VSS 25TH ANNIVERSARY

MAY 16-20

ST. PETE BEACH

ABSTRACTS

Schedule Overview

Wednesday, May 14, 2025			
9:00 am - 6:00 pm	Computational and Mathematical Models in Vision (MODVIS)	Satellite	Blue Heron
Thursday, May 15, 2025			
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 16, 2025			
7:00 - 8:00 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard
7:00 am - 6:00 pm	Registration Open	Registration	Grand Palm Colonnade
7:00 am - 10:00 pm	VSS Quiet Lounge	Lounge	Compass
7:00 am - 10:00 pm	VSS Social Lounge	Lounge	Royal Tern
8:00 - 10:00 am	The influence of language in higher-level visual cortex	Symposium	Talk Room 1
8:00 - 10:00 am	A vision scientist walks into a clinic...	Symposium	Talk Room 2
10:00 - 10:30 am	Coffee Break	Break	Garden Courtyard
10:30 am - 12:30 pm	Selective Visual Attention, Alive and Well!	Symposium	Talk Room 1
10:30 am - 12:30 pm	25 Years of Seeing 'Stuff' - Advances and Challenges in Material Perception	Symposium	Talk Room 2
12:30 - 1:00 pm	Coffee Break / Lunch (on your own) Lunch Available for Purchase in Garden Courtyard 11:30 - 2:30	Break	Garden Courtyard
1:00 - 3:00 pm	From insects to fish to mammals: Active vision in non-primate organisms	Symposium	Talk Room 1
1:00 - 3:00 pm	Model-optimized stimuli: more than just pretty pictures	Symposium	Talk Room 2
3:00 - 3:30 pm	Coffee Break	Break	Garden Courtyard
3:00 - 5:00 pm	Friday Afternoon Posters Visual Search: Memory Visual Search: Features, objects Decision Making: Perception Action: Navigation and locomotion Action: Miscellaneous Attention: Inattention, load Attention: Individual differences	Poster Session	Banyan Breezeway
3:30 - 4:45 pm	Eye Movements: Natural tasks, neural mechanisms	Talk Session	Talk Room 1
3:30 - 4:45 pm	Perceptual Organization: Neural mechanisms, models	Talk Session	Talk Room 2
4:45 - 5:15 pm	Coffee Break	Break	Garden Courtyard
5:15 - 7:15 pm	Honoring the Contributions of Eileen Kowler: Eye Movements as Windows to the Mind	Symposium	Talk Room 2
7:15 - 8:15 pm	VSS Undergrad Meetup — it all starts here!	Social	Banyan/Citrus
7:15 - 8:15 pm	VSS Newcomer Lemonade Social	Social	Pirate Island
7:30 - 9:30 pm	Opening Night Reception	Social	RumFish Beach
Saturday, May 17, 2025			
7:00 am - 10:00 pm	VSS Quiet Lounge	Lounge	Compass

7:00 am - 10:00 pm	VSS Social Lounge	Lounge	Royal Tern
7:30 am - 6:45 pm	Registration Open	Registration	Grand Palm Colonnade
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	Attention: Neural, objects, models	Talk Session	Talk Room 1
8:15 - 9:45 am	3D Processing	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Saturday Morning Posters Plasticity and Learning: Clinical Plasticity and Learning: Models Visual Memory: Neural mechanisms Visual Memory: Encoding and retrieval Visual Memory: Objects and features Perceptual Organization: Segmentation, grouping	Poster Session	Banyan Breezeway
8:30 am - 12:30 pm	Saturday Morning Posters Face and Body Perception: Experience, learning, expertise Face and Body Perception: Individual differences Face and Body Perception: Neural Face and Body Perception: Emotion Spatial Vision: Clinical Visual Memory: Models Theory Multisensory Processing: Visual-haptic and visual-vestibular integration	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break Sponsored by Neurobridge	Break	Garden Courtyard & Pavilion
10:45 am - 12:30 pm	Object Recognition: Categories and neural mechanisms	Talk Session	Talk Room 1
10:45 am - 12:30 pm	Temporal Processing	Talk Session	Talk Room 2
12:30 - 2:30 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard 11:30 - 2:30	Break	Garden Courtyard
12:45 - 2:15 pm	Virtual Reality + Eye Tracking for Research Organized by WorldViz VR	Satellite	Blue Heron
12:45 - 2:15 pm	Open Science & Publishing Workshop Organized by the VSS Student-Postdoc Advisory Committee (SPC)	Workshop	Banyan/Citrus
12:45 - 2:30 pm	Science across Countries and Cultures: Does Difference make a Difference?	Workshop	Palm/Sabal/Sawgrass
2:30 - 4:15 pm	Visual memory: General	Talk Session	Talk Room 1
2:30 - 4:15 pm	Face and Body Perception: Facial expressions, social relationships	Talk Session	Talk Room 2
2:45 - 6:45 pm	Saturday Afternoon Posters Eye Movements: Models, clinical, context Object Recognition: Reading Object Recognition: Frames of reference Perceptual Organization: Serial dependence Perceptual Organization: Neural mechanisms Pre-Data-Collection Poster Session	Poster Session	Banyan Breezeway
2:45 - 6:45 pm	Saturday Afternoon Posters Attention: Neural, spatial Attention: Reward	Poster Session	Pavilion

	Attention: Divided, tracking Multisensory Processing: Audiovisual integration Color, Light and Materials: Adaptation, constancy and cognition Color, Light and Materials: Surfaces and materials Motion: Illusions Motion: Biological, self-motion Temporal Processing: Duration, timing perception Temporal Processing: Clinical		
4:15 - 5:00 pm	Afternoon Coffee and Snack Sponsored by The Center for Visual Science (CVS) at the University of Rochester	Break	Garden Courtyard & Pavilion
5:15 - 7:00 pm	Decision Making	Talk Session	Talk Room 1
5:15 - 7:00 pm	Visual Search	Talk Session	Talk Room 2
7:30 - 8:30 pm	Keynote Lecture given by Anne Churchland The intersection of vision and movements in the mammalian brain	Keynote	Talk Room 2
Sunday, May 18, 2025			
7:00 am - 10:00 pm	VSS Quiet Lounge	Lounge	Compass
7:00 am - 10:00 pm	VSS Social Lounge	Lounge	Royal Tern
7:30 am - 6:45 pm	Registration Open	Registration	Grand Palm Colonnade
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	Spatial Vision: Crowding and eccentricity, clinical, models	Talk Session	Talk Room 1
8:15 - 9:45 am	Development	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Sunday Morning Posters Perceptual Organization: Ensembles 3D Processing: Shape Scene Perception: Ensemble Visual Search: Eye movements, scenes, real-world stimuli Undergraduate Just-In-Time 1	Poster Session	Banyan Breezeway
8:30 am - 12:30 pm	Sunday Morning Posters Action: Grasping, reaching, pointing, affordances Action: Perception and recognition Eye Movements: Saccades, remapping Eye Movements: Pursuit, learning, vergence Eye Movements: Neural mechanisms Binocular Vision: Rivalry and bistability, stereopsis, models, neural mechanisms Binocular Vision: Clinical, perception Attention: Neural mechanisms Attention: Spatial	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break	Break	Garden Courtyard & Pavilion
10:45 am - 12:30 pm	Visual Memory: Neural mechanisms of working memory	Talk Session	Talk Room 1
10:45 am - 12:30 pm	Object Recognition: Models	Talk Session	Talk Room 2
12:30 - 2:30 pm	Canadian Vision Science Social	Satellite	Sabal/Sawgrass

12:30 - 2:30 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard 11:30 - 2:30	Break	Garden Courtyard
12:45 - 2:15 pm	Making Vision Fun: Novel Approaches to Teaching	Satellite	Blue Heron
12:45 - 2:15 pm	Hands-on Data Visualization Brown Bag	Satellite	Glades/Jasmine
1:00 - 2:00 pm	Navigating International Research Funding	Workshop	Banyan/Citrus
2:30 - 4:15 pm	Perceptual Organization: Objects, events, ensembles	Talk Session	Talk Room 1
2:30 - 4:15 pm	Binocular Vision	Talk Session	Talk Room 2
2:45 - 6:45 pm	Sunday Afternoon Posters Visual Search: Models, strategy, sequential effects, context Decision Making: Actions Decision Making: Models Spatial Vision: Natural image statistics, texture Face and Body Perception: Body	Poster Session	Banyan Breezeway
2:45 - 6:45 pm	Sunday Afternoon Posters Visual Memory: Capacity and encoding of working memory Plasticity and Learning: Adaptation Development: Neural Development: Amblyopia, binocular Development: Infants, children Eye Movements: Pupillometry Eye Movements: Perception, fixational eye movements Motion: Local, higher-order, in-depth Motion: Models, neural mechanisms Multisensory Processing: Perception, neural, clinical	Poster Session	Pavilion
4:15 - 5:00 pm	Afternoon Coffee and Snack Sponsored by The Center for Visual Science (CVS) at the University of Rochester	Break	Garden Courtyard & Pavilion
5:15 - 7:15 pm	Attention: Neural mechanisms	Talk Session	Talk Room 1
5:15 - 7:15 pm	Color, Light and Materials: Cones to cognition	Talk Session	Talk Room 2
7:15 - 8:45 pm	FoVea: The hidden curriculum in vision sciences	Satellite	Banyan/Citrus
	Monday, May 19, 2025		
7:00 am - 2:30 pm	VSS Quiet Lounge	Lounge	Compass
7:00 am - 2:30 pm	VSS Social Lounge	Lounge	Royal Tern
7:45 - 8:45 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard & Pavilion
7:45 am - 2:30 pm	Registration Open	Registration	Grand Palm Colonnade
8:00 am - 12:30 pm	Exhibits Open	Exhibits	Pavilion
8:15 - 9:45 am	Action	Talk Session	Talk Room 1
8:15 - 9:45 am	Theory: Artificial neural networks	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Monday Morning Posters Visual Memory: Imagery, long-term Visual Memory: Memorability Visual Search: Attention, clinical Face and Body Perception: Parts and wholes Pre-Data-Collection Poster Session	Poster Session	Banyan Breezeway

