguishes the representations and computations of the ventral, dorsal, and lateral visual streams, and why do current computational models often fail to reflect these differences? Prevailing hypotheses suggest specialized functions for each stream: the ventral stream in object recognition, the dorsal stream in visually guided action, and the lateral stream in motion and social information processing. However, linear encoding models of deep neural networks (DNN) optimized for object categorization predict responses across the three visual streams similarly well. Such findings may indicate a failure to capture neural tuning in model-brain comparison tools, especially those using linear mappings. To address this question we first employed data-driven factorization to identify dominant sparse components within each stream. This method revealed face, place, body, text, and food-selective components in the ventral stream; social interaction, implied motion, and hand-selective components in the lateral stream; and some less interpretable components in the dorsal stream. To systematically assess this effect and its relation to models we propose a new technique – Sparse Components Alignment (SCA) – to measure model-brain alignment while remaining sensitive to neural tuning. Using the same methodological framework as RSA, we assessed stimulus-level representational dissimilarities. However, instead of relying on population geometry, SCA computes pairwise distances between stimuli based on the likelihood that they are processed by the same sparse component. We report three findings: (1) sparse representations differ strikingly across streams, (2) DNNs optimized for object categorization are more similar to the ventral visual stream in these sparse representations, and (3) the clarity of these differences is markedly enhanced with SCA compared to linear encoding or RSA methods. Thus, SCA reveals a notably stronger fit between DNNs and the ventral visual pathway than between DNNs and other pathways, underscoring the importance of characterizing neural tuning—above and beyond representational geometry—in assessing model-brain alignment.

This work was funded by NIH grant R01-EY033843

63.412 SPATIAL FILTERS IN NEURAL NETWORK MODELS OF VISUAL CORTEX DO NOT NEED TO BE LEARNED

Ananya Passi[1] (apassi1@jh.edu), Atlas Kazemian[1], Michael Bonner[1]; [1]Johns Hopkins University

The performance of convolutional neural networks as models of visual cortex relies on pre-training millions of parameters, optimizing them for a specific classification task. This process not only requires massive computational resources, but also results in learned features whose effectiveness is limited by the dataset richness. Furthermore, the time and resource intensive nature of this training process discourages iterative parameter studies, further reducing the interpretability of high-performing models of visual cortex. Here we propose a theoretically grounded convolutional architecture in which the training process is limited to learning linear combinations of pre-defined wavelet filters. This simplified model is based on an iterative process of expanding and subsequently reducing dimensionality in a deep hierarchy of modules, where each module consists of a filtering operation, followed by a non-linearly and channel mixing. We show that this model rivals a traditional pre-trained CNN in explaining stimuli-evoked neural responses to natural scenes in the human visual cortex Our model generates a useful set of features that can be combined to extract information from a wide range of stimuli, and it reduces the number of learned parameters by orders of magnitude. This model can enable neuroscientists to more efficiently perform in-silico analyses and controlled rearing experiments on deep learning models. Moreover, it can also give insights about how visual computation occurs in the brain, owing to its simple organization and reduced dependence on training.

63.413 SPATIAL FREQUENCY DECOUPLING: BIO-INSPIRED STRATEGY FOR NETWORK ROBUSTNESS

Suayb Arslan[1] (sarslan@mit.edu), Michal Fux[1], Hojin Jang[1], Matt Groth[1], Joydeep Munshi[2], Walt Dixon[2], Pawan Sinha[1]; [1]Massachusetts Institute of Technology, [2]GE Research

Convolutional Neural Networks (CNNs) are known to have an inherent bias towards texture and reliance on high spatial frequency elements. These characteristics compromise their classification robustness. How might we incorporate global shape information in the classification pipeline of such networks to capture long-range dependencies? Our electrophysiological studies with human participants provide some clues. We devised an experiment involving high-density EEG measurements from ten participants exposed to low-spatial frequency, high-spatial frequency, and full-resolution images comprising objects and faces. Analyses revealed an unexpected temporal staggering of high versus low spatial frequencies. Decoding of neural information to infer stimulus identity was feasible earlier in the timeline with low spatial frequencies than with high spatial frequencies. These findings have helped us formulate an analogous strategy of spatial frequency decoupling and temporal staging in convolutional network architectures. We find that CNNs endowed with this biologically-inspired feature in their architectural bias demonstrate superior resilience against challenging scenarios, such as viewpoint changes and turbulence. Based on these results, we propose that a staggered feedforward processing sequence, progressing from low to high frequencies, may be an important property to boost network resilience and secure effective out-of-distribution generalization.

This research is supported by ODNI, IARPA. The views are of the authors and shouldn't be interpreted as representing official policies of ODNI, IARPA, or the U.S. Gov., which is authorized to reproduce & distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

63.414 WHEN MACHINES OUTSHINE HUMANS IN OBJECT RECOGNITION, BENCHMARKING DILEMMA

Mohammad Javad Darvishi Bayazi[1,2,3] (mj.darvishi92@gmail.com), Md Rifat Arefin[1,2], Jocelyn Faubert[1,3], Irina rish[1,2]; [1]University of Montreal, [2]Mila - Québec AI Institute, [3]Faubert Lab

In the field of vision science, recent endeavours have aimed to assess the comparative performance of artificial neural network models against human vision. Methodologies often involve the utilization of benchmarks that intentionally perturb or disturb images, thereby measuring noise sensitivity to gain insights into important features for object recognition. Recent studies employing critical frequency band masking have unveiled a perspective, positing that neural networks strategically exploit a wider band and less stable frequency channel compared to the one-octave band of human vision. In this work, we extend the inquiry to encompass diverse modern computer vision models, it becomes apparent that a considerable number of recently developed models outperform human capabilities in the presence of frequency noise. This ascendancy is not merely attributable to conventional techniques such as input image data augmentation but also crucially stems from the proficient exploitation of semantic information within expansive datasets, coupled with rigorous model scaling. Conceiving semantic information from multimodal training as a variant of output augmentation, we posit that augmenting input images and labels holds the potential to improve artificial neural networks to go beyond human performance in the current benchmarks. These advantages establish the idea that these models can be complementary agents for humans, particularly in challenging conditions. Despite acknowledging this progress, we must recognize a limitation in computer vision benchmarks, as they do not comprehensively quantify human vision. Consequently, we emphasize the imperative for vision science-inspired datasets to measure the alignment between models and human vision.

This work was funded by the Canada CIFAR AI Chair Program and from the Canada Excellence Research Chairs (CERC) program and NSERC discovery grant RGPIN-2022-05122. We thank the Digital Research Alliance of Canada and Mila for providing computational resources.

63.415 VISUAL AND AUDITORY OBJECT RECOGNITION IN RELATION TO SPATIAL ABILITIES

Conor J. R. Smithson[1] (conor.smithson@vanderbilt.edu), Jason K. Chow[1], Isabel Gauthier[1]; [1]Vanderbilt University

Domain-general object recognition (o) is the ability to individuate members of an object category. Visual o is typically measured using novel objects (e.g. Greebles). Stimuli used to measure auditory o include birdsong, mechanical keyboard presses, and laughter. Previous work suggests a nearly perfect correlation across the visual and auditory modalities for this ability (Chow et al., 2023). However, until now the relationship between the two modalities has not been tested in a large sample. We also assess whether o can be distinguished from spatial ability, which has historically dominated measures of visual ability in psychometric studies. Using structural equation modeling with a large sample (n = 283), we estimate the relationships between these abilities at the construct level. We find that visual and auditory o are very closely related (r = .8, 95% CI [.68, .92]), but that this relationship is smaller once the influence of fluid intelligence (Gf) is controlled for (r = .6, 95% CI [.36, .83]). This supports the idea that visual and auditory o may rely substantially on a single cross-modal ability, but that they are nevertheless distinct. Model comparison further supports the claim that visual and auditory o are separable, as a model with distinct visual and auditory abilities had superior fit compared to a model with a single cross-modal ability. Spatial ability was measured using three tests (3D rotation, 2D rotation, paper-folding) and had sizable relationships with both visual (r = .74) and auditory (r = .69) o. However, the associations between spatial ability and both visual and auditory o were no longer significant once Gf was controlled for. O's partial independence from Gf and spatial abilities suggests that it could potentially offer incremental validity if used to predict performance in real-world domains requiring visual abilities.

This work was supported by the David K. Wilson Chair Research Fund and NSF BCS Award 2316474

WEDNESDAY, MAY 22, 8:30 AM – 12:30 PM, PAVILION

Face and Body Perception: Models

63.416 BAYESIAN ADAPTIVE ESTIMATION OF HIGH-DIMENSIONAL PSYCHOMETRIC FUNCTIONS: A PARTICLE FILTERING APPROACH

Lars Reining[1] (lars.reining@stud.tu-darmstadt.de), Rabea Turon[1], Philipp Hummel[1], Finn Radatz[1], Christine Lind[4], Angela Yu[1,3], Frank Jäkel[1], Thomas S. A. Wallis[1,2]; [1]Technical University of Darmstadt, Germany, [2]Centre for Mind, Brain and Behaviour (CMBB), Universities of Marburg, Giessen and Darmstadt, Germany, [3]HDSI, UC San Diego, [4]Electrical & Computer Engineering, UC San Diego

Many interesting stimulus spaces are high-dimensional. Exhaustively measuring perceptual decisions in these spaces is infeasible, creating the need for adaptive experimental methods that efficiently explore the space. Current adaptive methods are either not suited to classical, well characterized psychophysical tasks (e.g. Gibbs Sampling with People is not applicable to estimate multidimensional psychometric functions in binary response tasks) or do not scale well to more than 4 dimensions (e.g. QUEST+). Here, we propose a method that estimates the posterior distribution of a multidimensional (logistic) psychometric function with lapses. It selects the next stimulus in the experiment such that the expected information gain is maximized. We use a particle filtering approach to approximate the posterior distribution online. This allows us to update the posterior even for high-dimensional spaces fast enough to be feasible between trials (order of 1 sec). In simulations, we show that with this method the entropy decreases between two and three times faster than when sampling stimuli randomly for a 15-dimensional feature space (an 18-parameter psychometric function, with a 15-dimensional hyperplane, an intercept,

lower and upper asymptotes). We have tested the algorithm in up to 50 dimensions and found that it is still fast and reliable. We validate the algorithm in a human experiment on facial gender categorization. We compute Active Appearance Model (AAM) features (around 500.000 dimensions) for faces of the Chicago Face Database, perform dimensionality reduction to 15 dimensions using PCA and then create a pool of new face images by morphing between faces in the 15-dimensional space. Human participants label the faces as "male" or "female". As in the simulation we show that the adaptive method is at least twice as efficient as random sampling at minimizing the entropy. This method allows us to measure perceptual decision functions in stimulus spaces that were previously infeasible.

Funded by the Hessian research priority program LOEWE within the project "WhiteBox"

63.417 EFFICIENT INVERSE GRAPHICS WITH DIFFERENTIABLE GENERATIVE MODELS EXPLAINS TRIAL-LEVEL FACE DISCRIMINATIONS AND ROBUSTNESS OF FACE PERCEPTION TO UNUSUAL VIEWING ANGLES

Hakan Yilmaz[1] (hakan.yilmaz@yale.edu), Matthew Muellner[1], Joshua B. Tenenbaum[2], Katharina Dobs[3], Ilker Yildirim[1]; [1]Yale University, [2]Massachusetts Institute of Technology, [3]Justus-Liebig University Giessen

At a glance, we not only recognize the category or identity of objects, but also perceive their rich three-dimensional (3D) structure. Critically, this richness of perception is not brittle: Our percepts may degrade under unusual viewing conditions, but they do so gracefully, remaining far above chance, even when the best computer vision systems fail. What renders human perception so distinct—with efficiently inferred, rich representations that are nevertheless robust? Here, we present a new computational architecture of visual perception, Efficient Differentiable Inverse Graphics (EDIG), that integrates discriminative and generative computations to achieve fast and robust inferences of rich 3D scenes. In a bottom-up pass, EDIG uses a discriminatively trained deep neural network (DNN) to initialize a percept by mapping an observed real-world image to its underlying 3D scene. Crucially, EDIG can further refine this initial estimate via iterative, optimization-based inference over a differentiable graphics-based generative model. In a case study of face perception, we train EDIG on a dataset of upright face images, to learn to map these images to 3D scenes in a weakly supervised fashion. We also train an architecture-matched DNN with a standard supervised classification objective, using the same training dataset. We test EDIG, EDIG's bottom-up component, and this alternative on a behavioral dataset of 2AFC identity-matching tasks—with upright and inverted face conditions—consisting of 1560 unique trials per condition. We show that although EDIG and bottom-up only alternatives match average human accuracy on upright faces, only EDIG achieves human-level accuracy on inverted faces. Moreover, EDIG explains significantly more variance in trial-level human accuracy levels than alternatives. EDIG and humans also match qualitatively, both requiring extended processing to match inverted faces, relative to upright faces. These results suggest that human face perception integrates discriminative and generative computations, and provide a blueprint for building humanlike perception systems.