8:30 am - 12:30 pm	Monday Morning Posters Plasticity and Learning: Perceptual learning Attention: Temporal Attention: Features, objects Object Recognition: Neural mechanisms Object Recognition: Categories Perceptual Organization: Parts, wholes, shapes and objects Spatial Vision: Crowding and eccentricity Color, Light and Materials: Optics, models Color, Light and Materials: Lightness and brightness	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break	Break	Garden Courtyard & Pavilion
10:45 am - 12:15 pm	Eye Movements: Gaze strategies	Talk Session	Talk Room 1
10:45 am - 12:15 pm	Multisensory Processing	Talk Session	Talk Room 2
12:15 - 12:30 pm	Coffee Break / Lunch (on your own) Lunch Available for Purchase in Garden Courtyard 11:30 - 2:30	Break	Garden Courtyard & Pavilion
12:30 - 2:30 pm	VSS Awards Session Davida Teller Award, Ken Nakayama Medal, Young Investigator Award, 25th Anniversary Awards	Award	Talk Room 2
2:00 - 3:00 pm	3MT® Competition for Students and Postdocs	Student	Offsite
2:30 - 4:00 pm	The AI Revolution in Visual Science Organized by the VSS Student-Postdoc Advisory Committee (SPC)	Workshop	Banyan/Citrus
2:30 - 4:00 pm	CANCELLED - Psychophysics Software with MATLAB	Satellite	Jasmine/Palm
2:30 - 5:30 pm	Readability Workshop 2: Inclusive Metrics and Design	Satellite	Blue Heron
2:30 - 5:30 pm	plenoptic: synthesizing images to understand models	Satellite	Snowy Egret
3:00 - 4:00 pm	VSS Public Lecture given by Patrick Cavanagh Using Illusions and Art to Understand the Visual Brain	Other	Offsite
4:00 - 5:30 pm	Meet the Professors	Student	Banyan Breezeway
7:00 - 10:00 pm	Demo Night	Social	Talk Room 1-2
Tuesday, May 20, 2025			
7:00 am - 10:00 pm	VSS Quiet Lounge	Lounge	Compass
7:00 am - 10:00 pm	VSS Social Lounge	Lounge	Royal Tern
7:45 - 8:45 am	Morning Coffee and Continental Breakfast	Break	Garden Courtyard & Pavilion
7:45 am - 7:30 pm	Registration Open	Registration	Grand Palm Colonnade
8:00 am - 5:30 pm	Exhibits Open	Exhibits	Pavilion
8:15 - 9:45 am	Spatial Vision: Neural mechanisms	Talk Session	Talk Room 1
8:15 - 9:45 am	Visual Memory: Imagery, memorability, long-term	Talk Session	Talk Room 2
8:30 am - 12:30 pm	Tuesday Morning Posters Eye Movements: Social, individual differences, visual preferences Eye Movements: Natural or complex tasks Decision Making: Metacognition Object Recognition: Models Object Recognition: Visual preferences Perceptual Organization: Aesthetics Perceptual Organization: Individual differences, events and relations	Poster Session	Banyan Breezeway

8:30 am - 12:30 pm	Tuesday Morning Posters Eye Movements: Cognition Face and Body Perception: Social cognition, behavioural Face and Body Perception: Social cognition, neural mechanisms Face and Body Perception: Features Face and Body Perception: Development, clinical Object Recognition: Features and parts 3D Processing: Space, coordinate frames, virtual environments Color, light and materials: Neural mechanisms, clinical Temporal Processing: Neural mechanisms, models	Poster Session	Pavilion
9:45 - 10:30 am	Coffee Break	Break	Garden Courtyard & Pavilion
10:45 am - 12:15 pm	Attention: Temporal, spatial	Talk Session	Talk Room 1
10:45 am - 12:15 pm	Scene Perception	Talk Session	Talk Room 2
12:15 - 1:15 pm	VSS Business Meeting	Business	Talk Room 2
1:15 - 2:45 pm	Connect with Industry	Networking	Blue Heron
1:15 - 2:45 pm	Lunch (on your own) Lunch Available for Purchase in Garden Courtyard 11:30 - 2:30	Break	Garden Courtyard
1:15 - 3:15 pm	phiVis: Philosophy of Vision Science Workshop	Satellite	Banyan/Citrus/Glades
2:45 - 4:45 pm	Plasticity and Learning	Talk Session	Talk Room 1
2:45 - 4:45 pm	Motion: Models, Neural mechanisms	Talk Session	Talk Room 2
2:45 - 6:45 pm	Tuesday Afternoon Posters Spatial Vision: Models Spatial Vision: Neural mechanisms Attention: Visual search Attention: Capture Undergraduate Just-In-Time 2	Poster Session	Banyan Breezeway
2:45 - 6:45 pm	Tuesday Afternoon Posters Visual Memory: Working memory and attention Visual Memory: Working memory and visual functions Visual Memory: Neural mechanism of working memory Decision Making: Perception, memory Scene Perception: Categorization, memory, clinical, intuitive physics, models Scene Perception: Spatiotemporal factors Scene Perception: Natural images, virtual environments Scene Perception: Neural mechanisms	Poster Session	Pavilion
4:45 - 5:30 pm	Afternoon Coffee and Snack	Break	Garden Courtyard & Pavilion
5:30 - 7:15 pm	Eye Movements: Perceptual advantages and disadvantages	Talk Session	Talk Room 1
5:30 - 7:15 pm	Face and Body Perception: Development, clinical, individual differences, experience	Talk Session	Talk Room 2
8:30 - 10:00 pm	Visibility: A Gathering of LGBTQ+ Vision Scientists and Friends	Social	Banyan/Citrus
10:00 pm - 2:00 am	Club Vision	Social	Talk Room 1

Symposium Sessions

Symposium: Friday, May 16, 2025, 8:00 – 10:00 am, Talk Room 1

The influence of language in higher-level visual cortex

Organizers: Oscar Woolnough[1], Alex L White[2]; [1]UTHealth Houston, [2]Barnard College, Columbia University
Presenters: Alex L White, Alexia Dalski, Marisa Nordt, Marina Bedny, Oscar Woolnough, Emily X Meschke

Vision and language are often considered to be separate cognitive systems, studied by separate research teams. But they interact when we describe verbally the things we see and when we read written text. The development of these critical skills induces changes to brain organization, such that the visual and language systems influence one another. However, it has long been controversial whether apparently linguistic activity in the ventral visual stream, such as when reading, is purely the result of top-down feedback from the language network. To the contrary, recent work has demonstrated the critical nature of local language processing within visually-responsive areas of ventral temporal cortex. Language has multifaceted influences on visual cortex, from short-term modulations of activity shaped by ongoing linguistic processing, to long-term influences on functional organization across development. As children learn to read, they adapt their visual system to efficiently process written words, a cultural invention the visual system has not directly evolved to handle. They also need to be able to learn the names of and semantic relationships between objects in their environment. As we age, this ability to retrieve object names can degrade, resulting in anomias. In this symposium we will highlight the nature of the complex, bidirectional interplay between visual and linguistic processing. Specifically, there are three main themes: 1) How do the demands of linguistic tasks modulate visual processing? Alex White will begin with a review of top-down effects during visual word recognition and task-dependent functional connectivity between the 'visual word form area' and language areas. Alexia Dalski will describe how linguistic vs perceptual tasks change the nature of how visual cortex processes both words and emojis in adults and children. 2) How does learning language affect the development of visual cortex? Marisa Nordt will demonstrate how childhood language learning longitudinally shapes the development of category-selective visual cortex and how this reorganization predicts word and face recognition performance. Marina Bedny will describe how visual cortex repurposes to subserve language for speech and braille processing in blind readers and for sign language in Deaf signers. 3) How do semantics shape the organization of visual cortex? Oscar Woolnough will describe how mapping semantic and lexical processing in the ventral visual stream can predict stimulation-induced language disruptions and avoid post-surgical reading and naming deficits. Emily Meschke will report new discoveries about multimodal semantic representations that can be evoked by visual or linguistic input. Bringing together advances in neuroimaging, lesional, and computational methods, the speakers will outline a contemporary view on the nature of the interconnectivity between language and vision.

Talk 1

INTERACTIONS BETWEEN VISION AND LANGUAGE WHEN READING WORDS

Alex L White[1], Vassiki S Chauhan[1]; [1]Barnard College, Columbia University

Reading depends on a brain region known as the "visual word form area" (VWFA) in the left ventral occipitotemporal cortex. Its function is controversial. Some researchers emphasize its bottom-up visual selectivity for words (as compared to other visual stimuli), while others attribute its activity largely to feedback from the spoken language network. I will review recent fMRI studies that vary both task demands and visual stimulus properties to investigate the nature of top-down influences in the VWFA. The data show that the VWFA is uniquely modulated by a cognitive signal that is specific to voluntary linguistic processing. This signal differs from a more generic attentional effect because it enhances the response to some stimuli and suppresses others. Moreover, functional connectivity analyses suggest a source: communication between the VWFA and a left frontal language area increases when the participant is trying to read words. These results support a hybrid model: the VWFA is inherently selective for familiar orthography, but its linguistic processing is not fully automatic. Rather, the VWFA falls under control of the language network when the task demands it.

Talk 2

A PREFERENCE FOR WORD FORMS PRECEDES THE PREFERENCE FOR LINGUISTIC PROCESSING IN THE OTS-WORDS SUBREGIONS DURING DEVELOPMENT

Alexia Dalski[1], Antonia Schulz[1], Marie Klaes[1], Max Pirsch[1], Maria Meinhardt[1], Agon Ukaj[1], Laura Faßbender[2], Gudrun Schwarzer[2], Kalanit Grill-Spector[3], Mareike Grotheer[1]; [1]Philipps-Universität Marburg, [2]Justus-Liebig-Universität Giessen, [3]Stanford University

The visual word form areas in the occipito-temporal sulcus (mOTS-words and pOTS-words) are crucial for reading, and their responses are driven by both the visual features of text and linguistic task demands. We investigated the impact of text features and linguistic processing on OTS-words responses in adults (study 1, N=15) and longitudinally in children before and after first grade (study 2, N=15): Experiment 1 compared responses to text with other visual categories (faces, objects, limbs and houses) to localize OTS-words subregions. Experiment 2 compared responses during a linguistic and a color task performed either only on emojis (children) or on emoji and text stimuli (adults). In adults, both OTS-words subregions were identified by their text selectivity in experiment 1. In experiment 2, mOTS-words showed a preference for linguistic processing and a task-stimulus interaction, preferring emojis during the linguistic task. pOTS-words showed an overall preference for emojis. In children, OTS-words subregions showed no preference for text (limbs were the preferred category) or linguistic processing before schooling. After first grade, in experiment 1, a preference for text was observed in the OTS-words subregions, allowing for their reliable identification within individuals. In contrast to adults, the sub-regions did however not show a preference for linguistic processing in experiment 2 after first grade, suggesting that text selectivity may emerge earlier in development than the linguistic task preference. This indicates that while regions involved in visual word form processing emerge rapidly with literacy training, their functional specialization for linguistic processing requires additional experience.

TALK 3

LONGITUDINAL DEVELOPMENT OF CATEGORY REPRESENTATIONS IN HIGH-LEVEL VISUAL CORTEX DURING CHILDHOOD

Marisa Nordt[1,2], Jesse Gomez[3], Vaidehi S Natu[1], Alex A Rezai[1], Dawn Finzi[1], Holly Kular[1], Kalanit Grill-Spector[1]; [1]Stanford University, [2]RWTH Aachen, [3]Princeton University

Human ventral temporal cortex (VTC) contains category-selective regions that respond preferentially to ecologically relevant categories such as faces, bodies, and words. How do these regions develop during childhood? Here, I will present our longitudinal work examining this development using both univariate and multivariate measures. In a first project, we used functional magnetic resonance imaging to measure longitudinal development of category-selectivity in school-age children. We discovered that, from childhood to the teens, face- and word-selective regions in VTC expand, but limb-selective regions shrink and lose their preference for limbs. Critically, as a child develops, increases in word-selectivity are linked to decreases in limb selectivity, revealing that limb-selectivity in VTC is repurposed into word-selectivity. These results provide evidence for cortical recycling during childhood development. In a second project, we addressed the question how distributed category representations develop during childhood and if this development relates to behavioral changes in recognition. We longitudinally measured the development of distributed responses across VTC to 10 categories in school-age children over several years. Our results reveal both strengthening and weakening of category representations with age, which was mainly driven by changes across category-selective voxels. Representations became particularly more distinct for words in the left hemisphere and for faces bilaterally. Critically, distinctiveness for words and faces across category-selective voxels in left and right lateral VTC, respectively, predicted children's word and face recognition performance. These results suggest that the development of distributed representations in VTC has behavioral ramifications and advance our understanding of prolonged cortical development during childhood.

TALK 4

THE INTERACTION OF INNATE CONSTRAINTS AND EXPERIENCE AT THE LANGUAGE/VISION INTERFACE

Marina Bedny[1], Elizabeth J Saccone[1], Mengyu Tian[1,2], Marcin Szwed[4], Piotr Tomaszewski[3], Maria Zimmermann[1,3,4]; [1]Johns Hopkins University, [2]Beijing Normal University, [3]University of Warsaw, [4]Jagiellonian University

Although most humans learn their first language via speech, visual communication is part of our evolutionary heritage and one of our earliest ways of connecting to other people in infancy (e.g., looking at faces). Studies with people who have distinctive sensory experiences (i.e., people born deaf or blind) reveal how intrinsic connections between the visual and language systems enable a broad range of adaptive behaviors. The lateral ventral occipito-temporal cortex (vOTC) sits at the junctions of the visual and language systems and in sighted people develops specialization for visual print (i.e., visual word form area.) We find that in people born blind, the lateral vOTC shows enhanced responses to spoken and written (tactile braille) language. Language responses peak in the location of the so called 'fusiform face area' and extend throughout the lateral vOTC and into early visual circuits. Deaf speakers of visuo-manual sign languages robustly recruit the lateral vOTC during language comprehension. The lateral vOTC becomes synchronized across Deaf individuals when viewing Polish Sign Language stories and shows higher synchrony for stories and sentences than lists of unconnected words. During story comprehension, the vOTC shows functional connectivity with the fronto-temporal language network. Together these data reveal the intrinsic connectivity between the lateral vOTC and language systems as well as the capacity of this connectivity to adapt to varying behavioral needs of the individual.

TALK 5

DISSOCIATING VISUAL, SEMANTIC, AND LEXICAL PROCESSING IN HUMAN VENTRAL TEMPORAL CORTEX

Oscar Woolnough[1], Kathryn Snyder[1], Meredith McCarty[1], Elliot Murphy[1], Nitin Tandon[1]; [1]UTHealth Houston

Language-dominant ventral occipitotemporal cortex (vOTC) is crucial for multiple visual language tasks, including reading and naming.

Lesions in vOTC can result in selective impairment of either reading or naming, resulting in alexia or anomia. Across multiple tasks, activity in vOTC is sensitive to visual, semantic, and lexical factors, but how do these factors interact within vOTC and with broader brain networks? And which of these factors are predictive of post-lesion deficits? I will first review evidence from our large population intracranial recordings during reading and object naming. Sensitivity to multiple lexical and semantic factors, including word frequency and concreteness, originates earliest within mid-fusiform cortex before rapidly spreading backward to earlier visual cortex, and out to frontal and parietal cortex. This reveals complex spatiotemporal interactions, both within vOTC and with the broader semantic and language networks. I will also present causal evidence, from direct cortical stimulation and post-resection deficits, for the existence of a dissociable, unimodal visual word form area and a basal temporal language area, alongside multimodal cortex underlying both reading and naming. We demonstrate an anterior-posterior split within vOTC, with stimulation of more posterior sites eliciting perceptible visual phenomena, while more anterior sites disrupt comprehension without overt visual distortions. The anteromedial vOTC has the highest probability of producing naming disruption, while posterolateral regions result in greater reading-specific disruption. Together, these results suggest the existence of multiple causally distinct but interactive functional regions within vOTC underlying visual language.

Talk 6

VISUAL-SEMANTIC REPRESENTATIONS WITHIN THE DISTRIBUTED CONCEPTUAL NETWORK OF THE HUMAN BRAIN

Emily X Meschke[1], Jack L Gallant[1]; [1]University of California at Berkeley

Most vision studies focus primarily on the visual modality, but much of our ability to reason and communicate about the world is based on lexical-semantic and conceptual information that is derived from vision, from other modalities, and from our prior experience. A prior study from our lab showed that information from visual modality-specific networks enters the multi-modal conceptual network through a distributed set of category-specific channels arranged along the anterior border of the visual cortex (Popham et al., 2021). However, little is known about how visual information is represented in the distributed conceptual network itself. To investigate this issue we used fMRI and voxelwise encoding models to compare visual-semantic representations measured during movie watching to lexical-semantic representations measured during narrative comprehension in the same participants. Comparison of the fit encoding models across experiments revealed a set of multi-modal patches located within the larger distributed conceptual network. These results provide further insights about how visual information is used to form our conceptual understanding of the world.

Symposium: Friday, May 16, 2025, 8:00 – 10:00 am, Talk Room 2

A vision scientist walks into a clinic...

Organizer: Benjamin Backus[1,2]; [1]Vivid Vision Inc, [2]SUNY College of Optometry (emeritus)
Presenters: Geoffrey K. Aguirre, Preeti Verghese, Samantha A. Montoya, Deborah Giaschi, Benjamin T Backus

Three good things can happen when a vision scientist works with patients: (1) helping individual patients preserve their vision or see better, (2) helping to develop a treatment, diagnostic tool, assistive technology, or preventative for a disease or dysfunction, and (3) discovering something about how human vision operates by comparing to typically sighted control observers. These are three very different things, but they all require access to patients. In this symposium, the presenters will describe high quality, clinically-related vision science they've done that resulted in one or more of those three good things. We will make the case for clinical work, and share our observations about how to do it. We will keep the talks short to leave time for questions and observations from the audience, including, we hope, ones that cover additional areas of clinically relevant vision science. Questions to be addressed include: What does one need to know about how clinicians see the world, in order to collaborate with them? If you're a student or early in your career, should you also get an MD or OD along with your PhD? What should you do to be a good scientific colleague to a clinician and what should you look for in a clinician colleague? Under what circumstances can one get clinic floor time? What are the personal rewards from working with patients if you're not a medical doctor? What are the rewards from working with patient community groups? Geoffrey Aguirre will discuss sensory alteration in migraine, drawing insights from how patients describe their symptoms. Preeti Verghese will describe measuring and improving depth perception in patients with AMD, and will help us understand the role of institutional support. Samantha Montoya and Stephen Engel will discuss visual snow, an understudied and debilitating condition, and factors facing early- and later-career scientists doing clinical work, respectively. Deborah Giaschi will describe how motion perception differs between normal and amblyopic vision, and what this reveals about both. Benjamin Backus will discuss how vision scientists can contribute to improved clinical testing, and what it takes to get a new test into clinical use.

Talk 1

PHOTORECEPTORS, FACEBOOK, AND SENSORY ALTERATIONS IN MIGRAINE

Geoffrey K. Aguirre[1]; [1]Univeristy of Pennsylvania

A bright or flickering light can be unpleasant for anyone, but people with migraine have markedly increased sensitivity to these stimuli, even in the period between headaches. I will describe our efforts to understand the mechanistic basis of these experiences in a typical population and in people with migraine headaches. To be informative, studies of this kind require control over both stimulus and subject.

Photoreceptor targeting with tailored spectral modulations allows us to probe the relative contribution of cone and melanopsin signals to discomfort responses, and to measure how these signals are conveyed in post-retinal ganglion cell pathways. To find subjects to view these uncomfortably bright and flickering stimuli, we use targeted Facebook campaigns to direct candidate participants to automated diagnostic classification tools. In ongoing studies we are testing if migraine produces: 1) a post-retinal amplification of discomfort signals derived from the melanopsin-containing, intrinsically photosensitive retinal ganglion cells; 2) a mis-match between the expectation and experience of temporal modulations in the visual environment. Beyond these planned hypotheses, we have found that large-scale screening of clinical populations can yield its own insights. An analysis of over 300 patient responses has led us to wonder if touch signals from the face interact with the sensation of bright light, and how the natural statistics of flicker might vary across visual eccentricity. We recommend soliciting insights from patients using an open-ended question, which in our case was: "What should scientists study about headache and sensory symptoms?"

Talk 2

PERIPHERAL STEREOPSIS IN MACULAR DEGENERATION

Preeti Verghese[1]; [1]*The Smith-Kettlewell Eye Research Institute, San Francisco, CA*

Individuals with macular degeneration (MD) have vision loss in the central retina that typically involves the fovea and significantly impacts tasks of daily living. The type of functional loss depends on how much the scotomata in the two eyes overlap. When they occur in overlapping locations in the two eyes that include the fovea, the resultant binocular scotoma can significantly impact tasks that require high-acuity vision, such as reading and recognizing faces. Moreover, even when the scotomata in the two eyes do not overlap significantly, individuals experience loss of stereopsis in the part of the visual field that corresponds to a scotoma in either eye. To determine the extent of the stereoblind zone and the potential for intact stereopsis in the periphery, we developed a method to measure local stereopsis across the visual field. Local depth sensitivity was combined to generate a stereopsis map and compared to the union of monocular scotoma maps obtained from microperimetry in each eye. The "union" prediction aligned with the stereo map, showing impaired stereopsis within this region and residual stereopsis beyond. Importantly, the stereoblind region was more extensive than the binocular scotoma defined by the overlap of the scotomata. This explains why individuals with a small binocular scotoma may be severely compromised in tasks that benefit from stereopsis, such as eye–hand coordination and navigation. A big takeaway for individuals with MD is that they can learn to utilize the residual stereopsis in their binocular periphery through training.

Talk 3

VISUAL SNOW: SYMPTOMOLOGY AND MECHANISMS

Samantha A. Montoya[1], *Stephen A. Engel*[1]; [1]*University of Minnesota*

Visual snow is a recently isolated and surprisingly common (~2-3% of the population) symptom where people continuously perceive tiny flickering dots covering their entire visual field. When combined with other commonly co-occurring symptoms it can interfere with daily activities. The details of visual snow's symptomology and its underlying mechanisms are poorly understood, limiting the development of treatments. We recently provided the first detailed measurements of visual snow's appearance: Participants adjusted parameters of a simulation to match their symptom. On average, individual elements were very small and fast, and total contrast was low (~2.5% RMS). We also recently demonstrated that snow is dependent upon spontaneous neural activity in the visual system: In people with visual snow, adapting to high contrast external dynamic noise greatly reduced the strength of the snow (transiently) to the point that it disappeared in most observers, some of whom reported seeing the world without snow for the first time. Adaptation to visual noise reduces neural responsiveness in early visual cortex, suggesting that spiking there is necessary for the snow percept. And because effects were measured while viewing a blank screen, this activity must be spontaneous. Current work uses adaptation to identify precisely where and how the spontaneous activity arises, e.g. visual snow shows a motion aftereffect. Drawing upon expertise from the visual snow community, including one of us, was crucial for developing our experimental paradigms, which may in turn provide a window onto noise suppression processes in normally sighted individuals.

Talk 4

A VISION SCIENTIST WALKS INTO A CLINIC WITH THE GOAL OF IMPROVING TREATMENT OUTCOMES FOR AMBLYOPIA AND STRABISMUS

Deborah Giaschi[1]; [1]*University of British Columbia*

Amblyopia is a common developmental disorder that is defined clinically as impaired visual acuity in a healthy eye that cannot be immediately corrected with lenses. The non-amblyopic eye usually has normal visual acuity, but binocular vision is often disrupted. Treatment options are limited and often not successful, which results in lifelong health, education, and psychosocial effects. We showed that deficits in motion perception are common in children with amblyopia with viewing through either the amblyopic or nonamblyopic eye. This suggests that the current clinical definition of amblyopia is incomplete. Motion perception deficits do not improve with standard patching treatment aimed at improving visual acuity in the amblyopic eye, but they are better ameliorated with emerging binocular treatment approaches aimed at reducing interocular suppression. Strabismus (eye misalignment) is one of the main causes of amblyopia. Surgical realignment of the eyes is a common approach to prevent the development of amblyopia and to potentially restore binocular vision, but reoperation is often necessary. We have shown that presurgical levels of both stereopsis with large binocular disparities and global motion perception predict eye alignment 1 year after surgery. This work revealed a relationship between motion perception and binocular vision that may also be important for understanding healthy vision, and modernizes the definition of amblyopia. It also provides clinicians with a fresh perspective and new tools for making treatment decisions and assessing treatment outcomes. I will share my experience with getting access to patients and interacting effectively with clinicians and their staff.