63.418 EVIDENCE FOR EFFICIENT INVERSE GRAPHICS IN THE HUMAN BRAIN USING LARGE-SCALE ECOG DATA

Daniel Calbick[1] (daniel.calbick@yale.edu), Hakan Yilmaz[1], Greg McCarthy[1], Ilker Yildirim[1]; [1]Yale University

The efficient inverse graphics (EIG) hypothesis states that the initial few hundred milliseconds of visual processing unfolds a cascade of computations reversing generative models of how scenes form and project to images, thereby mapping retinal inputs to three-dimensional scene percepts. Here we ask: Is EIG a shared computational motif of face processing between non-human primates and the human brain? In two key domains of vision, perception of faces and bodies, EIG models have been shown to closely recapitulate processing hierarchies in macaque inferotemporal cortex. These studies used representational similarity analysis (RSA) and high spatio-temporal resolution recordings of single neurons, enabling effective model evaluation. To test EIG in the human brain, we turned to a large-scale electrocorticography (ECoG) dataset in epilepsy patients (N=60). All patients passively viewed a common screening task with faces and other stimulus categories. In addition, each patient viewed a stimulus set of only faces, with different subsets of patients seeing one of a number of distinct stimulus sets. Analyzing this data is challenging due to the idiosyncratic distribution of electrodes across patients and differences across stimulus sets. We address these challenges with a data-driven pipeline for functional and anatomical identification and registration of electrodes in ECoG: We exploit the shared screening task to identify face-informative electrodes in each patient and register electrodes to regions of interest (RoI; e.g., fusiform face area). Using RSA, we find that the similarity structure of EIG on a given stimulus set significantly correlates with the empirical similarity matrices, based on the stimulus-evoked time-frequency patterns of electrodes in individual patients. We also confirm this correlation using random shuffling of data. Finally, EIG explains the time-frequency patterns especially highly between 200-400 ms after stimulus onset. These results suggest EIG as a shared motif of visual computation across primates and humans.

Yale Center for Research Computing

63.419 FARET 2.1: ANATOMICALLY PRECISE MANIPULATION OF RACE IN 3D FACE MODELS AND A PIPELINE TO IMPORT REAL FACE SCANS

Emily Martin[1] (emart459@fiu.edu), Shamim Golafshan[1], Fabian Soto[1]; [1]Florida International University

FaReT is a free and open-source software developed to increase experimental control in face research via anatomy-driven 3D face modeling in MakeHuman. However, the software has been limited by (1) race models that may not reflect true morphological differences across ancestry groups and (2) a relatively small database of validated identity models. Currently, MakeHuman is equipped with default race targets (i.e., Caucasian, African & Asian), but the origins of these targets are unclear and are most likely qualitatively created for game development purposes. To develop valid race targets for scientific research, we conducted a literature review of anthropometric studies to identify significant face features for each major racial group and created novel race models within FaReT with these feature changes. We validated the final models by comparing anthropometric measurements of face landmarks with the values found in the literature for the same landmarks. To increase our database of identity models, rather than developing additional models, we constructed a pipeline to automatically transform face models described in the widely-used FLAME model space to FaReT space using a deep neural network. Thousands of real face scans are already publicly available in FLAME space, which opens the possibility of using this large dataset in MakeHuman for anatomically precise measurement and manipulation. In addition, it is possible to easily fit new face scans to FLAME and then MakeHuman using our pipeline.

This work was supported by the National Science Foundation.

63.420 FROM PERCEPTION TO ALGORITHM: QUANTIFYING FACIAL DISTINCTIVENESS WITH A DEEP CONVOLUTIONAL NEURAL NETWORK

Isabelle Boutet[1] (iboutet@uottawa.ca), Artem Pilzak[1], Maxime Ouimet[1], Alice J. O'Toole[2]; [1]University of Ottawa, [2]University of Texas at Dallas

Background: Face distinctiveness is a pivotal concept in face recognition, due to long-standing findings indicating that the rated typicality of a face is inversely related to its recognizability (e.g., Light et al., 1979). In traditional face space models (Valentine et al., 2016), distinctive faces are perceptually distant from the average or prototype face. Objectively measurable face spaces are available from deep convolutional neural network (DCNN), which are highly accurate at face recognition. Here we test whether DCNN models capture human-rated perceptions of face typicality, as well as other human-ratings (memorability, attractiveness, etc.). Method: We utilized FaceNet (Schroff et al., 2015, 2018), a pre-trained DCNN, to derive a distinctiveness index based on the average distance of a face to other faces in the DCNN space. First, we quantified distinctiveness for 418 male and 631 female faces using FaceNet's 512-dimensional embedding space and the cosine as a measure of distances. Second, we computed correlations between ach face's distinctiveness and various human ratings of the same faces (Bainbridge et al., 2013). Results: Male faces classified as more distinctive are rated as less common ($r = 0.27$) and atypical ($r = 0.21$). Female faces classified as more distinctive are rated as less common ($r = -0.35$), more attractive ($r = 0.37$), more egotistical ($r = 0.22$), and more confident ($r = 0.25$) (all $p < 0.001$). Conclusion: DCNN provides a model of memory-related traits that relate to face distinctiveness across both genders. For female faces, distinctiveness also strongly correlates with personality and social attributes. This underscores the utility of DCNN models for understanding distinctiveness and face recognition. Current efforts focus on extending these findings using an alternate face distinctiveness model and another stimulus set to bolster the generalizability of our results.

This research is funded by a Natural Sciences and Engineering Research Council of Canada Discovery grant (2022-03998) to IB.

63.421 MODELING FACE-IDENTITY “LIKENESS” WITH A CONVOLUTIONAL NEURAL NETWORK TRAINED FOR FACE IDENTIFICATION

Connor J. Parde[1,2] (connor.parde@vanderbilt.edu), Alice J. O'Toole[2]; [1]Vanderbilt University, [2]The University of Texas at Dallas

Certain face images are considered a better “likeness” of an identity than others. However, it remains unclear whether perceived likeness is driven by: a) the proximity of a face image to an averaged prototype comprising all instances of an identity (i.e., a “central prototype”), or b) how closely a face image approximates the most-common appearances of the identity (i.e., the “density” of known instances). Convolutional neural networks (CNNs) trained for face recognition can be used to instantiate face spaces that model both within- and between-identity variability. These networks provide a testbed for investigating human perceptions of face likeness. We compared human-assigned likeness ratings with likeness ratings based on identity-prototypes and local image density from a face-identification CNN. Participants (n=50) viewed 20 face images simultaneously (5 viewpoints x 4 illumination conditions) of each of 72 identities and adjusted a slider bar to indicate whether each image was a “good likeness” of the identity being shown. Responses were collapsed across participants to generate a single likeness rating per image. Next, we used face-image descriptors from a CNN to generate likeness ratings based on either the proximity of a descriptor from the “prototype” created by averaging all corresponding same-identity descriptors (prototype-proximity likeness) or the density of same-identity descriptors located around a given descriptor in the output space generated by the CNN (density likeness). For all measures of perceived likeness (human-assigned ratings, prototype-proximity and density), likeness differed across viewpoint ($p < 0.0001$) and illumination ($p < 0.001$). However, only the CNN density-based likeness ratings mirrored the pattern of human likeness ratings. These results demonstrate that density within an identity-specific face space is a better model of human-assigned perceived-likeness ratings than distance to an identity prototype. In addition, these results show that viewpoint and illumination influence perceived-likeness ratings for face images.

Funding provided by National Eye Institute Grant R01EY029692-04 to AOT

63.422 NORM-REFERENCED ENCODING SUPPORTS TRANSFER LEARNING OF EXPRESSIONS ACROSS STRONGLY DIFFERENT HEAD SHAPES

Martin A. Giese[1] (martin.giese@uni-tuebingen.de), Michael Stettler[1], Alexander Lappe[1], Nick Taubert[1], Ramona Siebert[1], Peter Thier[1]; [1]Hertie Institute / CIN, University Clinic Tuebingen

Humans can recognise human facial expressions spontaneously even from non-human heads, cartoons, or animal faces, even if they have never seen such expressions before. The computational basis of this generalization capability is unknown. We propose a novel deep neural network architecture that exploits norm-referenced encoding of facial expressions. This means that facial expressions are represented in terms of the shifts of landmarks relative to the neutral face. RESULTS: We tested the model by training it with human expressions and only a single neutral face of each other non-human head shape. We demonstrate that the developed architecture accomplishes 100% classification accuracy on a matching expression dataset, which we developed, that contains the same expressions retargeted to very different head shapes (humans, monkeys, and cartoon avatars). As opposed to our new architecture, established deep networks machine learning architectures accomplish maximally 63.1% accuracy on this transfer learning task (chance level accuracy: 14.3%). At the same

time, the proposed neural representation reproduces the gradual analogous encoding of expression strength, which has been observed in cortical face-selective neurons in the superior temporal sulcus in monkey cortex. We also demonstrate that the proposed architecture scales up in a data-efficient manner. The proposed algorithm performs better on the FERG cartoon dataset (Aneja et al. 2016) than the original methods proposed by these authors (92.15% as compared to 89.02%). CONCLUSION: Norm referenced encoding provides an interesting theoretical concept not only to explain the tuning properties of face-selective neurons, but also with respect to optimized transfer learning of expressions across different facial shapes.

Supported by ERC 2019-SyG-RELEVANCE-856495, SSTeP-KiZ BMG: ZMWI1-2520DAT700, and NVIDIA Corp.

63.423 READING MINDS IN THE EYES WITH GPT4-VISION

Scott Murray[1] (somurray@uw.edu), Geoffrey Boynton[1]; [1]University of Washington

This study investigates the capabilities of GPT-4, an advanced language model with integrated vision capabilities, in interpreting complex mental states using the Reading the Mind in the Eyes Test (RMET). The RMET involves identifying subtle emotional and mental states from photographs of the region immediately around the human eyes. It is comprised of 36 photographs along with four descriptors of the person's mental state. Like in human studies, we prompted GPT4 (API = gpt-4-vision-preview) to "Choose which word best describes what the person in the picture is thinking or feeling. You may feel that more than one word is applicable, but please choose just one word, the word which you consider to be most suitable. Your 4 choices are: …" We conducted five iterations of the RMET. GPT-4 produced an average of 25.4 items out of 36 correct (SD = 0.89), aligning closely with 'typical' general population human performance range (~25 – 26 items correct). Notably, inverting the images led to a 30% decrease in performance, which is less than the 50% decrease seen in humans, revealing a reliance on global, wholistic processing. Block scrambling the images (2 x 5 grid format), which preserved eye-size features but renders the images nearly unrecognizable to human observers, had almost no impact on GPT-4's performance (24 items correct). This surprising finding suggests that GPT-4's analysis of visual information may prioritize local features (eye gaze, eyebrow characteristics, etc.), over more global aspects of the image. These results provide insight into understanding an AI's visual processing mechanisms, indicating an interplay of feature-specific and holistic image analysis. Overall, the findings show that GPT-4 demonstrates a significant level of competence in recognizing a range of mental states, indicating its potential in applications requiring sophisticated emotional and cognitive understanding.

NIH R01MH131595

63.424 TRAINING DEEP LEARNING ALGORITHMS FOR FACE RECOGNITION WITH LARGE DATASETS IMPROVES PERFORMANCE BUT REDUCES SIMILARITY TO HUMAN REPRESENTATIONS

Nitzan Guy[1] (guynitzan@gmail.com), Mandy Rosemblaum[2], Adva Shoham[3], Galit Yovel[4]; [1]Tel Aviv University

The perceptual representation of facial identity is influenced by factors such as familiarity and experience variability. Yet the impact of overall experience with faces - specifically, the number of identities, amount of exposure for each identity, and the variability of head pose - on the nature of face representations remains unknown. In this work, face-trained deep neural networks were used to answer these questions by manipulating the number of identities, exposure for each identity, and variations in head pose during model training. Model quality evaluation included accuracy assessment with a standard face benchmark (Labeled Face in the Wild dataset), testing for human-like face effects (e.g., the inversion effect), and examining the correlation of models with human similarity representations. Our findings reveal that the number of identities and images per identity significantly influenced model performance. Intriguingly, while increased experience improves accuracy, correlations with human representations were higher for models trained with limited experience (e.g., models trained on only 500 identities and 300 images per identity) compared to models with extensive experience (e.g., CLIP and VGG16 trained on Vggface2 dataset that includes more than 8000 identities and hundreds of images per identity). With respect to head pose, we limited training to poses that varied between frontal to 20 degrees (frontal-only model) and compared to a model that was trained on 25-45 degrees (three-quarter-only model). Whereas both models reached similar performance level, the frontal model was more similar to human representations than the three-quarter model. Taken together, our findings show that similarity between face-trained DNN and human representations does not correspond with model performance, and may not require extensive training with the large datasets of faces that are commonly used to train deep learning models.

ISF 917/21

63.425 VIEW-SYMMETRIC REPRESENTATIONS OF FACES IN HUMAN AND ARTIFICIAL NEURAL NETWORKS

Tim Andrews[1] (ta505@york.ac.uk), David Watson[1], Daniel Rogers[1], Xun Zhu[2]; [1]University of York, [2]Wenzhou University

Models of face processing propose that the neural representation of face identity is initially view-specific, but then becomes view-invariant to enable recognition across different images. Recent studies have suggested that view-symmetry may be an important intermediate representation between view-specific and view-invariant representations. In this study, we compared view-symmetry in humans and a deep convolutional neural network (DCNN) trained to recognize faces (VGG-Face). First, we asked whether view-symmetry is an emergent property of the DCNN for different rotations of the head. We compared the output in the early convolutional layers and the later fully-connected layers of the DCNN to changes in viewpoint caused rotations in yaw (left-right), pitch (up-down) and roll (in-plane rotation). We found that there was an initial view-specific representation in the convolutional layers for yaw, but that a view-symmetric representation emerged in the fully-connected layers. We also found that the ability to differentiate identity was greater across symmetrical compared to non-symmetrical viewpoints. In contrast, we did not find a similar transition from view-specific to view-symmetric representations for either pitch or roll. Next, we compared patterns of response in the DCNN to changes in viewpoint for yaw with corresponding behavioural and neural responses in humans. We found that responses in the fully-connected layers of the DCNN correlated with judgements of perceptual similarity. We also found that the response of the convolutional layers of the DCNN correlated with responses in early visual areas, but that the response of the fully-connected layers correlated with responses in higher visual regions. These findings suggest that view-symmetry emerges when opposite rotations lead to mirror images. The difference in the response to same identity and different identity faces suggests that view-symmetric representations may be important for the recognition of faces in humans and artificial neural networks.

63.426 VISUALIZING THE OTHER-RACE EFFECT WITH GAN-BASED IMAGE RECONSTRUCTION

Moaz Shoura[1] (moaz.shoura@mail.utoronto.ca), Dirk Walther[1], Adrian Nestor[1]; [1]University of Toronto

The other-race effect (ORE) describes the advantage of recognizing faces of one's own race better than other-race faces. While this effect has been extensively documented, its representational basis remains elusive. This study aims to bridge this gap by employing style-based generative adversarial networks (i.e., styleGAN2), a deep learning technique for generating photorealistic images (Karras et al., 2020), in conjunction with facial image reconstruction to investigate the characteristics and mechanisms underlying the ORE. Specifically, we explored how the ORE manifests in styleGAN2, by analyzing the similarity in face representations between GANs and adult participants. This involved assessing the pairwise visual similarity of GAN-generated face images by East Asian and Caucasian participants (N = 106). We then compared the structure of the human face space with that of the GAN latent face space and of other neural network face models (i.e., VGG16 and InsightFace). Our findings suggest that GANs offer insights into face recognition that are not captured by existing models. Furthermore, by leveraging the representational similarity between GANs and human participants, we were able to reconstruct perceptual face representations associated with viewing East Asian and Caucasian face stimuli. Last, we identified latent vector features associated with the ORE and we visualized systematic differences associated with the perception of other-race faces. In conclusion, this research provides a novel perspective on the ORE by integrating generative deep learning techniques in the behavioral study of face perception. The ability of GANs to complement other models of face space structure and perceptual bias underscores their potential as a tool in the study of face perception. Our findings not only contribute to the theoretical understanding of the ORE but also demonstrate the utility of GANs and image reconstruction in behavioral research.

Wednesday, May 22, 8:30 am – 12:30 pm, Pavilion

Face and Body Perception: Disorders, individual differences

63.427 AN ONLINE REPLICATION OF THE ASSOCIATION BETWEEN FACE PERCEPTION ABILITIES AND THE AMOUNT OF VISUAL INFORMATION REQUIRED TO IDENTIFY A FACE

Laurianne Côté[1] (lauriannecote31@gmail.com), Isabelle Charbonneau[1], Caroline Blais[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais

The role of holistic vs. featural information in face processing and the importance of replication are recurrent topics in the field. While some support the importance of holistic processing for individual differences in face identification, others reveal the central role of face parts. Royer and colleagues (2015) found a strong negative correlation between the amount of information required for identification and individual abilities. They conclude that the main mechanism underlying individual differences in face processing ability does not rely on the whole face. Since the original study involved a small, homogeneous sample (35 French Canadians), and considering that this result may seem surprising, an online replication was attempted using Pack&Go, an online platform enabling real-time modifications to stimuli. Online replication allowed for a more diverse sample in terms of socioeconomic level and country of origin. Participants (N=115) completed the same paradigm in which they had to match a target face (front-view or side-view) with one of two front-view faces partially revealed with randomly positioned small Gaussian apertures (Bubbles; Gosselin & Schyns, 2002). The number of bubbles was controlled using QUEST so that each participant achieved a target performance of 75%. Participants also completed a face matching task (GFMT2), which represented a measure of their processing abilities with whole faces. We observed two significant negative correlations between face processing ability and the last number of bubbles in each condition (rhoside = -0.52, pside < 0.001, rhofront = -0.48, pfront < 0.001), indicating that individuals with superior face processing skills require less facial information for accurate task performance. These correlations are of a similar magnitude to the one reported in the original article. These results not only provide further evidence for the pivotal role of part-based processing during face identification, challenging conventional theories, but also showcase the flexibility of online testing.

63.428 EXPLORING SPATIAL FREQUENCY AND ORIENTATION TUNINGS FOR FACE RECOGNITION IN EIGHT CULTURAL GROUPS

Francis Gingras[1,2] (francis.gingras16@gmail.com), Arianne Richer[2], Cousineau Alex[2], Justin Duncan[2], Daniel Fiset[2], Frédéric Gosselin[3], Caroline Blais[2]; [1]Université du Québec à Montréal, [2]Université du Québec en Outaouais, [3]Université de Montréal

East Asians use lower spatial frequencies (SF) compared to Westerners while processing faces (Tardif et al., 2017). These differences have been attributed to culture; however, the underlying mechanism remains unclear. While many hypotheses exist (e.g. social orientation, urbanisation), having data for only two cultural groups makes generalisation difficult/iffy at best. The present study addresses this limitation by measuring SF tunings across eight cultural groups. Preliminary data was collected in Sub-Saharan Africa (n = 70), East Asia (n = 45), Eastern Europe (n = 83), English speaking countries (n = 63), Latin America (n = 89), Middle East (n=50), Southern Asia (n=72) and Western Europe (n=78). Targeted sample size is n=80 for all groups, as pre-registered on the OSF. Participants completed 600 trials of a same/different face matching task online using VPixx Pack & Go (VPixx Technologies, 2021). Target stimuli were filtered using SFO Bubbles, allowing for the sampling of all combinations of SF and orientations (Gingras et al., 2022). A weighted sum of all filters was computed to reveal SFO use for each participant as a 2D classification image. Preliminary analyses comparing the top 1% of t-scores across cultures reveal no differences in orientation tunings but reveal that Eastern and Southern Asians, as well as Sub-Saharan Africans, use lower SF compared to Western Europe/English countries. This is inconsistent with the recently proposed urbanization hypothesis (Caparos et al., 2012), according to which African cultures should show a local bias (and therefore use higher spatial frequencies). While the social orientation hypothesis is more consistent with our results, it fails to predict other visual effects, such as the Ebbinghaus illusion (Caparos et al., 2012) or eye movements (Gingras et al., in press). Other theories, applicable not only to East Asia, but to Southern Asia and Sub-Saharan Africa as well, should be explored.

63.429 INDIVIDUAL DIFFERENCES IN FUSING THE FACE IDENTIFICATION DECISIONS OF HUMANS AND MACHINES

P. Jonathon Phillips[1] (jonathon.phillips@nist.gov), Géraldine Jeckeln[2], Carina Hahn[1], Amy Yates[1], Peter Fontana[1], Alice O'Toole[2]; [1]NIST, [2]The University of Texas at Dallas

Combining the decisions of two individuals (fusion) can increase face identification performance. However, fusion can be beneficial or detrimental, depending on the particular individuals paired (cf. medical imaging, Kurvers et al., 2016). Here, we establish the circumstances under which fusing two particular individuals (human-human) or one individual and a machine (human-machine) reliably benefit face-identification accuracy. We modeled pairwise human-human and human-machine fusions, using human data (27 forensic facial examiners, 32 students, 14 controls) from White et al. (2015) and machine data from a deep convolutional neural network (Parkhi et al., 2015). Both were tested with 84 face-identity matching trials, and indicated whether images showed the same identity or different identities (humans: 5-point scale; machine: similarity score between embeddings). Human and DCNN accuracy were measured as the area under the curve (AUC). For each pair, we averaged the independent responses (human-human, human-machine) and measured: the fused AUC (from averaged responses); the fusion benefit (difference between the fused AUC and the AUC of the more accurate person in the pair), and ΔAUC (the absolute difference in baseline accuracy for the two performers (human-human, human-machine). Fusion benefit decreased as ΔAUC increased [human-human: $r(2626) = -0.7385$, $p < 0.001$; human-machine: $r(71) = -0.8996$, $p < 0.001$]. We used this result to implement "selective fusion" for each human-machine pair. Specifically, ΔAUC was used to determine whether to fuse responses or to retain the better performer's responses. Selective fusion yielded greater accuracy than fused AUC in 88% of trials. To achieve optimal human-human pairing, we implemented graph theory (Galil 1986) and show that optimal pairing (AUC = 0.97) significantly outperformed randomly generated pairs (AUC M = 0.94, SD = 0.006). Optimizing collaboration benefits at the level of individual performers (human or machine) is critical for reducing face-identification errors in applied settings (e.g., forensic face examination).