Talk 5

VISUAL PERIMETRY: MORE INTERESTING THAN YOU THINK

Benjamin T Backus[1,2]; [1]Vivid Vision Inc, San Francisco, CA, USA, [2]SUNY College of Optometry (emeritus)

What could be more boring than visual field testing, which consists of badly measuring the luminance increment threshold of a small stimulus at each of 54 retinal locations, one eye at a time? Our initial goal was simply to port this test to inexpensive mobile VR headsets. However, the test is difficult to take, as it demands several minutes of steady fixation, unrelenting visual attention, and a fixed internal criterion for responding. Thus, eye doctors keep visual field tests short to avoid fatigue. But visual behavior and cognition are better understood now than in the 1970's when the conventional test was developed, so we updated the test to make it longer and collect more data. With the Yvonne Ou lab at UCSF, we demonstrated that 20 out of 20 patients with glaucoma were each able to perform ten 18-minute tests at home. 100% adherence to anything done at home is remarkable. The scientific benefit of this approach is that visual sensitivity can be measured with better precision. That in turn makes the test useful for pharmaceutical companies, which are using it in clinical trials of new therapeutics. Precision is also essential for research on the relationship between "structure" (which is measured quickly with OCT) and "function" (which is measured laboriously with visual field tests) which may allow OCT to replace visual field testing. Use of the test in direct patient care remains a goal. Many patients find significant meaning in contributing to research on their condition.

Symposium: Friday, May 16, 2025, 10:30 am – 12:30 pm, Talk Room 1

Selective Visual Attention, Alive and Well!

Organizers: Marisa Carrasco[1], Miguel Eckstein[2], Anna C Nobre[3]; [1]New York University, [2]University of California Santa Barbara, [3]Yale University
Presenters: Marisa Carrasco, Sam Ling, Farran Briggs, Miguel Eckstein, Anna C Nobre

Visual attention is critical for perception, enabling prioritized and selective information processing in human and non-human primates. This symposium encompasses a broad range of studies on covert visual attention across spatial [Carrasco, Eckstein, Briggs], feature [Ling, Briggs], and temporal [Nobre] domains. Studies employ diverse methods and levels of analysis, including psychophysics [Carrasco, Briggs, Nobre], multi-unit recording [Briggs], MEG/EEG [Nobre], neurostimulation [transcranial-magnetic stimulation, TMS, Carrasco], and computational modeling using Convolutional Neural Networks [CNNs; Eckstein]. Central and common to these studies is the careful definition, operationalization, and manipulation of attention, enabling us to isolate and characterize the roles and mechanisms of different types of attention. Together, these methods, level of analyses, and approaches advance our understanding of visual attention at both behavioral and neural levels. The symposium will begin with an overview of the central questions addressed by each of the five speakers, including the types of attention studied, levels of analysis employed, the strengths and limitations of the respective protocols and methods, and the theoretical contributions. Each presentation will be followed by a brief Q&A session. Presentations: Marisa Carrasco discusses how endogenous (voluntary) and exogenous (involuntary) covert spatial attention shape human perception by modulating visual processing in distinct ways. Her methodological approach isolates the perceptual effects and neural correlates of each attention type. Sam Ling explores how feature-based attention alters population responses within the human visual cortex (V1-V3), altering the contrast response and spatial-frequency tuning for attended items. Farran Briggs investigates neuronal mechanisms of spatial and feature attention, using neurophysiological measurements at the neuronal and circuit scale in non-human primates performing discrimination tasks that involve shifting spatial and feature attention. Miguel Eckstein examines how feedforward Convolutional Neural Networks (CNNs), despite lacking built-in attention mechanisms, exhibit behavioral signatures of human-like covert spatial attention. Moreover, they show some "neuronal unit" responses analogous to documented physiological signatures of attention, as well as other unit responses not yet observed in neurophysiological studies. Kia Nobre presents recent MEG and EEG studies showing how goals, temporal expectations and attention modulate sensory and motor processing stages, enabling the prioritization of items in perceptual and working-memory tasks in humans. The symposium will conclude with a Panel discussion summarizing the contributions of each talk through concise answers to the posed questions and facilitating an open dialogue with the audience, enhancing our collective understanding of visual attention.

Talk 1

HOW SPATIAL ATTENTION SHAPES PERCEPTION

Marisa Carrasco[1], Antonio Fernández[1], Laura Dugué[2,3], Nina Hanning[1,4]; [1]New York University, [2]Université Paris Cité, [3]CNRS, [4]Humboldt-Universität zu Berlin

We investigated how endogenous (voluntary) and exogenous (involuntary) covert spatial attention differentially modulate visual perception, at behavioral and neural levels. To understand the effects of endogenous and exogenous attention, we conducted head-to-head comparisons using psychophysical, computational, neuroimaging (fMRI), and neurostimulation (TMS) methods. In each empirical study, we kept participants, stimuli, and discrimination tasks constant, varying only the attentional cue manipulation. This approach allowed us to isolate the distinct effects of each type of attention. First, I highlight findings from psychophysical experiments showing how endogenous and exogenous covert spatial attention differentially affect performance in tasks mediated by basic visual dimensions (e.g., contrast sensitivity) and their featural representations (using reverse correlation methods). Next, I present findings from fMRI studies, revealing distinct neural correlates of endogenous and exogenous attention. For example, in a pre-cue condition, endogenous attention increasingly modulated stimulus-evoked activity along the visuocortical hierarchy (V1 to LO1), with largest effects in higher cortical areas; however, this modulation remained constant in a post-cue condition. In contrast, exogenous attention produced consistent modulation across the visual hierarchy in a pre-cue condition but showed no modulation in a post-cue condition. Lastly, I discuss how we used TMS experiments to uncover that brief disruption of V1/V2 activity eliminates exogenous attention effects, whereas disruption of the (human homologue of) the right frontal eye fields (rFEF+) significantly reduces endogenous attention effects. Together, these studies reveal how endogenous and exogenous attention shape perception by selectively modulating the processing of basic visual dimensions across different brain regions.

TALK 2

HOW FEATURE-BASED ATTENTION ALTERS REPRESENTATIONS IN HUMAN VISUAL CORTEX

Sam Ling[1], Joshua Foster[1], Feiyi Wang[1], Luis Ramirez[1]; [1]Boston University

I will describe human neuroimaging research that examines the modulatory effects of feature-based attention on two fundamental building blocks in vision: contrast response and spatial frequency tuning. To assess the influence of attention on the gain of the contrast response function (CRF), we measured BOLD activity evoked by a probe stimulus in one visual hemifield while participants attended to the probe orientation (attended condition), or to an orthogonal orientation (unattended condition), in the other hemifield. To measure CRFs in visual areas V1-V3, we parametrically varied the contrast of the probe stimulus. In all three areas, feature-based attention increased contrast gain, improving sensitivity by shifting CRFs toward lower contrasts. In extrastriate cortex, we also found an increase in response gain that was greatest in the parafovea. In a second study, we used fMRI to measure population spatial-frequency tuning (pSFT), which allowed us to examine how feature-based attention alters the response profiles of neural populations in early visual cortex. Participants were cued to attend to one of two spatially competing letter streams, each defined by low or high frequency content. Concurrently, we measured pSFT in a task-irrelevant hemifield to examine how the known spatial spread of feature-based attention influenced the SF tuning properties of neurons sampled within a voxel. We discovered that attention elicited attractive shifts in SF preference, towards the attended SF. Taken together, these two studies reveal reliable signatures of feature-based attention on the gain and tuning preferences of population responses within human visual cortex.

TALK 3

CORTICAL AND THALAMIC CONTRIBUTIONS TO SPATIAL AND FEATURE ATTENTION

Farran Briggs[1,2], Alesandra Martin[1], Sabrina Mai[1], Shraddha Shah[1]; [1]University of Rochester, [2]National Eye Institute

Much is known about how shifting attention between locations in visual space (spatial attention) or across different stimulus features (feature attention) alters the activity of visual neurons. However, the mechanisms guiding these effects at the scale of individual neurons, neuronal circuits, or neuronal populations remain a mystery. We compare attention effects across early visual system structures and across spatial and feature attention tasks as a way to infer similarities and differences in underlying mechanisms at the neuronal and circuit scale. We trained non-human primates (NHPs) to perform feature discrimination tasks involving shifting spatial attention and switching attention between multiple stimulus features (contrast, orientation, and color). Neuronal recordings were made with multi-electrode arrays implanted in the visual thalamus or primary and secondary visual cortex. Measurements of neuronal feature tuning were made independent of the attention task. Our observations suggest attention effects depend critically on areal organization, measurement scale, neuroanatomical connections, and neuronal feature selectivity. First, attentional modulation of multi-unit activity is quite robust in visual cortex but absent in visual thalamus. Second, attentional modulation of correlated variability among neurons in visual cortex depends on anatomical connectivity: attention strongly reduces correlated variability among neurons connected via excitatory synapses. Third, attentional modulation of cortical neurons depends on the match between their feature selectivity and the task-relevant feature, even when spatial attention is held constant. These findings show that attention effects in NHPs depend critically on factors that span anatomical scales and feature encoding dimensions, necessitating exploration across scales and task protocols.

TALK 4

CONVOLUTIONAL NEURAL NETWORK UNCOVERS EMERGENT NEURALLY-PLAUSIBLE MECHANISMS MEDIATING HUMAN-LIKE COVERT ATTENTION

Miguel Eckstein[1], Sudhanshu Srivastava[1]; [1]University of California Santa Barbara

Covert spatial attention allows the brain to select different regions of the visual world without concurrent eye movements. Cues predictive of a target location orient covert attention and improve perceptual performance. In most computational models, researchers explicitly incorporate an attentional mechanism that alters processing at the attended location (gain, noise reduction, divisive normalization, biased competition, Bayesian priors). Here we assess the emergent neuronal mechanisms of feedforward Convolutional Neural Networks (CNNs), with no built-in attention mechanism, which shows behavioral signatures of human-like covert attention. We use a system-neuroscience-inspired approach to analyze 1.8M computational units ("neurons") of CNNs trained on a spatial cueing task. Consistent with neurophysiology, we show early layers with retinotopic units separately tuned to the target or cue, and later layers with units with joint tuning and increased influence of the cue on target responses. We find cue-inhibitory and location-opponent units unreported by neurophysiologists. The cue influences the mean neuronal unit responses and changes target sensitivity through three mechanisms: cue-weighted summation and opponency across locations, and interaction with the thresholding Rectified Linear Unit (ReLU). CNN computational stages mirror a Bayesian ideal observer (BIO), but with more gradual transitions, and also include computations (the opponency and ReLU interaction) distinct from the BIO and not discussed in the physiology literature. Together, the findings establish a likely system-wide characterization of the brain computations that mediate the behavioral signatures of covert attention and provide a theoretical framework for neurophysiologists and vision scientists.