63.430 MASKED-FACE RECOGNITION LEADS TO LEARNING OF NEW PERCEPTUAL ABILITIES

Hyerim Kim[1] (lssk211037@gmail.com), Nayoung Kim[1], Huynu Kim[2], Oakyoon Cha[1]; [1]Department of Psychology, Sungshin Women's University, [2]Department of German Language, Literature and Culture, Sungshin Women's University

The mask mandate during the COVID-19 pandemic brought attention to the challenge of recognizing masked faces. Common intuition suggests that the difficulty arises from the fact that only the upper half of a masked face is visible and the shapes of chin and mouth do not contribute to recognition. This study sought to understand whether the difficulty stems from a lack of visual information or from the recruitment of new cognitive abilities rarely used in face recognition. Specifically, we investigated whether the same cognitive abilities were employed for the recognition of masked faces and non-masked faces. We selected two sets of face images from the Karolinska Directed Emotional Faces database – each comprising face images of 4 men and 4 women displaying various facial expressions (non-masked face image set). Subsequently, masked versions of these faces were generated for each set (masked face image set). Participants' face recognition abilities were measured for each of the two non-masked image sets and two masked image sets separately. They memorized two face images portraying different identities, and then after a brief interval, selected one face image depicting an identity not present in the previous display. We found a positive correlation between recognition performance for the two masked face image sets, after controlling for the influence of non-masked face recognition. Whereas Bayesian evidence favored the absence of correlation between recognition performance for the two non-masked face image sets, after controlling for the influence of masked face recognition. These results suggest unique variability in masked-face recognition that cannot be explained solely by individual differences in face recognition and that facial features covered by a mask are not a significant source of difficulty in masked-face recognition.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. 2022R1C1C1008628 and No. RS-2023-00211668).

63.431 COLOR ROBUSTLY AFFECTS THE INTENSITY OF FACIAL DISTORTIONS IN TWO CASES OF PROSOPOMETAMORPHOPSIA

Antônio Mello[1], Daniel Garside[2], Daniel Stehr[1], Krzysztof Bujarski[1], Bevil Conway[2], Brad Duchaine[1]; [1]Dartmouth College, [2]National Eye Institute, National Institutes of Health

Prosopometamorphopsia (PMO) is a striking perceptual disorder in which faces appear distorted. Mello et al. (2023) reported that viewing faces through color filters modulated the intensity of face distortions in case V.S., but it was unclear if this effect was independent of luminance and saturation. Here, we investigated this question in case V.S. (59-year-old man) and a new case, Ezri (32-year-old woman), using a replicated single-case experimental design (SCED). Each participant looked at two faces in real life through combinations of colored and neutral-density filters that allowed us to test the impact of hue while controlling for chroma and luminance. We tested perception through a neutral filter (control) and through filters of eight hues that evenly sampled the CIELAB color space at two lightness levels ($L^* = 53, 73$) and two chroma levels (chroma = 32, 52). Each participant rated the intensity of the distortions on a scale from zero ("no distortions") to six ("extreme distortions") across 270 randomized trials (30 per hue and control). Results from multiple regression analyses tailored for SCEDs (Manolov & Onghena, 2018) showed that, in both cases, specific hues significantly modulated the distortions even when considering effects of luminance, chroma, and face identity (V.S.'s adjusted $R2 = 0.75$, Ezri's = 0.38; both $p < .001$). Surprisingly, cool and warm colors had opposite effects for the participants. For V.S., green and cyan decreased distortions ($p < .001$), and lavender, pink, and red increased distortions ($p < .001$); for Ezri, green, cyan, and light blue increased distortions ($p < .001$), and pink decreased distortions ($p < .05$). These results demonstrate that hue can robustly modulate PMO distortions and that the interaction of color and face distortion can show substantial individual differences. In addition, they suggest color may be tightly linked with face perception in typical visual systems.

We thank the Hitchcock Foundation for funding this research.

63.432 FALSE-ALARM RATE AND INTER-TRIAL PRIMING PREDICT HALLUCINATION PRONENESS IN THE SIGNAL DETECTION PAREIDOLIA TEST

Nathan H. Heller[1], Marvin R. Maechler[1], Adithi Jayaraman[1], Vidyulata Kamath[2], Alexander Y. Pantelyat[2], Alyssa Tiedemann[2], Susan H. Magsamen[2], Peter U. Tse[1]; [1]Dartmouth College, [2]Johns Hopkins Medicine

Pareidolia, the phenomenon of seeing faces in clouds, is a form of perceptual false alarm. The recently developed Noise Pareidolia Test (NPT) is used diagnostically to assess hallucination proneness in Parkinsonian psychosis patients (Yokoi et al., 2014). It utilizes images with salient (i.e., high signal) faces embedded in visual noise and pure-noise images. In the NPT, patient populations produce false alarms, but the general population does not. This suggests that hallucinations and pareidolia may share a common neural basis. Normal pareidolic processes in the general population may be amplified in patients. However, the NPT is not suited to test this hypothesis because it does not systematically illicit false alarms in the general population. To address this, we evaluated the Signal Detection Pareidolia Test (SDPT), designed to measure hallucination proneness across populations. In the SDPT, half of the images consist of gaussian-filtered 1/f noise. The rest of the images consist of noise with embedded faces. Four face signal-to-noise ratios are produced by mixing percentages of noise pixels (0.2, 0.35, 0.5, 0.65) with face pixels, and then applying a gaussian filter. In an online study, participants (N=57) viewed each image for 3 seconds before reporting yes-face or no-face. They then rated their confidence on a scale of 1-4. To measure hallucination proneness, participants completed the Cardiff Anomalous Perception Scale (CAPS). Results: two measures of the SDPT significantly correlated with hallucination proneness: 1) false-alarm rate ($r=0.33$, $p=0.012$) and 2) inter-trial priming ($r=0.28$, $p=0.037$), quantified by the conditional probability of reporting yes-face on trial N given a yes-face response on trial N-1. The first result validates that the SDPT can index hallucination proneness in the general population. The second result suggests that perceptual expectation plays a role in hallucination proneness. The latter finding supports a leading theory of hallucinogenesis (Sheldon et al., 2022).

We thank Pat Bernstein and Blink to See for generously funding pareidolia research.

63.433 PERSON COLORS

Adam Reeves[1] (reeves0@gmail.com), Catherine Craver-Lemley[2]; [1]Northeastern University, [2]ElizabethTown College

Synesthete DR experiences a rare form of synesthesia in which the sight of people simultaneously elicits imaginary colors located in space in front of her. We evaluated her convictions that these synesthetic 'people colors' are both stable and associated with their personality traits, as with synesthete S (Collins, 1929, Journal of General Psychology, 2:1, 12-27.) DR's people colors were elicited by previously unknown undergraduates, who each rated themselves on the Ten Item Personality Inventory (TIPI). In Study 1, DR briefly viewed each of 10 subjects in silence through a half-silvered mirror, for two sessions. In Study 2, DR could both see and talk briefly to each of 12 new subjects, for three sessions. Each time DR both detailed her synesthetic colors and filled out her TIPI ratings of each subject. Ratings and colors were reliable across sessions. Her colors were captured as color washes by an artist, to DR's satisfaction. We rated the two most prominent colors in each wash on the yellow-blue and red-green axes. The correlation of DR's yellow-blue color with the subjects' Extraversion, as coded from the TIPI self-reports, rose from $r=0.53$ in Study 1 to $r=0.74$ in Study 2. All other relations were weak ($-0.2<r<0.30$). We conclude that DR's person colors, though uninformative in general, are stably and non-accidentally related to extraversion.

none

63.434 THE EYES STILL HAVE IT: EYE PROCESSING IS A DISTINCT DEFICIT IN DEVELOPMENTAL PROSOPAGNOSIA

Joseph DeGutis[1,2] (deguti@gmail.com), Asha Jotwani[1,2], Mia Haidamus[1,2], Regan Fry[1,2], Leah Kirsch[1,2], Maruti Mishra[1,2,3], Alison Campbell[1,4]; [1]Boston Attention and Learning Laboratory, VA Boston Healthcare System, Boston, MA, [2]Department of Psychiatry, Harvard Medical School, Boston, MA, [3]Department of Psychology, California State University Bakersfield, CA, [4]Department of Psychiatry, Boston University Chobanian & Avedisian School of Medicine, Boston, MA

Numerous studies have demonstrated that efficient processing of the information-rich eye region is important for face recognition. However, the extent to which eye processing is impaired in developmental prosopagnosia (DP) and how these potential impairments contribute to DPs' face perception deficits remains unclear. To address this, we had 110 DPs and 133 controls perform validated face matching assessments (Cambridge Face Perception Test, Computerized Benton Facial Recognition Test, Same/Different Face Matching Test) and the Georges and Part-Whole tasks to measure feature (eye, nose, mouth, chin, forehead sensitivity) and holistic processing (part-whole holistic advantage) abilities. Regarding feature processing, DPs were most impaired on the eye conditions (Georges eye spacing and size trials, Part-Whole part eye trials) but showed relatively normal mouth (Georges mouth size trials and Part-Whole part mouth trials, though deficient Georges mouth spacing trials) and chin performance. DPs also showed a significantly reduced Part-Whole holistic advantage. We next ran factor analyses in DPs and controls using all the facial feature conditions and holistic advantage. In both groups, we found a three-factor solution: a clearly distinct eye processing factor as well as less distinct lower facial feature (mouth and chin trials) and holistic

processing factors. Notably, in regression models using factor scores to predict face perception ability (face matching performance), both DPs and controls showed that eye processing ability predicted the most unique variance, while holistic processing also predicted unique variance. Finally, in a follow-up experiment to determine the generality of DP eye impairments, we examined eye gaze perception and reading emotion from the eyes and found subtle impairments in DPs. In sum, these results demonstrate that, in both DPs and controls, eye processing ability is distinct from more general facial feature processing and holistic processing and is a crucial contributor to DP face perception deficits.

Funded by National Eye Institute R01EY032510 awarded to JD

Wednesday, May 22, 8:30 am – 12:30 pm, Pavilion

Face and Body Perception: Social cognition

63.435 "I FELT SUCCESSFUL!" ASSESSING AUTISTIC ADOLESCENT GAME USABILITY FROM RANDOMIZED CONTROL TRIAL TO IMPROVE SENSITIVITY TO EYE-GAZE CUES

Sumaiya Muhammad[1] (sxm6488@psu.edu), Myles Arrington[1], K. Suzanne Scherf[1]; [1]Penn State University

We demonstrated that Social Games for Adolescents with Autism (SAGA) effectively improves perception and understanding of eye gaze cues among autistic adolescents. SAGA is a serious game in which participants learn to interpret non-verbal social communicative behaviors (e.g., pointing, head turns, eye gaze cues) from characters to solve narrative-related quests. Importantly, SAGA was initially evaluated using a waitlist control. In our current study, SAGA2, we used feedback from autistic adolescents to refine the intervention game and develop an active control game. The intervention and control games were identical in narrative structure and game logistics, including increasing task difficulty and personalized progression. However, instead of focusing on the characters' non-verbal social cues, the control game required participants to complete an object matching task on a screen adjacent to the characters' face. In this randomized controlled trial, 46 autistic adolescents (10-18 years) were randomly assigned to the intervention or control condition. Participants received computers to play the game at home for 45 minutes, thrice a week, over 12 weeks. Post-training, we administered a game usability survey which inquired about participants' impressions regarding the game logistics (e.g., The instructions were clear, My character was easy to control), and their affect experiences during and after the training (e.g., I thought the game was fun, I felt frustrated). Participants in both conditions reported similar levels of ease in game navigation and a general willingness to recommend the game to their friends. Both games elicited similar levels of positive and negative affect during, and after the training. However, the control game elicited slightly more positive affect after the training, likely due to participants getting farther along in the narrative. Together, these usability findings indicate that the control game is well matched to the intervention game in terms of motivating factors, engagement, and difficulty from participants' perspective.

Clinical Trial Number: NCT03690661 Funding Information: R33 MH110624

63.436 EXAMINING EFFECTIVENESS OF A RANDOMIZED CONTROLLED TRIAL TO ENHANCE UNDERSTANDING OF EYE GAZE CUES IN AUTISM: INCORPORATING AN ACTIVE CONTROL GAME IN SAGA

Hunter Mattern[1] (hmm5712@psu.edu), Lark Caboy[1], Myles Arrington[1], Sumaiya Muhammad[1], Justin Wortman[1], K. Suzanne Scherf[1]; [1]Penn State University

Shifts in eye gaze provide non-verbal social communicative information, that is challenging for autistic individual to process. Serious Games for Adolescents with Autism (SAGA) is a intervention that previously showed effectiveness in improving sensitivity to these eye gaze cues in autistic adolescents. Importantly, SAGA was originally tested against a waitlist control. Here, we present data from an ongoing randomized controlled trial (RCT) in which we are evaluating the effectiveness of SAGA2, which includes an active control game. The control game was identical to the intervention game in all respects with one exception; in the intervention game participants view characters portraying non-verbal social communicative behaviors including shifts of eye gaze that can be used to solve narrative-related quests in the game. In the active control game, the characters remain in the scene, but the participants must match objects on a table to target objects presented in a "tv screen." Levels increased in difficulty by altering the target items (e.g., rotated, partially occluded). Forty-six autistic adolescents (ages 10-18) were randomly assigned to either the intervention or control game for a training period of three months, in which participants were instructed to play the game for 45 minutes, three times a week. Participants in the control condition did not differ from those in the intervention condition in the total number of hours spent playing the game, $t(41) = 0.96$, $p = 0.64$. Importantly, they were also comparable to the intervention participants in the number of hours spent engaging specifically in the core tasks of the game (following non-verbal-behavior cues, object-matching), $t(41) = 1.37$, $p = 0.18$. The RCT is ongoing; we will present effectiveness data about the intervention that includes tasks of near transfer (understanding eye gaze) and far transfer (face recognition) of learning.

Clinical Trial Number: NCT03690661 Funding Info: R33 MH110624 – PI Scherf

63.437 A CROWD AMPLIFICATION EFFECT IN THE PERCEPTION OF SOCIAL STATUS

Phyu Sin Myat[1] (myapm-24@rhodes.edu), Matthew Weeks[1], Jason Haberman[1]; [1]Rhodes College

Ensemble perception is the ability to perceive sets of similar objects using a summary statistical representation. However, efficient representations of ensembles can come at a cost, as information about individual differences is necessarily lost through averaging. This work is centered on understanding the socio-cognitive consequences of a system that sacrifices information in favor of the average, explored in the context of social status. Social status is a multidimensional construct that includes an array of factors, including social economic status, trustworthiness, race, among others. Social status is commonly evaluated on a hierarchical ladder (the higher the rung, the higher the status). Our experiments were divided into two phases. In Phase 1, observers rated the social status of full bodies presented singularly. In Phase 2, observers rated the average social status of those same items presented in ensembles. In Experiment 1, race varied but the ensembles, constructed in Phase 2 from the individual ratings of Phase 1, came from the middle of the ladder. Results revealed an amplification effect, whereby the ensemble was rated higher than expected compared to the average of its constituents. Interestingly, there was an interaction with race — black individuals' social status increased more than white individuals' when presented in a crowd. In Experiment 2, sets were composed of only white individuals, but came from multiple rungs on the ladder (i.e., set social status varied). Results replicated our amplification effect, but the direction of amplification varied depending on where on the ladder stimuli came from. Low status sets were rated even lower than what was expected, while high status sets were rated even higher. Overall, our results reveal that perceptions of social status are altered when judging people within the context of a crowd versus in isolation.

63.438 CONTINUED PREFERENCE FOR REVERSED IMAGES OF SELF

Jessica Huang[1] (huajj-24@rhodes.edu), Jordan Suchow[2], Jason Haberman[1]; [1]Rhodes College, [2]Stevens Institute of Technology

Our visual experience of our own face is facilitated through reflections in mirrors and depictions in photographs, although our exposure to each format differs. Previously (2017), we reported a strong preference of the self in mirror reflections, likely mediated by greater exposure to the self in the mirror. With the continued rise of social media, including years of online interactions using applications such as Zoom, it is reasonable to predict people have had increased exposure to depictions of themselves, akin to how they appear in photographs (i.e., not mirror reversed). Such exposure might mitigate the effects we previously observed. In the current experiment, we replicated our original methods, six years after original data collection. We asked participants (N=30) to view their likeness in photographs that were reversed (such as when viewed in a mirror) or not reversed (such as when viewed in a photograph). Participants also adapted (or not) to the reversed or non-reversed photographs in a 2 x 2 design: 1) reversed; 2) non-reversed; 3) adapt to reversed for 45 seconds and judge non-reversed; 4) adapt to the non-reversed image for 45 seconds then view the reversed image. Photographs of each participant were taken just prior to the experiment. Participants then rated (on a scale of 1 – 7) how much each image looked like them and how much they liked the image. Each participant was run in each condition three times in random order. Our original results replicated — participants still preferred their mirror-reversed selves compared to their non-reversal. Their preferences significantly declined when judging non-reversal after adaptation to their mirror reflection. Although it is reasonable to assume familiarity to one's non-reversed self has increased due to digital exposure, our results do not reflect any substantive change in self-representation.

63.439 FACIAL EXPRESSIONS OF APOLOGY COMPRISE COMPLEX SOCIAL SIGNALS

Yichen Wu[1] (2651802w@student.gla.ac.uk), Jonas Nölle[1], Chaona Chen[1], Oliver Garrod[1], Robin Ince[1], Philippe Schyns[1], Rachael Jack[1]; [1]University of Glasgow

Human society relies on maintaining healthy social bonds. When transgressions are made against others, expressing apology is critical to successfully repairing those bonds. Yet, current understanding about apology signals remains controversial. Some theories suggest that apology signals should comprise negative emotions, such as sadness or shame, while others argue that smiles are necessary for appeasement. To resolve this debate, we modelled dynamic facial expressions of apology using a 3D dynamic generative model of the human face and the data-driven method of reverse correlation. On each experimental trial, we generated a facial animation comprising a random selection of facial movements (i.e., Action Units, AUs), eye gaze direction (up, down, left, right, direct) and blushing (present/absent) displayed on a randomly generated face identity (white, 18–35 years, male/female). Participants (N = 60, white Western European, Mage = 20.6 years, sex-balanced) rated how sorry each person looked in response to a given transgression (e.g. "He stepped on your foot") on a 5-point scale from 'Not sorry at all' to 'Extremely sorry'. Each participant completed 2400 such trials in a between-subjects design (n = 20 per transgression) with sex of stimulus face blocked and counterbalanced. We then modelled the specific facial cues of apology by quantifying the statistical relationship between each facial cue (e.g. presence of an AU) and the participant's responses using Mutual Information and non-parametric permutation testing to establish the statistical threshold. Results showed that facial expressions of apology comprise frowning, lateral lip stretching, smiling, and averted gaze (e.g., looking down); blushing did not have a significant effect. Together, our results show that communicating apology involves a unique constellation of facial movements that reflect a combination of negative emotions and appeasement, thereby unifying current accounts. Future work will examine whether and how apology signals vary across transgression types, relationship types, and cultures.

63.440 GAZE ALLOCATION TOWARDS CONTEXTUAL INFORMATION PREDICTS PERFORMANCE IN A DYNAMIC EMOTION PERCEPTION TASK

Jefferson Ortega[1] (jefferson_ortega@berkeley.edu), David Whitney[1]; [1]University of California, Berkeley

When inferring others' emotions, humans often use contextual information (e.g. other people, scene information, etc.) in addition to facial expressions. Previous research has found that observers can accurately infer and track others' emotions by using only contextual information (Inferential Emotion Tracking). However, it's unknown whether using contextual information is what leads to accurate emotion recognition, even when faces are available. Here, we measured human observers' gaze while they continuously inferred the emotion (valence-arousal) of a character in a video, in real-time. Observers completed either the context-only condition, which had the target character blurred out while the context (background scene) remained visible, or the ground-truth condition, where no visual information was blurred out. We calculated between-subject agreement in observers' gaze by calculating the pair-wise Pearson correlation between observers. We found that observers had lower intersubject gaze correlations during the context-only compared to the ground-truth condition ($t(51) = 7.37$, $p < 0.001$). This may be an unsurprising consequence of the reduced information and increased difficulty in the context-only condition. Interestingly, inferential emotion-tracking performance in the context-only condition was significantly correlated with context-only intersubject gaze correlations (spearman $r = 0.535$, $p < 0.01$), but performance in the ground-truth condition was not significantly correlated with ground-truth intersubject gaze correlations (spearman $r = 0.317$, $p = 0.243$). This indicates that knowing where to look in the context leads to more accurate inferences about others' emotions. Additionally, we found observers in the ground-truth condition who had high intersubject gaze correlations with observers in the context-only condition had higher emotion-tracking performance in the ground-truth condition ($p = 0.0072$, permutation test). Our approach provides a means of identifying emotionally relevant information in the context and suggests that individuals who know where emotional clues are in the context are more accurate at emotion recognition, even when faces are available.

This study was funded by National Institutes of Health (1R01CA236793-01).

63.441 INFERENTIAL TRUSTWORTHINESS TRACKING REVEALS FAST CONTEXT-BASED TRUSTWORTHINESS PERCEPTION

Yifan Fang[1] (yfang2024@berkeley.edu), Jefferson Ortega[1], David Whitney[1]; [1]University of California, Berkeley

The ability to infer trustworthiness is fundamental in all social interactions. Humans constantly evaluate others' trustworthiness to avoid potential risks and to build potential relationships with others. Recent studies show that people are able to continuously perceive others' trustworthiness based solely on context (Fang, et al, 2023), but it remains unclear how efficient estimates of trustworthiness from context are, especially in naturalistic dynamic situations. In the present study, we investigated the relative latency with which the visual system uses contextual versus face-specific information to form judgments of trustworthiness in natural movies. Participants continuously tracked and reported the trustworthiness of target characters in movie clips using a Likert-scale, in real-time. Participants were split into one of three conditions: the context-only condition ($n = 40$), which masks the faces and bodies of the target character; the character-only condition ($n = 42$), which masks the background contextual information, and the fully-informed condition ($n = 42$), with no mask. We calculated the cross-correlation function (CCF) between the context-only ratings and the fully-informed ratings and between the character-only ratings and the fully-informed ratings. We found no lag or lead for the character-only condition, and only a 27 msec lag for context-only trustworthy judgments (median lag: -27 msec, bootstrapped 99% CI: -54 to 0 msec). A direct CCF between context and character conditions confirmed the minimal differential latency. The CCF was not confounded by within-subject dependence or memory because each observer only participated in one condition. Our results suggest that surrounding context is available and used for trustworthiness judgments nearly as quickly as face and body information. Contextual information may be processed far more efficiently than previously assumed, and potentially even in parallel to faces.

63.442 PICTURE A SCIENTIST: CLASSIFICATION IMAGES OF SCIENTISTS ARE SEEN AS WHITE, MALE, AND SOCIALLY INEPT

Maheen Shakil[1] (shakilm@mcmaster.ca), Hasan Siddiqui[1], Mel Rutherford[1]; [1]McMaster University

Stereotypes and biases towards social categories are often reflected in mental representations of faces. The current study used a two-phase reverse correlation procedure to visualize mental representations of the face of a Scientist, a Hero, a Genius, and a Person. In the first phase, 20 participants completed four blocks of a two-image forced-choice task. In each block, they selected which face out of a pair looked like one of the four categories. The images they selected were averaged to create classification images (CIs) which are proxy images for their mental representations of the four categories. In the second phase of the study, 251 naive participants rated the CIs on a number of valenced and demographic characteristics. We found that the scientist image was rated as the most White and male, which reflects stereotypes about who pursues scientific careers. The scientist image was also rated more negatively than the other CIs on several characteristics, which might reflect negative biases towards scientists as unsociable, poor communicators, and incompetent authority figures, especially during the COVID-19 pandemic. These findings extend our understanding of the way social categories are represented, and how the classification image method can be used to uncover stereotypes and attitudes regarding these social categories.