Talk 5

MOMENTS OF ATTENTION

Anna C Nobre[1]; [1]Yale University

Visual selective attention functions – that anticipate, select, prioritize, and prepare contents – are infrastructural and core to guiding adaptive perception and behavior. Decades of psychophysical and neuroscientific research across levels of analysis have advanced our understanding of how internal signals related to experience and goals flexibly and proactively modulate processing according to receptive field properties of relevant anticipated stimuli. The next challenge is to situate attention functions within the dynamic flux of natural experience. Active agents move through changing environments, with evolving goals and expectations. Research on selective temporal attention investigates how the brain focuses on specific moments of predictable or relevant events to guide perception, action, or mnemonic access. From the diverse and multilayered sensory stream, the brain extracts various types of temporal structures to tune the processing of events according to their timing. I will discuss recent studies in our group using MEG and EEG that show how goals and temporal expectations modulate sensory and motor stages of processing to prioritize items in perceptual and working-memory tasks in humans. Modulatory functions can proceed in tandem and independently. The findings highlight the robust impact and flexibility of temporal attention, implicating short-term and long-term sources of predictions that interact with various receptive-field properties to enhance perception and memory through multiple mechanisms. Time is pressing for the fourth dimension of attention research.

Symposium: Friday, May 16, 2025, 10:30 am – 12:30 pm, Talk Room 2

25 Years of Seeing ‘Stuff’—Advances and Challenges in Material Perception

Organizers: Vivian C. Paulun[1], Roland Fleming[2]; [1]Massachusetts Institute of Technology, [2]Justus Liebig University Giessen and Center for Mind, Brain and Behavior (CMBB), Universities of Marburg, Giessen and TU Darmstadt
Presenters: Edward Adelson, Roland Fleming, Vivian Paulun, Wenyan Bi, Bei Xiao, Maarten Wijntjes

Vision is more than just “know[ing] what is where by looking” (Marr, 1982). A central aspect of visual perception is working out the physical properties of surfaces and objects. To interact effectively with the world, we need to recognize whether something is made of wood, marble, chocolate or silk and whether it is soft, slippery, brittle or elastic. From delicate porcelain plates to sticky honey, from polished oak wood to fluffy merino wool—our world is filled with an incredible diversity of materials. Materials are ubiquitous in our world so material perception is deeply intertwined with all levels of visual processing from color and texture to shape and motion, from scenes and intuitive physics to visually-guided actions. In the past 25 years we have seen substantial progress in our understanding of material perception and shed light on the underlying cues and computations. Novel insights and experimental results have fueled theoretical debates that go far beyond material perception, spanning the role of image statistics and heuristics as well as analysis-by-synthesis and mental physics simulation in visual perception. The aim of this symposium is to review the most exciting advances in the field, identify common themes as well as discrepancies and highlight the most pressing challenges in material perception research today. The symposium will include six talks from speakers at various career stages, with complementary backgrounds and scientific approaches. The presentations will showcase work investigating material perception using a breadth of methodologies spanning psychophysics, neuroimaging, computational modeling and AI, computer graphics and art history. Edward Adelson will kick off the symposium by setting material perception in context, introducing fundamental questions about how we perceive the material world and drawing important parallels between human and computer vision. Roland Fleming will provide a broad overview of new directions in material perception research enabled by technological developments and theoretical advances. Next, Vivian Paulun will focus in on the perception of dynamic materials, by asking how the visual system draws rich inferences about mechanical properties like viscosity, elasticity and from the way materials flow, bounce or deform. Wenyan Bi will describe how generative AI models can be used to investigate the cues underlying judgments about mechanical properties like cloth stiffness. Bei Xiao will discuss the (mis-)alignment of visual and semantic representations in material perception. Finally, Maarten Wijntjes will showcase what we can learn about material perception from studying the art of material depiction. Each talk will be 20 minutes long, including 3 minutes for Q&A. There is little overlap of this symposium with the regular VSS program or any symposium in the past 15 years. Unlike other sessions at the conference, this symposium is dedicated solely to materials, and it will allow speakers to zoom out to the bigger picture questions and theoretical frameworks. We expect this topic to be of interest to VSS members from a multitude of subfields, such as the perception of color, shape, motion, objects, events, relations, and scenes, intuitive physics, perception and action, and visual neuroscience.

Talk 1

ON SEEING STUFF

Edward Adelson[1]; [1]Massachusetts Institute of Technology

It is gratifying to know that this paper has had a lasting impact. It contains no experimental results and no theorems, and therefore was not published in a peer reviewed journal. The paper is an invitation to share in two key activities, which are looking and thinking. The paper expresses the hope that it might be possible to understand material perception by starting with a small number of simple concepts and measurements. Is this still viable? The success of deep learning suggests that visual perception involves representations and computations of extraordinarily high dimension, and that we may be fooling ourselves when we imagine that we can describe perception in terms of simple principles. Nonetheless, we have no choice but to press forward and extract as much understanding as we can. We have to keep looking and thinking and experimenting and introspecting.

TALK 2

PROGRESS ON SEEING STUFF

Roland Fleming[1]; [1]Justus Liebig University Giessen and Center for Mind, Brain and Behavior (CMBB), Universities of Marburg, Giessen and TU Darmstadt

Over the last 25 years we have seen massive progress in our understanding of material perception. When we started out, computer graphics was just about reaching physical realism and the range of material properties we could feasibly investigate was extremely limited. Our conceptual frameworks were legacies of lightness and colour constancy research. In the meantime, the scope of scientific questions our field is addressing has exploded to encompass topics spanning from the role of low- and mid-level cues in the estimation of diverse optical and mechanical properties, to the relationship of material perception to high-level concepts, learning, language, action and multimodal representations. Nowadays, material perception research even poses challenges for widely-held tenets of vision science, like inverse optics. I will summarise the massive changes we have witnessed in material perception research and highlight the impact of recent and emerging developments like crowdsourcing, interactive computer graphics, motion capture, deep learning and generative AI. In the process I will describe some major theoretical advances in our understanding of material perception and the consequences these have for how we should think about vision science more broadly.

TALK 3

BOUNCING, BENDING, BUBBLING: SEEING DYNAMIC STUFF

Vivian Paulun[1]; [1]Massachusetts Institute of Technology

We can draw rich inferences from observing how materials flow, deform, and interact, from how gooey honey slowly drips off a spoon, powdery sand trickles between someone's fingers or how jelly bouncily wobbles back and forth on a plate. Visually inferring materials and their properties from dynamic interactions yields powerful perceptual effects despite unique computational challenges. Computationally, the task is ill-posed: The observable behavior of any material depends on a multitude of factors, its physical properties and the forces applied. The motion of a liquid in a bowl depends on its viscosity as well as the movement of the whisk. How a box sinks into a couch cushion depends on the properties of both, the cushion and the box. Because the observed behavior can vary infinitely, this task cannot be solved using simple pattern recognition—an immense challenge to state-of-the-art AI. Yet, humans robustly use dynamic information to visually determine the kind of 'Stuff' in front of them, e.g., liquid, solid, or jelly, and estimate its mechanical properties, e.g., elasticity or viscosity. When brought into conflict with texture cues, dynamic material information clearly dominates our percept. Despite its inherent ambiguity, the brain can infer materials from dynamics with minimal information, e.g. indirectly from how other objects move in response to an interaction. I will characterize such perceptual effects and draw conclusions about the cues and computations underlying dynamic material perception. Furthermore, I will show evidence for distinct neural representation of dynamic 'Stuff' in both the ventral and dorsal visual pathways.

TALK 4

INTUITIVE PHYSICS UNDERLIES MATERIAL PERCEPTION: COMPUTATIONAL, PSYCHOPHYSICAL, AND NEURAL EVIDENCE

Wenyan Bi[1], Ilker Yildirim[1]; [1]Department of Psychology, Yale University

From the wrinkles and folds a soft object makes, how do we see, not just these changes in geometry, but also physical material properties, including their mass and stiffness? A common view states that the brain relies on high-level image and motion statistics that differentiate the degrees of these physical properties (e.g., discriminating a soft cloth from a stiff cloth). I'll counter this view with an alternative framework, in which the brain inverts an internalized, physics-based generative model to arrive at the scene-level causes underlying visual inputs. In this account, material perception is cast as posterior inference of physical properties under a generative model of "soft body dynamics" (a game-engine style description of how non-rigid materials move and react to external forces) and simple graphics to project these scenes to sensory measurements. I'll present a computational model that implements this framework and evaluate it in psychophysical and neural experiments. First, I'll show that the physics-based model explains both the successes and failures in material perception across multiple match-to-sample tasks. It outperforms a performant DNN model that solves material perception by acquiring high-level image and motion statistics for discriminating the physical properties. Next, I'll evaluate these models in a new fMRI experiment. I'll present evidence for a double dissociation, where a set of higher-order frontoparietal regions aligns with the physics-based model and an occipitotemporal region aligns with DNN. Together, these findings suggest that visual material perception transcends image statistics to also involve intuitive physics—formalized as probabilistic simulations of soft-body dynamics.

TALK 5

PROBING VISUAL AND SEMANTIC REPRESENTATIONS IN MATERIAL PERCEPTION USING PSYCHOPHYSICS AND DEEP LEARNING

Bei Xiao[1], Chenxi Liao[1], Masataka Sawayama[2]; [1]Department of Computer Science, American University, [2]University of Tokyo

The look and feel of materials are an integral part of our daily experience. Seeing and understanding materials allows us to interact with materials in diverse tasks. In the past decades, significant progress has been made on visual inference of material properties. Relatively little is known about the semantic representation of materials. Much of this knowledge is represented symbolically in language, which allows us to articulate material qualities and appearances. The immense diversity, complexity, and versatility of materials present challenges in verbalization. In this talk, I will discuss our recent progress in material perception using psychophysics and deep learning methods. First, I will introduce a deep generative framework to synthesize realistic and diverse material appearances and learn an interpretable latent space that can capture perceptually relevant visual information. I will discuss how the learned latent space can be used to probe human perception of translucency. In the second part, I will discuss how we extend our framework to elucidate the relationship between language and vision in material perception in both familiar and unfamiliar ambiguous materials. By comparing the representations derived across modalities, our results reveal the alignment and misalignment of vision-language connection and underscore the importance of leveraging the vision and semantic features to reveal behavioral relevant features in material perception. I will close by presenting our latest results on modeling the relationship between material categorization and material discrimination, and the future directions of integrating our approach to discover neural correlates of material perception and related cognitive tasks.

Talk 6

25 YEARS OF SEEING STUFF? 2500 YEARS OF DEPICTING STUFF!

Maarten Wijntjes[1]; [1]Delft University of Technology

About 2500 years ago, "[Parrhasius] entered into a pictorial contest with Zeuxis, who represented some grapes, painted so naturally that the birds flew towards the spot where the picture was exhibited. Parrhasius, on the other hand, exhibited a curtain, drawn with such singular truthfulness, that Zeuxis, […] haughtily demanded that the curtain should be drawn aside to let the picture be seen." With an important role for optical (grapes) and mechanical (curtain) material properties, Pliny the Elder accounts the birth of material depiction. The distinction between optical and mechanical properties is not the only parallel to be drawn with "On seeing stuff". Adelson (2001) argues that to understand material perception, one needs to understand how images are made. In other words, to understand vision, one needs to understand depiction. Whereas in vision the dichotomy between stuff and things dominates, in depiction the more relevant dichotomy is that between stuff and space. These two 'formal elements' of depiction have interesting histories related the invention of new media (stuff) and new projection techniques (space). This contribution will discuss de vision and depiction of material properties through time and medium. I will discuss the handling of highlights, demonstrate a 'distant viewing' approach to the history of material depiction and compare the medium of engravings with oil paint.

Symposium: Friday, May 16, 2025, 1:00 – 3:00 pm, Talk Room 1

From insects to fish to mammals: Active vision in non-primate organisms

Organizers: Lisa Kroell[1], Lisa Fenk[1]; [1]Max Planck Institute for Biological Intelligence
Presenters: Lisa Fenk, Lisa Bauer, Eva Naumann, Jason Kerr, Philip Parker

Certain questions have occupied active vision researchers for decades: How is visual perception modulated yet ultimately undisturbed by frequent movements of the eye, head and body? How is external world motion distinguished from motion signals caused by effector movements? And what functions may eye movements serve beyond the foveation of relevant information? Within the field of human and non-human primate vision, these questions are extensively discussed by the VSS community. Among the remaining 99.98% of living animal species, however, a universe of active visual behaviors waits to be discovered. While strikingly similar questions arise across species, many non-primate organisms face a markedly different set of preconditions. They might, for instance, possess uniform acuity across a broad visual field, need to account for the visual consequences of both gait and flight, or even be able to move their eyes independently of one another. Simultaneously, the study of non-primate organisms opens up a world of methodological possibilities: Depending on the species, neuronal activity can be controlled using optogenetics, neurophysiological recordings can be performed during unconstrained behavior and, due to the reduced complexity of the underlying neural systems, visual circuits can be dissected in remarkable detail. We gathered five active vision specialists from four international institutions who harness this exceptional toolset to investigate visual perception in behaving insects, fish and non-primate mammals. Our speakers—all of them first-time VSS attendees—represent a variety of career stages, from graduate student to early independent researcher to full professor. Despite their use of diverse model systems, they pursue a common aim: understanding how movements of the eyes, paws, wings, tails or fins shape and even support vision. First, Lisa Fenk will introduce the audience to the recently discovered retinal movements in fruit flies. Through video-based eye tracking and whole-cell patch clamp recording, she investigates the properties of these movements, as well as the consequences of spontaneous movements for visual processing in general. Lisa Bauer will subsequently venture into underwater habitats and describe how spontaneous and stimulus-induced saccadic eye movements in zebrafish alter the activity of visual neurons. Eva Naumann will follow up by showcasing the vast potential of combining several sophisticated optical techniques: Through two-photon microscopy and 3D holographic optogenetic photostimulation, she pinpoints how visual signals are translated to motor commands during the zebrafish optomotor response. Next, Jason Kerr will demonstrate that freely pursuing ferrets execute stabilizing eye movements as well as saccades that align optic flow-fields with the high-resolution area centralis and with the intended direction of travel. Philip Parker will conclude by showing that neurons in primary visual cortex of unconstrained mice demonstrate a saccade-locked coarse-to-fine processing sequence akin to what is observed in primates. By combining this diverse yet highly complementary expertise, we hope to open the conference up to novel audiences, stimulate future comparative work and, potentially, uncover unifying principles of active vision across species.

Talk 1

NEURAL MECHANISMS FOR ACTIVE EYE MOVEMENTS IN DROSOPHILA

Lisa Fenk[1]; [1]Max Planck Institute for Biological Intelligence

Our work focuses on understanding two fundamental aspects of active vision. How do brains ignore aspects of the changing sensory stream that are not informative for the task at hand? And, perhaps more remarkably, how do brains actively move their sensors to create sensory patterns of activity that enhance their perception of the world? We use the Drosophila visual system to study both of these sensory challenges in a genetic model organism. During fast flight turns, we observe motor-related inputs to Drosophila visual cells whose properties suggest that they briefly abrogate the cells' visual-sensory responses. Rather than a wholesale shutdown of the visual system during flight turns, fly visual neurons receive targeted inputs that are precisely calibrated to abrogate each cell's expected visual response, suggesting that they function as "efference copies." While flies suppress the perception of self-generated visual motion during flight turns, they also purposefully generate visual motion in other circumstances. We recently discovered that fruit flies move their retinas via tiny muscles, both seemingly spontaneously and in response to visual motion. These movements share surprising similarities with our vertebrate eye movements. We now leverage fly retinal movements as a relatively simple model to examine the cellular underpinnings of active visual processing. We aim to understand how fly eye movements are controlled neuronally, how the brain processes input from moving eyes, and how visual perception ultimately benefits from eye movements.

TALK 2

UNDERSTANDING VISUAL PROCESSING DURING SACCADES USING ZEBRAFISH

Lisa Bauer[1], J. C. Donovan[1], H. Baier[1]; [1]Max Planck Institute for Biological Intelligence

Saccadic eye movements are fundamental to vertebrate visual perception, yet it is still unclear exactly how circuits smoothly process the resulting shifts in visual input. The zebrafish model, with its experimental tractability, offers an ideal system for investigating the underlying sensorimotor circuits. Zebrafish perform spontaneous and visually induced saccades as early as 4 days post fertilization (dpf). Here, we combined eye tracking and two-photon calcium imaging to investigate the neuronal correlates of saccades in larval zebrafish (6-8 dpf). We focus on the optic tectum (OT), the fish equivalent of the mammalian superior colliculus (SC), as well as the largest visual brain area. Using a two-photon microscope custom-modified with a remote focusing path to enable rapid multi-plane imaging, we record single-cell resolution neuronal activity across the OT at 5 volumes per second. Even in the absence of visual stimuli we find neurons in the OT that showed increased activity correlated with spontaneous saccades. To investigate how visual stimuli are integrated, we recorded from the same neurons during various visual stimulus paradigms. Our findings reveal that most spontaneous saccade-correlated neurons in the OT respond similarly regardless of visual environment. Moreover, spontaneous saccade-responding neurons are a subset of visually induced saccade neurons. Notably, while many neurons were active around the time of a saccade, certain neurons' activity peaked before, suggesting a role in anticipatory motor planning. Our results underscore the effectiveness of the larval zebrafish as a model for functional investigation, enabling experimental approaches that are challenging to implement in primate models.

TALK 3

FUNCTIONAL CONNECTIVITY CONSTRAINED SIMULATIONS OF VISUOMOTOR CIRCUITS IN ZEBRAFISH

Eva Naumann[1]; [1]Duke University School of Medicine

Visual motion processing in the brain is critical for generating movements with appropriate speed and vigor. However, single-cell mechanistic characterizations in vertebrates remain challenging due to the complexity of mammalian brains. The translucent larval zebrafish provides an important model for studying brain-wide visual computations at the cellular level. A key visuomotor transformation in zebrafish is the optomotor response (OMR), where fish stabilize their body position in response to optic flow. The underlying neural circuits involve the retinorecipient pretectum (Pt) and descending motor command neurons in the midbrain nucleus of the medial longitudinal fasciculus (nMLF). By modeling these circuits in a physics-based neuromechanical simulation, we show that the functional connections between these populations are critical for accurate speed adaptation in the simulation. To causally map how neurons interact and compute visual motion information, we integrated volumetric two-photon microscopy with simultaneous 3D holographic optogenetic photostimulation during visual stimulation and tail tracking. Using these all-optical methods, we uncovered the cellular level Pt-nMLF functional connectivity, defined as a neuron's functional identity or 'receptive field' and its functional role in the circuit's computation or 'projective field'. Our findings reveal that specific visually responsive Pt subtypes differentially modulate specific nMLF neural activity, forming correlation-based functional connectomes that guide motor output. We applied these experimentally derived functional connectivity weights to update our model, improving its behavioral response to variable-speed visual stimuli. These results highlight how all-optical methods can map functional connections to provide new insight into brain-scale sensorimotor transformations in vertebrates.

TALK 4

HOW FREELY MOVING ANIMALS MOVE THEIR EYES DURING PREDATOR/PREY INTERACTIONS AND NAVIGATE

Jason Kerr[1]; [1]Max Planck Institute for Neurobiology of Behavior

During prey pursuit, how eye-rotations, such as saccades, enable continuous tracking of erratically moving targets while simultaneously enabling an animal to navigate through the environment is unknown. To better understand this, we measured head and eye rotations in freely running ferrets during pursuit behavior. By also tracking the target and all environmental features we reconstructed the animal's visual fields and their relationship to retinal structures. In the reconstructed visual fields, the target position clustered on and around the high acuity retinal area location, the area centralis, and surprisingly this cluster was not significantly shifted by digital removal of either eye saccades, exclusively elicited when the ferrets made turns, or head rotations which were tightly synchronized with the saccades. Here we show that, while the saccades did not fixate the moving target with area centralis, they instead aligned the area centralis with the intended direction of travel. This also aligned the area centralis with features of the optic flow pattern, such as flow direction and focus of expansion, used for navigation by many species. While saccades initially rotated the eyes in the same direction as the head turn, saccades were followed by eye rotations countering the ongoing head rotation, which reduced image blur and limited information loss across the visual field during head-turns. As we measured the same head and eye rotational relationship in freely moving tree shrews, rats and mice, we suggest these saccades and counter-rotations are a generalized mechanism enabling mammals to navigate complex environments during pursuit.

TALK 5

NEURAL CODING AND CIRCUITRY OF ACTIVE VISION IN MICE

Philip Parker[1]; [1]Rutgers University

Visual perception is an active process: we constantly move our eyes, head, and body to fully perceive the world around us. Research over the last century has yielded incredible insight into how neurons in the visual system process information, yet our understanding is largely limited to conditions of ‘passive’ rather than ‘active’ vision. This is primarily due to the fact that experiments are traditionally performed under physically restrictive conditions that prevent the natural sensory consequences of movement (e.g. head-restrained animals presented with isolated stimuli). How does visual processing occur under ethological conditions, where animals freely explore complex visual environments in a goal-dependent manner? We addressed this question by performing visual physiology in freely moving mice, recording the activity of more than 100 V1 neurons while measuring the visual input with a head-mounted camera. In addition to mapping spatiotemporal receptive fields in freely moving animals, we found that V1 neurons jointly code for eye and head position - a coding scheme useful for performing retinocentric-to-egocentric reference frame transformations. We also found that V1 neurons fire in a sequence around saccadic eye movements according to increasing spatial frequency preference, consistent with coarse-to-fine models of visual scene processing. To address how these phenomena relate to goal-directed behavior, ongoing work in the lab is focused on the neural circuits and coding underlying visual distance estimation. Together, we are moving toward a greater understanding of how our visual systems operate under real-world conditions.