This research was funded by a Natural Sciences and Engineering Research Council (Canada) grant to MDR.

63.443 SOCIAL INTERACTIONS CAUSE SPATIAL DISTORTIONS IN VISUAL MEMORY, NOT PERCEPTION

Tim Vestner[1]; [1]Leeds Trinity University, Leeds, UK

Visual perception of social interactions is an essential everyday skill that has recently drawn substantial attention. Notable effects include increased attention to interacting dyads and better memory for individuals within these dyads. It has also been established that individuals in socially interacting dyads are judged to be closer together than those that are not interacting. It is not known whether this spatial contraction of interindividual distance originates in perception or visual memory. Investigating this question is particularly difficult as perceptual distortions would affect both original stimulus and any reproduction ("El Greco fallacy"). The present series of studies uses the well-established kappa effect to solve this problem: When presented with a series of flashing lights, observers will judge the interval between flashes as longer when the lights are perceived to be further apart. Experiment 1 used a variation of this effect and presented participants with three faces, two of which were arranged into a facing dyad, the other facing away from that dyad. Faces flashed in repeated succession and participants were asked to adjust the temporal position of the middle face such that the two resulting intervals were of equal duration. No effects consistent with perceptual distortions of distance were found. Experiment 2 used a memory task where participants were presented with facing or non-facing dyads. After a blank interval that same dyad reappeared at a different location. Participants were required to assess whether this dyad kept the same interindividual distance as before or whether it had changed. Responses were consistent with memory distortions such that facing dyads were remembered as closer together. These results firmly place spatial distortions due to social interactions into visual memory and provide details on the magnitude of these distortions at various timepoints throughout memory.

This research was funded by the BA/Leverhulme Small Research Grants Scheme (SRG2223\231093).

63.444 UNVEILING MENTAL SELF-IMAGES FROM FACE PERCEPTION AND MEMORY

Arijit De[1] (arijit.de@mail.utoronto.ca), Shao Feng Liu[1], Yong Zhong Liang[1], Moaz Shoura[1], Adrian Nestor[1]; [1]University of Toronto

Human cognition and behavior are intricately shaped by self-perception, self-concept and self-esteem. While prior work has begun to uncover visual traits of self-representations, detailed depictions of mental self-images and their evaluation with respect to systematic biases are still largely missing from the field. To address this, our study aims to uncover self-images from perception and memory, as well as to assess their sensitivity to specific perceptual abilities and personality traits. To this end, female Caucasian adults (N=30) evaluated the visual similarity between pairs of female face stimuli, including images of their own faces, as well as between mental images of themselves, as recalled from memory, and other face stimuli. Perception and memory-based self-images were then derived through behavior-based image reconstruction as applied to similarity data and assessed with respect to their visual content. Our investigation revealed significant levels of reconstruction accuracy relative to actual face images of the participants. It further revealed systematic correspondence between perception and memory-based representations of the self. Further, it demonstrated the impact of specific factors (e.g., as captured by self-attractiveness ratings) on the content of self-images. Thus, our findings shed new light on self-representations, on their visual content and on the factors that impact their veracity.

Natural Sciences and Engineering Research Council of Canada

63.445 VISUALIZING FACE REPRESENTATIONS AFTER ADAPTATION

KAZUSA MINEMOTO[1] (k.minemoto@gmail.com), Yoshiyuki Ueda[1]; [1]Kyoto University

A distorted perception from face adaptation has been reported broadly and robustly, and it is used for exploring face representation. While the effect of face adaptation is typically estimated based on changes in participants' judgments regarding subsequent facial stimuli, we aimed to visualize participants' mental representations of faces by using reverse correlation image classification—a psychological method for revealing representation (Mangini & Biederman, 2004). In the experiment, participants adapted to either attractive or unattractive faces for five minutes; following this, the averaged face (i.e., base face) of attractive and unattractive faces was presented for 20 seconds, and participants were asked to memorize it. After a five-minute filler task, participants were asked to judge which of two presented faces—that is, the base faces covered with the same noise pattern but inverted (white and black)—was more similar to their memorized base face. Each participant judged 450 face pairs, and based on their judgments, the noises of the selected and unselected stimuli were each averaged and used for visualizing the representation of each participant. The experiment included three conditions (attractive face adaptation, unattractive face adaptation, and no adaptation), and each was conducted over three days. The first day featured either the attractive or the unattractive face as the adaptation stimulus, and the next day featured the other face. The no-adaptation condition was performed one week after the first day. The results of the visualized images showed that the base face appeared more attractive after the participants had adapted to an unattractive face compared to after they had adapted to an attractive face. This method enables us to reveal how adaptation changes face representation visually. The visualized representation images are presented in our poster.

JSPS KAKENHI Grant Number 19H00628

63.446 WHO YOU LOOKIN' AT? PERCEPTION OF GAZE DIRECTION IN GROUP SETTINGS DEPENDS ON NATURALNESS OF GAZE BEHAVIOR AND CLUTTER

Ruth Rosenholtz[1] (rrosenholtz@nvidia.com), Niall Williams[2]; [1]NVIDIA Research, [2]University of Maryland, College Park

Moving to online meetings inherently impairs our ability to perceive social cues such as the gaze direction of meeting participants. Much of the previous work on human gaze estimation has examined gaze near the forward direction, i.e. toward the camera. However, in-person group interactions include gaze directions over a much wider range. Here we examine the accuracy of estimating the gaze of human-like avatars looking at each other around a table. Gaze direction in this scene varies in magnitude from 22.5 to 157.5 degrees relative to the camera, and from 0 to 67.5 degrees relative to the avatar body. In the realistic condition, each avatar rotated its torso, head, and eyes according to previous work on human gaze behavior. In the simple turn condition, each avatar moved only its head and eyes, by the same angle. We also varied scene richness; richer scenes may provide stronger cues to depth and body angle, while also adding clutter. Ten participants judged the gaze direction of five randomly chosen MetaHuman avatars in mixed conditions for a total of 300 trials. Results show a consistent bias in perceived gaze direction, as previously reported for smaller viewing angles, but not for realistic torso-head-eye orienting. Uncluttered scenes led to better performance. Gaze uncertainty varied modestly over the wide range of gaze directions, as a function of angle relative to the camera (slope <0.05 degrees/degree of viewing angle), and angle relative to the viewer (slope approximately 0.10 degrees/degree of viewing angle). Over all gaze directions studied, subjects estimated gaze with an average standard deviation of 10.6 degrees, as compared to previous work showing standard deviations of 1-3 degrees for gaze directions within 15 degrees of the direction towards the camera.

63.447 DON'T LOOK AT THE CAMERA: ACHIEVING EYE CONTACT IN VIDEO CONFERENCING PLATFORMS

Samyukta Jayakumar[1] (samyukta.jayakumar@email.ucr.edu), Marcello Maniglia[1], Alice Gao[2], Brian Curless[2], Ira Kemelmacher[2], Steve Seitz[2], Aaron Seitz[1,3]; [1]University of California, Riverside, [2]University of Washington, Seattle, [3]Northeastern University, Boston

Eye contact and gaze are important social cues as they convey information about attention, awareness, emotion and intent. For single subjects photographed by a camera, conventional wisdom tells us that

looking directly into the camera achieves eye contact. Is this actually correct? The answer is fundamental in times of increasingly intensive use of video conferencing platforms like Zoom, Microsoft Teams and Google Meet, where, in the absence of in-person non verbal cues, eye-contact can promote positive social interaction. In this study we used subjective and objective measures to assess where subjects should direct their gaze relative to a camera lens to optimize perception of eye contact by an external observer (e.g., a person on the other side of a video call). We collected pictures of 4 actors looking at 11 vertically displaced points (including straight-ahead at the camera) and verified their point of gaze using an eye-tracker. These pictures were then presented to 17 participants who were asked to rate their subjective perception of eye contact as well as whether the actor seemed to be looking up or down. Results suggest that looking just below the camera (~2°) appears to be ideal for achieving the perception of eye contact. We discuss the implications of the current study, offer insights into future directions and explore considerations for correcting for the perception of eye contact in video conferencing platforms.

Wednesday, May 22, 8:30 am – 12:30 pm, Pavilion

Attention: Reward, motivation, emotion

63.448 EMOTIONAL CONSEQUENCES OF EXPENDING PERCEPTUAL EFFORT

Emma Wiedenmann[1] (emma.wiedenmann@gwu.edu), Julia Danett[1], Myeong-Ho Sohn[1]; [1]The George Washington University

Cognitive effort, or the amount of mental energy it takes to complete a task, has paradoxical consequences. Empirically, the emotional consequence of processing effort is task-dependent (Tae et al., 2021). However, cognitive conflict resolution is not the only process demanding cognitive resources. Notably, perceptual effort influences different processing stages, with cognitive conflict resolution demanding computation and perceptual effort demanding encoding. For example, the word “BROCHURE” will be encoded more easily than “BROCHURE”, although the cognitive demand is the same once encoding is successful (Dreisbach & Fischer, 2012). The current study examined whether perceptual effort induces, and if so, whether such influence is task-dependent. Using an immediate priming paradigm, participants performed a prime task with a male or female face, presented in either clear condition or blurry condition. The prime task was followed by a target task which is to recognize emotion from a different face. In Experiment 1, the prime task was to view a perceptually effortful (i.e., blurry) or effortless (i.e., clear) face stimulus passively. In Experiment 2, the prime task was to respond to the perceptual quality of the prime stimulus (e.g., clear or blurry). In Experiment 3, the prime task was to overcome perceptual difficulty and as the participants had to judge the apparent gender of the prime stimulus (e.g., female or male). In the passive viewing task, the perceptual fluency significantly modulated the emotion processing; positive emotions were processed faster after clear trials in comparison to blurry trials. The other two experiments showed no significant interaction between perceptual fluency and emotion response. These results suggest that perceptual effort is subjected to the same emotional process as cognitive effort and is task-dependent.

63.449 THE INFLUENCE OF FEEDBACK AND RISK ON LEARNING TO LINK STIMULUS FEATURES TO REWARD

Hana Taha[1] (hanataha@iu.edu), Thomas W. James[1]; [1]Indiana University Bloomington

Risk as a factor in decision-making has been well-studied in both economics and psychological literatures. However, research on risk, especially in the psychological literature, has focused more on avoidance of negative outcomes and less on the adaptive nature of risk-taking. Thus, this study sought to explore how the choice of more high vs low risk options influenced learning the association between stimulus features and good outcomes. Our task was a multiarmed bandit with six visually distinct slot machines that produced points as outcomes in a 2x3 design with two levels of expected value (EV = 12 or 20) and three levels of risk (low, medium, high). The number of points gained/lost was displayed as feedback on every trial, however, on trials with extreme outcomes (which were more common with high-risk machines), extra feedback was given. The primary dependent measure was the proportion of choices for each machine. Performance was defined as learning to choose the machines that produced higher gains (EV-20). Subjects who did not learn were analyzed separately and the others were divided into high and low performers. In the positive feedback condition, both groups showed a preference for high-risk over medium-risk machines, suggesting that positive feedback increased risk-taking. High performers showed a preference for low-risk over medium-risk machines, suggesting that making more low-risk choices enhanced learning. An analysis of sequences of choices suggested that low performers were more likely to explore across all six machines, while high performers tended to exploit their preferred machines.