SYMPOSIUM: FRIDAY, MAY 16, 2025, 1:00 – 3:00 PM, TALK ROOM 2

Model-optimized stimuli: more than just pretty pictures

Organizers: William Broderick[1], Jenelle Feather[1]; [1]Center for Computational Neuroscience, Flatiron Institute
Presenters: Ruth Rosenholtz, William F. Broderick, Arash Afraz, Jenelle Feather, Binxu Wang, Andreas Tolias

Experiments in vision science rely on designing the appropriate stimulus set to test specific properties of visual systems. Data collection time is inherently limited, so choosing stimuli that get the most "bang for your buck" when collecting behavioral or neural data is imperative. While early experiments relied on simple stimuli such as points and light and visual gratings to test specific hypotheses about visual representations, the advent of large-scale neural recordings and large benchmark datasets encouraged many researchers to make a jump to natural stimuli such as photographs and video, where one could test predictive models on these large datasets. However, both of these approaches have limitations. A model based purely on simple synthetic stimuli may not generalize to natural environments, but when relying solely on natural stimuli, the hypotheses about the underlying representations are often unclear and results can be muddled with stimulus correlations. We now have tools to combine these approaches, allowing for the design of complex stimuli that test specific hypotheses, guide responses, or decouple confounding underlying properties of the stimuli. In this symposium, we cover recent progress in utilizing model-optimized stimuli to probe visual perception and cognition. Ruth Rosenholtz will describe how stimulus synthesis facilitated the development of a model of peripheral crowding, developing intuition about its predictions and correcting conceptual errors along the way. William Broderick will describe "plenoptic," an open-source python package that offers a framework for synthesizing model-optimized stimuli such as metamers and eigendistortions, enabling the broader vision science community to use these approaches for their own models. Arash Afraz will detail work with "Perceptography", a technique that combines behavior, optogenetics, and machine learning tools to design stimuli that capture the subjective experience induced by local cortical stimuluation. Jenelle Feather will discuss how behavioral experiments with model metamers and other types of model-optimized stimuli can reveal differences between biological and artificial neural networks, and show how these methods can point towards model improvements. Binxu Wang will describe ongoing work using artificial neural networks (ANNs) to fully explore the dynamic range of macaque visual cortical neurons, enabling improved model performance. Andreas Tolias will describe the development and use of "digital twins", performing in-silico experiments on brain foundation models to improve our understanding of the brain while simultaneously developing more comprehensible, energy-efficient AI models. Choosing the right stimuli to test a hypothesis is a fundamental aspect of experimental design, and this symposium will be of general interest to VSS members interested in testing the properties of neural populations underlying visual perception. This symposium will give attendees an overview of the various uses of these stimuli, and inspire them to apply such methods in their own research.

TALK 1

SYNTHESIZING MODEL PREDICTIONS SUPERCHARGES UNDERSTANDING

Ruth Rosenholtz[1], Benjamin Balas[2]; [1]NVIDIA, [2]North Dakota State University

At VSS 2008 we first presented our model of peripheral crowding. Serendipitously, we captured existing intuitions about crowding with Portilla & Simoncelli’s (2001) texture analysis/synthesis algorithm. Doing so allowed us to synthesize images in which distorted and retained image structure revealed the visual information available, ambiguous, or lost according to the model (Balas et al, 2009; Freeman & Simoncelli, 2011; Rosenholtz, Huang, & Ehinger, 2012). Many researchers have similarly used a wide range of image-computable models to visualize predictions, and the value of this approach cannot be overstated. Image-computability combined with synthesized outputs supports rapid development of predictions for a wide range of stimuli and tasks. Formal experiments, almost as an afterthought, can quantify those intuitions as needed, but the instant, informal psychophysics afforded by this framework is often sufficient. In our case, this enabled us to quickly build novel intuitions about crowding, search, reading, scene perception, maze-solving, change blindness, illusions, choice of fixations, and vision for action. This methodology helped correct conceptual errors; for example, that nearby flanker things crowd (a transitive verb) a target thing, unless some other mechanism intervenes to relieve crowding. Instead, model syntheses clearly reveal the importance of stimulus and task, and the generality of crowding well beyond the original empirical phenomena.

TALK 2

PLENOPTIC: A PYTHON LIBRARY FOR SYNTHESIZING MODEL-OPTIMIZED VISUAL STIMULI

William F. Broderick[1], Edoardo Balzani[1], Kathryn Bonnen[2], Hanna Dettki[3], Lyndon Duong[4], Pierre-Étienne Fiquet[1], Daniel Herrera-Esposito[5], Nikhil Parthasarathy[6], Thomas Yerxa[3], Xinyuan Zhao[3], Eero P. Simoncelli[1,3]; [1]Center for Computational Neuroscience, Flatiron Institute, [2]Indiana University, [3]New York University, [4]Apple, [5]University of Pennsylvania, [6]Google Deep Mind

In sensory perception and neuroscience, new computational models are most often tested and compared in terms of their ability to fit existing data sets. However, experimental data are inherently limited and complex models often saturate their explainable variance, resulting in similar performance across models. Here, we present "Plenoptic", a python software library for synthesizing model-optimized visual stimuli for understanding, testing, and comparing models. Plenoptic provides a unified framework containing three previously-published synthesis methods -- model metamers (Freeman and Simoncelli, 2011), Maximum Differentiation (MAD) competition (Wang and Simoncelli, 2008), and eigen-distortions (Berardino et al. 2017) -- which enable visualization of different aspects of model representations. The resulting images can then be used to experimentally test model alignment with biological visual systems. Plenoptic leverages modern machine-learning methods to enable application of these synthesis methods to any computational model that satisfies a small set of common requirements: the model must be image-computable, implemented in PyTorch, and end-to-end differentiable. The package includes examples of several previously-published low- and mid-level visual models, as well as a set of perceptual quality metrics, and is compatible with the pre-trained machine learning models included in PyTorch's torchvision library. Plenoptic is open source, tested, documented, modular, and extensible, allowing the broader research community to contribute new models, examples, and methods. In summary, Plenoptic leverages machine learning tools to tighten the scientific hypothesis-testing loop, facilitating the development of computational models aligned with biological visual representations.

TALK 3

HOW IS VISUAL PERCEPTION CONSTRUCTED BY VISUALLY RESPONSIVE NEURONS?

Arash Afraz[1]; [1]NIH

Local perturbation of neural activity in high-level visual cortical areas alters visual perception. Quantitative characterization of these perceptual alterations holds the key to understanding the mapping between patterns of neuronal activity and elements of visual perception. Nevertheless, the complexity and subjective nature of these perceptual alterations make them elusive for scientific examination. Here, combining high throughput behavioral optogenetics with cutting edge machine learning tools, we introduce a new experimental approach, "Perceptography", to develop graphical descriptors (pictures) of the subjective experience induced by local cortical stimulation in the inferior temporal cortex of macaque monkeys. According to the "labeled line hypothesis" the causal contribution of inferior temporal neurons to visual perception is expected to be a constant feature determined by the best visual driver of each neuron. However, our results clearly demonstrate that the perceptual events induced by local neural stimulation in inferior temporal cortex highly depend on the contents of concurrent visual perception, refuting the labeled line hypothesis.

TALK 4

SYNTHESIZING STIMULI FOR TARGETED COMPARISON OF BIOLOGICAL AND ARTIFICIAL PERCEPTION

Jenelle Feather[1]; [1]Center for Computational Neuroscience, Flatiron Institute

The past decade has given rise to artificial neural networks that transform sensory inputs into representations useful for complex visual behaviors. These models can improve our understanding of biological sensory systems and may provide a test bed for technology that aids sensory impairments, provided that model representations resemble those in the brain. Here, I will highlight recent lines of work probing aspects of complex model representations using model optimized stimuli, and detail how these stimuli can be used for comparing the representations of the models with human representations. I will first describe work using "model metamers"–stimuli whose activations within a model stage are matched to those of a natural stimulus. Metamers for state-of-the-art supervised and unsupervised neural visual network models were often completely unrecognizable to humans when generated from late model stages, suggesting differences between model and human invariances. While targeted model changes improved human recognizability of model metamers, they did not fully eliminate the human–model discrepancy. Notably, human recognizability of a model's metamers was well predicted by their recognizability by other models, suggesting that models contain idiosyncratic invariances in addition to those required by the task, and that removing these idiosyncrasies may lead to better models of visual perception. To conclude, I will discuss how behavioral results on model metamers, adversarial examples, and synthetic "out-of-distibution" stimuli can show differences between models even when traditional brain-based benchmarks of similarity do not, demonstrating how coupling behavioral measures with targeted stimuli can be an effective tool for comparing biological and artificial representations.

Talk 5

ON THE IMPORTANCE OF DYNAMIC RANGE AND SAMPLE STATISTICS IN FITTING NEURONAL ENCODING MODELS

Binxu Wang[1], Carlos R. Ponce[1]; [1]Harvard University

Understanding visual neurons involves examining how their responses vary with input stimuli. Treating neuronal response prediction as a regression problem, we identify two key data factors for achieving high R^2: sufficient variance in predictors (images) and responses (neuronal firing). Traditionally, we lack control over neuronal firing before image selection. Recently, neuron-guided image synthesis allows us to generate images that control neuronal responses in real time, providing an opportunity to study the effect of response range and image covariance on encoding models. From experiments on hidden units in CNNs and visual cortical neurons, we found that randomly selected natural stimuli often underestimate the full response range, particularly at higher levels visual cortices. Neuron-guided optimization can find stimulus set with higher activation and variance. Encoding models trained on generated images along optimization trajectory can predict graded neuronal responses within the dynamic range for held-out generated images. However, randomly selected natural images, lacking the dynamic range, often have worse R^2 value. We also compared encoding models trained on pre-selected natural images and neuron-guided image samples in different generative spaces (e.g., DeepSim, BigGAN), via feature visualization. Advanced generative models exhibit image statistics similar to those of natural images, helping generalization; however, these priors also bias the encoding model, leading it to infer that neurons respond to more complex features than they do. In contrast, simpler generative models lead to more parsimonious features in encoding model, aiding interpretation. Our study offers valuable insights for the design of image sets for future vision experiments.

Talk 6

A LESS ARTIFICIAL INTELLIGENCE

Andreas Tolias[1]; [1]Stanford University

Neural activity fundamentally shapes our perceptions, behaviors, and cognition, propelling one of neuroscience's greatest quests: decrypting the neural code. This challenge is hindered by our limited ability to precisely record and manipulate extensive neuronal networks under complex conditions and to accurately model the relationships between stimuli, behaviors, and brain states within the natural world's complexity. Recent advancements have started addressing these barriers. Concurrently, advancements in AI now enable analysis of this complex data, facilitating the construction of brain foundation models. These models, akin to AI systems like Video-LLaMA, which decipher video and language relationships, can systematically compile large-scale neural and behavioral data from diverse natural settings. These digital twins of the brain allow for unlimited in silico experiments and the application of AI interpretability tools, enhancing our understanding of neural computations. By applying these insights to AI, we aim to develop more robust, energy-efficient, and comprehensible systems, advancing beyond Big Tech's practice of scaling models with just more behavioral data. Additionally, brain foundation models could revolutionize the diagnosis and treatments for neuropsychiatric disorders. To effectively build these models, we must now decisively move away from traditional hypothesis-driven neuroscience and commit to generating extensive, combined neural and behavioral data across a range of diverse natural tasks.

Talk Sessions

Talk Session: Friday, May 16, 2025, 3:30 – 4:45 pm, Talk Room 1

Eye Movements: Natural tasks, neural mechanisms

Moderator: Austin Roorda, UC, Berkeley

Talk 1, 3:30 pm, 14.11

CONSEQUENCES OF TEMPORAL MODULATIONS ON FOVEAL VISION

Ruitao Lin[1], Alessandro Benedetto[1,2], Michele Rucci[1]; [1]University of Rochester, [2]University of Florence

Previous research has shown that temporally modulating the stimulus improves acuity in the visual periphery. Here we investigate whether temporal modulation can also be used to enhance visual acuity in the foveola. In a forced-choice task, human observers (N=7) judged the orientation of Snellen E optotypes embedded in a 1/f noise background at 0-degree or 7.5-degree eccentricities. The stimulus luminance either remained constant or was temporally modulated in a square-wave manner at 3, 6, or 9 Hz. As expected, visual acuity improved in the periphery when the stimulus was modulated at low temporal frequencies (3 Hz), yielding significantly higher performance than during exposure to a non-modulated stimulus. In contrast, no improvement was observed in the foveola irrespective of the modulation frequency. A possible explanation of this result is that the foveola is particularly sensitive to the luminance modulations introduced by ocular drifts, the persistent fixational eye movements that continually modulate visual input signals. To test this hypothesis, we repeated the experiment with stimuli in the foveola using a custom apparatus to counteract the consequences of eye movements and maintain the stimulus immobile on the retina. We compared performance between optotypes of fixed luminance and temporally modulated at 5Hz. Notably, temporal modulation of stimulus luminance improved acuity under retinal stabilization, even though no improvement was visible during normal viewing. Our results indicate that normal fixational eye movements generate spatiotemporal signals that in the foveola are sufficient for discriminating high acuity stimuli. These results suggest ways for enhancing vision in observers with abnormal fixational motion.