63.450 THE INFLUENCE OF ADULT RELATIONSHIP ATTACHMENT STYLE ON THE NETWORKS OF ATTENTION

Ralph Redden[1] (rredden@stfx.ca), Jack Lukeman[1], Veda Haggarty[1]; [1]St. Francis Xavier University

Relationship attachment style has been found to influence performance associated with two of the three networks of attention. Adaptations to information processing for individuals with insecure attachment styles are thought to be a result of increased vigilance to external threats. Attachment avoidance has been shown to influence the executive network of attention, wherein individuals higher in attachment avoidance show reduced interference in the standard Eriksen Flanker task. Attachment anxiety has been shown to influence the executive network as well, in addition to the orienting network of attention, wherein individuals higher in attachment anxiety show larger endogenous cueing effects. The current study was designed to replicate these prior findings, and to examine the heretofore untested relationship between attachment styles and the alerting network of attention, all in a single task. AttentionTrip is an engaging, gamified version of the Attention Network Task, played on an iPad tablet. Observers pilot a ship through a tunnel, using helpful signals (spatial and temporal) to prepare for upcoming targets. The correct weapon must be used to destroy targets, which can be flanked congruently or incongruently by distractors. As in previous work, attachment styles were assessed using the Experience in Close Relationships-Revised Questionnaire. Across individuals, robust network scores were observed for each network of attention. Attachment styles modulated the network effects in the same manner as previously observed, although the sizes of the effects of attachment appear to be smaller in our task. Pertaining to the relationship between attachment and the alerting network, no statistically significant finding was observed. However, individuals with insecure attachment showed slightly faster overall aggregate RTs, lending additional support to the hypothesis that adaptations to attention due to insecure attachment are a consequence of increased vigilance.

St. Francis Xavier University Council for Research

63.451 CAN FOREKNOWLEDGE OF DISTRACTOR TYPE REDUCE THE EMOTION-INDUCED BLINDNESS EFFECT?

Ho Ming Chan[1] (frauzip@connect.hku.hk), Jeffrey Saunders[1]; [1]The University of Hong Kong

This study tests whether foreknowledge of distractor type can reduce the emotion-induced blindness (EIB) effect caused by erotic images. A previous study found that warning about the incoming emotional distractor improved target detection and reduced EIB. One explanation is that foreknowledge allows subjects to better prepare for the distractor and apply proactive inhibition. Another explanation is that the warning encourages more effort. To distinguish these potential explanations, we tested a wrong-warning condition in addition to the correct-warning and no-warning conditions. We performed a preregistered experiment with N=56 subjects. Two types of distractor images were presented in RSVP streams: erotic images and neutral images of people. In most trials with a warning before the RSVP stream (80%), the warning was consistent with the incoming distractor type (correct-warning condition). In remaining warning trials (20%), the warning was inconsistent with the incoming distractor type (wrong-warning condition). In the no-warning condition, only the word “unknown” was shown. If the previously observed benefit of warning was due to the distractor foreknowledge, higher accuracy would only be observed in the correct-warning condition. If the benefit was due to the increased effort, higher accuracy would be observed in both the correct and wrong-warning conditions. Contrary to these hypotheses, neither warning condition showed a benefit compared to the no-warning condition. In all conditions, erotic images impaired detection of target images after a 200ms lag. There was a trend toward higher accuracy with correct warning, suggesting that foreknowledge might provide a small benefit. Unlike the previous experiment, we did not include a baseline condition, so subjects might have exerted higher effort in all our conditions. We were not able to replicate an effect of warning on EIB, and our results suggest that knowing the distractor type provides little or no reduction in the EIB effect.

This research was supported by a grant from the Hong Kong Research Grants Council, GRF 1740914

63.452 YOU SEE FIRST WHAT YOU LIKE MOST: VISUALLY PRIORITIZING POSITIVE OVER NEGATIVE SEMANTIC STIMULI

Sihan He[1] (mikehe0605@hotmail.com), Jay Pratt[1]; [1]University of Toronto

In our complex world, we often encounter situations with multiple objects almost simultaneously entering our visual fields. Identifying the temporal order of these stimuli is thus crucial for scene and event segmentation, and guiding task prioritization. Attending to a stimulus has been found to make it perceived earlier than others, with various attention-modulating factors contributing to this advantage (e.g., reward, ownership, perceived spatial depth). However, the impact of affective valences (positivity or negativity), a significant subjective factor influencing attention selection and processing speed, on temporal order perception remains unexplored. To investigate this issue, we used a cueless Temporal Order Judgement (TOJ) task in three experiments. Observers always saw two Chinese characters on the left and right sides of a central fixation, and there was a variable onset delay between the two characters, ranging between 0 ms and 100 ms (in 20 ms intervals). The observers were instructed to indicate which of the two stimuli appeared first. Different pairs of stimuli valences were used in each experiment: positive and negative (Experiment 1), positive and neutral (Experiment 2), and negative and neutral (Experiment 3). The results of the first and second experiments indicated that people reliably perceived positive stimuli earlier than negative stimuli but not neutral stimuli; the third experiment showed that neutral stimuli were perceived earlier when presented with another negative one. Our findings revealed a general temporal prioritization towards semantically positive stimuli modulated by the strength of affective contrasts.

63.453 APPROACH AND AVOIDANCE VISUAL CUES ARE PROCESSED SIMILARLY IN THE BRAIN

Yuqian Ni[1] (yuqni@iu.edu), Thomas James[1], Eunice Tshibambe[2], Kefu Mu[1]; [1]Indiana University Bloomington, [2]Université catholique de Louvain

Approach-avoidance motivations have been studied for decades but our understanding of motivation under dynamical environments is still limited. In this fMRI study, subjects (N = 30) played a predator-prey maze-navigation game. The game had six types of predator/prey characters (one type per trial), differentiated by color and shape, that had different chances of hurting or healing the player’s character. Subjects were instructed to approach prey or avoid predators to keep their health above zero. One predator and one prey character was designed to only hurt or only heal (non-conflict), whereas the other four characters could both hurt and heal with different probabilities (conflict). Subjects behaved as expected, avoiding non-conflict predators and approaching non-conflict prey and producing a mixture of approach/avoidance behaviors for conflict characters. A mass univariate whole brain analysis comparing the different character types was conducted. A contrast of avoiding predators > approaching prey showed significant upper-tail clusters in dorsal attention network and dorsal striatum and lower-tail clusters in default mode network. Upper-tail clusters overlapped with the conjunction of predator > rest & prey > rest and lower-tail clusters overlapped with the conjunction of predator < rest & prey < rest. Thus, predators produced stronger activation and “deactivation” than prey in regions that were “activated” and “deactivated” by playing the game. A contrast of conflict > non-conflict characters showed upper-tail clusters in regions that matched central executive network. In conclusion, that gameplay-(de)activated regions were the same for predator and prey suggests that approach and avoidance motivation systems are overlapping. Although dorsal attention network, dorsal striatum, and default network showed differences between predator and prey characters, the differences could be explained by consistently stronger (de)activation by predators than prey. Lastly, gameplay with conflict characters recruited regions in central executive network suggests that coactivation of approach-avoidance motivation produces extra demands.

63.454 OPPOSITE POLARITIES IN ALPHA-BAND POWER IN EEG WERE INDUCED BY REWARD AND AROUSAL: AN INITIAL DISCOVERY IN THE PSYCHOPHYSIOLOGICAL REALM THAT DISTINCTLY DISSOCIATES REWARD FROM AROUSAL

Yusuke Nakashima[1] (yusuke_nakashima@brown.edu), Yuka Sasaki[1], Takeo Watanabe[1]; [1]Brown University

Neuromodulatory signals, such as reward and arousal, significantly influence perceptual processing (Seitz et al., 2009; Vinck et al., 2015). While reward and arousal are mediated by distinct neural mechanisms, their psychophysiological responses, like pupil dilation, often exhibit such similarity that discerning between them becomes challenging (Aston-Jones & Cohen, 2005; Bijleveld et al., 2009). In our study, we observed clear differences in alpha-band power in EEG recordings. We conducted experiments measuring pupillary responses and alpha power in EEG following presentations of reward and sound. Each participant underwent a four-hour fasting period before engaging in three conditions. In the reward condition, a drop of water was provided via a tube. In the sound condition, a clicking sound was presented to increase arousal levels. The reward+sound condition involved simultaneous presentation of both water and sound. During each condition, while participants passively viewed a dynamic sequence of Mondrian patterns for a total of 15 sec, water and/or sound was presented 1 to 4 times. Our findings showed a significant increase in pupil size in all three conditions, suggesting that reward and arousal have similar effects on pupillary responses. The alpha power in the occipital region significantly decreased from the pre-stimulus baseline in the sound condition. Conversely, alpha power increased in the reward and reward+sound conditions. These results indicate that alpha power can effectively differentiate between the effects of reward and arousal. Given that decreased alpha power is associated with excitation including arousal, it is plausible that increased alpha power could be involved in inhibition linked to reward.

NIH R01EY019466, R01EY027841, R01EY031705

63.455 REDUCED ATTENTIONAL CAPTURE FOLLOWING MORE VARIABLE REWARDS

Sojung Youn[1] (syoun@tamu.edu), Brian Anderson[1]; [1]Texas A&M University

Value-driven attentional capture is a robust phenomenon in which a reward-associated stimulus automatically captures our attention despite being irrelevant to the task at hand. Last year, we investigated the occurrence of value-driven attentional capture in the presence of reward variability. We did not observe a difference in attentional bias for stimuli associated with consistent or variable reward (expected value equated). Rather, our results suggested that individual preferences for variable or consistent reward, as revealed by a decision-making task, predicted which reward schedule more strongly influenced attention. This year, we will present a follow-up study that probes the relative strength of value-driven attentional biases when the reward variance is considerably more extreme. In this follow-up study, participants were presented with color squares predictive of reward feedback during a training phase. One color was associated with a consistent amount of reward (five cents) while another color predicted reward with high variance (one dollar or no reward), with expected value equated between colors. After training, attentional bias was measured using a task in which participants fixated on a target circle while ignoring a square distractor, each of which could appear in the colors experienced during training. Finally, a manual decision-making task was presented where participants selected between the same colors experienced during training in an effort to maximize earnings. The results showed a significant attentional bias towards the consistent reward color, and a marginally significant correlation between the frequency of distractor fixations and reward preference during the decision-making task. These results replicate our previous finding in which attentional bias and choice preference are linked. Strikingly, in contrast to some more recent reports, we also demonstrate that more variable rewards do not always lead to more robust attentional biases, and that with sufficient variability, consistent rewards will more strongly drive attention.

Wednesday, May 22, 8:30 am – 12:30 pm, Pavilion

Attention: Exogenous, endogenous, gaze

63.456 DECONSTRUCTING THE TASK-EVOKED PUPILLARY RESPONSE

Sean R. O'Bryan[1] (sean_obryan@brown.edu), William Kemball-Cook[1], Joo-Hyun Song[1]; [1]Brown University

Task-evoked pupillary responses can provide a useful biomarker of dynamic internal cognitive states such as effort, attentional control, and surprise. These responses are typically quantified by subtracting pre-trial baseline pupil diameter (PD) from peak PD following a stimulus or motor response, corresponding to response amplitude. However, other physiological components such as baseline PD, dilation velocity, and peak latency all affect the shape of the canonical pupillary response, and moreover, each of these components can directly affect amplitude estimates. Here, we systematically tested the relationships between each of these measures by analyzing pupillometry data across three independent cognitive control experiments: a Stroop task (N = 87), a visual working memory task (N = 74), and a visuomotor adaptation task (N = 30). Furthermore, we examined how and whether each PD measure tracked cognitive load and individual differences in performance. Notably, subject- and trial-level correlation analyses revealed that measurements of amplitude were highly (negatively) correlated with baseline PD, and that the strength of these associations may lead to collinearity issues in studies seeking to separately test the contributions of PD amplitude and baseline to behavior. We also found that although amplitude and velocity measures were positively correlated, velocity explained less than 20% variance in amplitude at the trial level, such that velocity and amplitude measures are not interchangeable. Surprisingly, model comparison revealed that peak latency (i.e., time elapsed between stimulus onset and maximum PD) provided the best fit accounting for both task demands and individual differences in behavior across all three experiments: PD latency explained Stroop interference effects, accounted for load- and capacity-related differences in working memory, and tracked reach error under perturbed feedback conditions in visuomotor adaptation. Together, these results represent a critical step toward better understanding the capabilities (and limits) of pupillometry as a tool in the cognitive sciences.

NSF BCS-1555006

63.457 ENDOGENOUS ATTENTION SAMPLES RHYTHMICALLY UNDER SPATIAL UNCERTAINTY

Xiaoyi Liu[1] (xl1549@princeton.edu), Michael Arcaro[2], Sabine Kastner[1]; [1]Princeton University, [2]University of Pennsylvania

Theories of rhythmic perception propose that the environment is sampled in cycles, with alternating states of enhanced or diminished perceptual sensitivity at an attended location. These cycles involve periodic adjustments in functional communication between cognitive control areas and either the sensory or motor regions (Fiebelkorn et al., 2019). Previous research has shown robust rhythmic sampling with exogenous spatial cues that automatically direct attention (e.g., Landau & Fries, 2012). However, it remains unclear whether the same applies to endogenous, voluntary, attention, which relies on different neural mechanisms, and how spatial uncertainty affects the rhythmic process. To investigate, we conducted an experiment that combined elements of perceptual decision making and spatial attention tasks. In each trial, participants maintained central fixation while viewing two gratings displayed at either side in their periphery. A centrally presented cloud of red and green dots served as a cue, with the red-to-green ratio of the colored dots indicating the location of an upcoming target with 80% validity. Spatial uncertainty was manipulated by varying the ratio of the colored dots. After cue offset, the target, a brief near-threshold orientation change, appeared at one of the two grating locations at a randomly selected cue-target interval (CTI) from 300 to 1100 ms. We observed a significant endogenous attention effect, where target detection performance (accuracy and reaction time) was better at the cued vs. uncued location. This effect diminished with increased spatial uncertainty. When analyzing performance as a function of the CTI, we found a strong behavioral oscillation in the theta band and an anti-phase relationship between cued and uncued locations. This finding supports our hypothesis that endogenous attentional sampling is a rhythmic process. Importantly, preliminary data revealed a larger behavioral oscillation under high versus low spatial uncertainty, indicating an increasing reliance on periodic functional connectivity reweighting to resolve increasing uncertainty.

63.458 GAZE PATTERNS MODELED WITH A LLM CAN BE USED TO CLASSIFY AUTISTIC VS. NON-AUTISTIC VIEWERS

Amanda J Haskins[1] (ajh.gr@dartmouth.edu), Thomas L Botch[1], Brenda D Garcia[1], Jennifer McLaren[2], Caroline E Robertson[1]; [1]Dartmouth College, [2]Dartmouth Hitchcock Medical Center

Atypical visual attention is a promising marker of autism spectrum conditions (ASC). Yet, it remains unclear what mental processes guide individual and group-level gaze differences in autism. This is in part because eyetracking analyses have focused on properties of external visual stimuli (e.g., object categories) and failed to investigate a key influence on gaze: the viewer's own internal conceptual priorities. Disambiguating these influences is crucial for advancing gaze as an endophenotype for autism. Here, we tested the hypothesis that gaze differences in autism stem from abstract conceptual-level information, rather than object categorical information. Adult participants (N = 40; 20 ASC) viewed real-world photospheres (N = 60) in VR. We characterized conceptual-level scene information using human captions, which we transformed into sentence-level embeddings using a large language model (BERT). For each participant, we obtained a "conceptual gaze model": the linear relationship between each participant's gaze and conceptual features (BERT; dimensionality reduced using PCA). To compare the influence of internal, conceptual-level information ("for sale", "sports fan") with external, image-based properties ("hat"), we also modeled gaze patterns using a vision model with comparable transformer architecture (ViT). Using a support vector machine (SVM) iteratively trained to classify participant pairs using conceptual gaze models, we find that individual classification for both autistic and non-autistic participants significantly exceeds chance (62% overall, $p < 0.001$); moreover, individual classification for conceptual gaze models is higher than classification for visual categorical models ($t(39) = 4.9$, $p < 0.001$). Next, using a binary SVM to evaluate group-level differences in autistic gaze patterns, we found higher group classification accuracy for left-out participants when training the SVM on conceptual vs. categorical gaze models ($t(399) = 3.88$, $p < 0.001$). These results suggest that gaze differences are reliable within autistic individuals, and that group-level gaze differences are particularly driven by conceptual-level informational priorities.

63.459 THE DISTINCT ROLE OF HUMAN PIT IN ATTENTION CONTROL

Siyuan huang[1,2] (huangsaiyuan15@mails.ucas.ac.cn), Lan Wang[1], Sheng He[1,2,3]; [1]State Key Laboratory of Brain and Cognitive Science, Institute of biophysics, Chinese Academy of Sciences, [2]University of Chinese Academy of Sciences, [3]CAS Center for Excellence in Brain Science and Intelligence Technology

Attention is a mechanism for brain to direct its resource, including both endogenous and exogenous processes, with the priority map serving to bridge these two processes. Recent studies have revealed that the posterior inferotemporal cortex (PITd) in the monkey brain and its functional homologue in the human brain (hPIT) serves as a node in the endogenous attentional control network, the current fMRI study investigated whether the human PIT area in fact functions as a priority map. The hPIT was robustly identified with strong spatial attentional modulation across motion, color, and shape discrimination tasks. The distinct role in attention control manifests in three ways. Compared to aIPS, FEF, and TPJ (classical nodes of attention), the hPIT showed stronger attentional modulation across all tasks regardless of the presence of visual inputs; and in addition, the hPIT uniquely showed more robust attentional modulation in the presence compared to the absence of visual input, demonstrating its role in both endogenous and exogenous attention which is a key aspect for attention priority map. Consistently, the hPIT also demonstrated function connectivity to both dorsal and ventral attention networks. Further, the modulation on hPIT is sensitive to attention load while invariant to the stimulus category. Our findings demonstrate the distinct role of hPIT in attention control, as an attentional priority map that integrates endogenous and exogenous processes.

63.460 IDIOSYNCRATIC SEARCH: BIASES IN THE DEPLOYMENT OF COVERT ATTENTION.

Nathan Trinkl[1] (ntrinkl@bwh.harvard.edu), Ava Mitra[1], Jeremy Wolfe[1,2]; [1]Brigham and Womens Hospital, [2]Harvard Medical School

Eye tracking of visual search tasks shows that the probability that the eyes will move from the current fixation to a nearby target on the next saccade is only ~50%. How can observers fail to find clearly identifiable targets close to fixation (even if they find it later)? One possibility is that processing within the Functional Visual Field (FVF) around fixation is not homogenous. If so, is that inhomogeneity random or systematic? To answer this question, we asked observers to move their eyes to a cue. 300 msec after cue onset, a ring of 7 black Ls and one black T was briefly flashed. Observers made 4AFC decisions about the orientation of Ts. After response, a new fixation location appeared, and this process repeated for two blocks of 360 trials. We found reliably idiosyncratic patterns of accuracy as a function of radial angle (10 of 16 observers were significantly different from normalized group average accuracy, assessed by Chi-sq, $p<0.001$. Four more $p<0.05$). Is idiosyncratic accuracy a function of idiosyncratic deployment of attention or retinotopic variation in basic visual processing? To test this, we made the T red, allowing it to summon attention without need for search. Duration was staircased to produce ~25% errors. This eliminated systematic idiosyncrasies in accuracy (Only 1 of 20 observers with $p<0.05$). Did the original idiosyncrasies depend on making successive saccades to new fixation points? We repeated the experiment with fixation held at a single location. Idiosyncratic patterns were seen, though they seem weaker than with a moving fixation (7 of 20 observers with $p<0.001$. Three more with $p<0.05$). We do not yet know if the idiosyncratic patterns for one observer would be the same with saccades and with steady fixation. These results suggest that attentional deployment is systematically inhomogeneous in the immediate vicinity of fixation.

National Science Foundation (NSF), Grant #2146617 & Economic and Social Research Council (ESRC) — UK Research and Innovation (UKRI), Grant: ES/X000443/1

63.461 MEASURING INDIVIDUAL DIFFERENCES IN MULTITASKING ABILITY

Lauri Oksama[1] (lauri.oksama@mil.fi), Jaakko Kulomäki[1], Jukka Hyönä[2], Esa Rantanen[3]; [1]Finnish Defence Research Agency, Finland, [2]University of Turku, Finland, [3]Rochester Institute of Technology

Multitasking is an essential skill in many job environments where multiple subtasks must be performed simultaneously. However, when several subtasks occupy the same sensory channel simultaneously, such as displays placed at different points in the visual field, operators of the systems must temporarily choose which subtask to monitor and pay closer attention to. In real-life multitasking, operators must prioritize subtasks, allocating their attention to the most important tasks at the right time, even at the expense of less important tasks. In demanding safety-critical environments, such as aviation, it is necessary to be able to measure the ability of those aspiring to the field to cope with challenging multitasking environments. Recently, we devised a demanding multitask environment that requires subtask prioritization, intelligent and flexible attention control, and time-sharing for successful performance (see Kulomäki et al., 2022). The task consists of four subtasks (like dynamic meters) requiring continuous monitoring and precise motor responses at exactly the right time. In our multitasking environment, subtasks had distinct event rates defining their priorities; the subtask with the highest event rate requires more attention and responses than the subtask with a lower event rate. To motivate participants to perform at their maximum level, a game-like composite scoring system was used, based on penalty and reward scores derived from the accuracy of responses to the subtasks. Here, we present evidence of the content validity of the composite score of the new multitask by examining its correlation with other tests measuring cognitive constructs, such as fluid intelligence, and with other multitasking environments like SYNWORK (Elsmore, 1994). We report correlational results from Finnish Air Force applicants (two data sets, N=456 and N=196). Our findings provide evidence that our new multitask environment shows better discriminant validity (lower correlations) with fluid intelligence compared to previous multitasks like SYNWORK (Elsmore, 1994).

63.462 MIND-WANDERING DURING ENCODING IMPAIRS RECOGNITION FOR BOTH FORGETTABLE AND MEMORABLE COMPLEX SCENES

Shivang Shelat[1] (sshelat@ucsb.edu), Brecken Marome[1], Barry Giesbrecht[1], Jonathan W. Schooler[1]; [1]University of California, Santa Barbara

Mind-wandering is an all-too-familiar experience for us all: the tendency for one's attention to shift away from their external environment to an unrelated internal stream of thought. One consequence is that mind-wandering while studying visual stimuli impairs later recognition performance. Here, we examine how image memorability affects the typical reduction in recognition of images learned during mind-wandering episodes by investigating how spontaneous lapses in sustained attention affect recognition memory for highly memorable and forgettable scenes. Subjects ($n = 60$) viewed a set of either memorable or forgettable images (matched for low-level features and number of objects) and were intermittently probed for mind-wandering. After a digit span task, they completed an old/new memory test for the images intermixed with matched scene category foils. Linear regression analyses found that subjects who saw highly forgettable images exhibited lower hit rate ($\beta = -0.087$, $p = 0.025$), lower d-prime ($\beta = -0.595$, $p = 0.020$), and higher criterion ($\beta = 0.181$, $p = 0.033$), but non-different false alarm rate ($\beta = 0.022$, $p = 0.414$). Mind-wandering was associated with lower hit rate ($\beta = -0.071$, $p = 0.002$), lower d-prime ($\beta = -0.586$, $p < 0.001$), and lower false alarm rate ($\beta = 0.050$, $p = 0.002$), but not with criterion ($\beta = 0.027$, $p = 0.558$). Mind-wandering and memorability did not interact, suggesting that attention and memorability are independent predictors of recognition memory. Future work will explore the effects of conscious accessibility and foil memorability on adopted criterion thresholds during test.

This work is supported by a National Science Foundation Graduate Research Fellowship awarded to the first author under Grant No. 2139319 and by contract W911NF-19-2-0026 from the US Army Research Office.