This work was supported by National Institutes of Health grants EY018363, P30 EY001319, and University of Florence (Progetti competitivi 2025-2026).

TALK 2, 3:45 PM, 14.12

DETERMINING WAVELENGTH-IN-FOCUS FOR POLYCHROMATIC VISUAL STIMULI

Benjamin M Chin[1], Martin S Banks[1], Derek Nankivil[2], Austin Roorda[1], Emily A Cooper[1]; [1]University of California, Berkeley, [2]Johnson & Johnson Vision Care

Visual stimuli encountered in the natural environment are typically polychromatic, comprising a combination of visible wavelengths. But the human eye can only bring one wavelength into focus at a time, due to chromatic aberration in its optics. Models of human vision often assume that the eye focuses on light near the peak of the luminosity function, ~555 nm. But in reality, the wavelength-in-focus likely varies depending on the stimulus (Finch et al., 2024). The goal of the present study was to identify the wavelength that the human eye brings into focus for stimuli that vary in color. First, we measured the magnitude of each participants' (n=9) longitudinal chromatic aberration (LCA). This was accomplished via a psychophysical task in which participants indicated the orientation of a Gabor flashed briefly at one of nine virtual distances on an OLED display. The distance at which peak performance occurred for different colors was used to constrain a model of LCA by Thibos et al. (1992). Next, we measured the point spread functions (PSFs) of the participants with a Shack-Hartmann wavefront sensor recording at 30Hz as they focused on three-letter words for three seconds. Stimuli varied in their relative proportions of long, middle, and short wavelengths. Combined with the individual LCA curves, we could then determine the wavelength that was in best focus for each stimulus. We defined best focus as the absence of defocus aberration in the wavefront for a point source on the stimulus. Across stimuli of different colors, the wavelength-in-focus varied notably: it shifted towards longer wavelengths when the stimulus had more long wavelengths, and vice versa for shorter wavelengths. Preliminary modeling suggests this behavior could be driven by a cone-opponent mechanism with negative weights on S- or M-cones and positive weights on L-cones.

This work was supported by the NSF (#2041726) and the NIH (T35EY007139, K99EY036497).

TALK 3, 4:00 PM, 14.13

THE OPTOKINETIC RESPONSE IN FRUIT FLIES IS TUNED TO TYPICAL VISUAL CONSEQUENCES OF SELF-MOTION

Lisa M. Kroell[1] (lisa.m.kroell@gmail.com), Etienne Serbe-Kamp[1], Lisa M. Fenk[1], Martin Rolfs[2]; [1]Max Planck Institute for Biological Intelligence, [2]Humboldt-Universität zu Berlin

Recent evidence shows that Drosophila move the retinas below their rigid compound eyes to smoothly track and thereby stabilize visual image shifts (Fenk et al., 2022). Here, we suggest that this optokinetic response is tuned to the idiosyncratic visual consequences of self-motion. Male flies were reared in darkness and, at 8–10 days of age, exposed to 40 minutes of visual stimulation in a closed- or open-loop environment. In both conditions, head-fixed flies walked on a floating ball while viewing a vertical square-wave grating. In the closed-loop setting, a rotation of the ball along the yaw-axis produced a horizontal translation of the grating on screen (at gains of 0.8 or 6), simulating retinal image shifts during natural locomotion. In the open-loop condition, we presented a replay of stimulus translations produced by flies in the closed-loop setting. We thus obtained pairs of flies that had been exposed to identical temporal frequency (TF) information throughout their lifespan, yet only half of them had generated the underlying image shifts through active self-motion. To measure the temporal tuning of the optokinetic response, we moved a vertical grating horizontally across the screen at 0.4–10 Hz. We simultaneously recorded the position of the deep pseudopupil, a virtual image on the fly retina, with video-based infrared tracking. Across both gain conditions, flies that had actively produced a certain TF range during the exposure phase now followed these frequencies more readily with their retinas than their open-loop counterparts. Moreover, while dark-reared flies initially executed slower retinal movements than light-reared conspecifics, the closed-loop, high-gain exposure condition raised retinal movement velocities significantly above the light-reared baseline. Our findings suggest a surprising plasticity of the fruit fly optokinetic response: To stabilize image shifts during active locomotion with maximum efficacy, retinal movements preferentially follow visual transients that match the perceptual consequences of self-motion.

This research was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [865715 – VIS-A-VIS] to MR).

TALK 4, 4:15 PM, 14.14

TEMPORAL DYNAMICS OF OCULOMOTOR AND PERCEPTUAL ADAPTATION IN RESPONSE TO VISUAL-VESTIBULAR CONFLICT

Phillip Guan[1] (philguan@meta.com), Zhetuo Zhao[1,3], Xiuyun Wu[2], T Scott Murdison[2]; [1]Reality Labs Research, Meta, [2]Reality Labs, Meta, [3]University of Rochester

The consistency between vestibular signals and retinal image motion during head movement is crucial for both the vestibulo-ocular reflex (VOR) and the perception of visual stability. Near-eye optics in prescription eyewear and head-mounted displays (HMDs) can introduce optical distortions to visual input while vestibular signals remain unaffected, leading to visual-vestibular conflict (VVC). This conflict may negatively affect the efficacy of the default VOR response, potentially leading to perceptual errors in visual stability and triggering adaptations in VOR response and visual perception. These two adaptation processes have mostly been studied separately, and the relationship between them—whether one determines the other or they operate as distinct mechanisms—remains unclear. In this work, we characterize perceptual and VOR gain adaptations to five patterns of VVC and disentangle their respective contributions to motor and perceptual changes. Our study is conducted using a custom-built system to facilitate repeatable VOR head motions, accurate head and gaze tracking, and distortion-free, wide field-of-view stimulus presentation. Our results suggest that the perceptual changes in visual stability results from a combination of both motor and perceptual adaptation. We observe both VOR adaptation to minimize demands for smooth pursuit and visual-vestibular recalibration driven by the discrepancies between empirical and expected retinal image motion. Furthermore, we map the evolution of these changes over time (at one minute intervals over nine minutes of adaptation), and we find that VOR adaptation is most pronounced when motor adaptation demands align with retinal motion errors, which also leads to greater shifts in visual stability judgement.

Talk 5, 4:30 pm, 14.15

DIFFERENTIAL EFFECTS OF PERIPHERAL AND CENTRAL VISION LOSS ON SCENE PERCEPTION AND EYE MOVEMENT PATTERNS

Byron A. Johnson[1] (byron_johnson@ucsb.edu), Michael Beyeler[1], Miguel P. Eckstein[1]; [1]University of California, Santa Barbara

Peripheral (PVL) and central vision loss (CVL) are irreversible visual impairments that significantly affect visual tasks like search and reading. Studies have shown that patients with vision loss exhibit reduced visual search accuracies and reading speeds. While previous studies have documented reduced performance in these domains, less is known about how PVL and CVL impact the perception of natural scenes and social cues. To investigate, we tested 32 sighted observers using a gaze-contingent simulation for PVL, CVL, or no impairment. Participants viewed 120 natural scenes (half depicting social interactions and half from the MS COCO dataset) and generated descriptions after one or three saccades. PVL was simulated with a 10-degree clear window surrounded by a Gaussian blur, while CVL applied a 10-degree Gaussian blur to the center of fixation. Eye movement data was analyzed by determining the correlation between fixation heat maps for each viewing condition and scene. Description quality for each scene was rated for semantic similarity to gold-standard descriptions. Results revealed a significant three-way interaction between viewing condition, scene type, and saccade count ($F=3.978$, $p=.018$). Interestingly, when viewing social interaction scenes with one saccade, descriptions generated with PVL were rated lower than CVL ($p=.0001$, Cohen's $D = .3045$) and no impairment ($p < .0001$, Cohen's $D = .5206$). Fixation heat map correlations between CVL and no impairment were lowest across scene types and saccades ($F=52.907$, $p < .0001$), suggesting greater changes in fixation patterns compared to no impairment ($p < .0001$, Cohen's $D = 1.29$) and PVL ($p < .0001$, Cohen's $D = 1.02$). These findings suggest distinct effects underlying scene perception in PVL and CVL: PVL reduces semantic understanding of scenes, while CVL alters gaze behavior. This work underscores the need for tailored interventions based on impairment type to improve daily functioning for individuals with vision loss.

Talk Session: Friday, May 16, 2025, 3:30 – 4:45 pm, Talk Room 2

Perceptual Organization: Neural mechanisms, models

Moderator: Yuanning Li, ShanghaiTech University

Talk 1, 3:30 pm, 14.21

INVESTIGATING SEMANTIC EXPECTATION AND PREDICTIVE ERROR IN THE VISUAL CORTEX WITH A LARGE FMRI VISION-LANGUAGE DATASET

Shurui Li[1] (lishr2022@shanghaitech.edu.cn), Zheyu Jin[1], Ru-Yuan Zhang[2], Shi Gu[3], Yuanning Li[1]; [1]ShanghaiTech University, Shanghai, China, [2]Shanghai Jiao Tong University, Shanghai, China, [3]University of Electronic Science and Technology of China, Chengdu, China

Classical models of visual processing in the brain emphasize a predominantly feedforward hierarchical coding scheme, where lower-level features are progressively integrated into higher-level semantic representations. However, this view fails to fully account for the complex and dynamic nature of semantic information processing in the visual cortex, which involves interactions that extend beyond passive feedforward pathways. To study the neural coding of semantic information in the visual hierarchy, we collected a large-scale fMRI vision-language dataset, where each participant processed over 4,400 paired stimuli consisting of a text caption followed by a naturalistic image, with the task of evaluating their semantic congruence. Driven by the predictive coding theory, we hypothesized that the early visual cortex can represent semantic expectations and predictive errors. First, we observed that when subjects were presented with images that matched their semantic expectations, the early visual cortex responded significantly less than they were presented with unexpected images. To explain the neural coding, we built neural encoding models using features extracted from large language and vision models. We found that language model features could predict the early visual cortex's response after the subjects viewed a text caption. This indicates that the early visual cortex can generate semantic expectations. Next, we found that neural activity from V1 to V3 encode prediction mismatch of visual stimuli, representing a progression from low-level to high-level prediction errors. Finally, we found that the degree of response amplitude reduction correlated to the neural coding of high-level prediction errors. To sum up, using a large fMRI vision-language dataset, we provide evidence of cross-modal semantic expectation and predictive error coding in the visual cortex.

This work is supported by the National Natural Science Foundation of China (32371154, Y.L.), Shanghai Rising-Star Program (24QA2705500, Y.L.), Shanghai Pujiang Program (22PJ1410500, Y.L.), and the Lingang Laboratory (LG-GG-202402-06, Y.L.).

Talk 2, 3:45 pm, 14.22

MESOSCALE FUNCTIONAL CONNECTIVITY IN HUMAN V1 REVEALED BY HIGH-RESOLUTION FMRI.

Marianna E. Schmidt[1,2], Iman Aganj[3,4], Jason Stockmann[3,4], Berkin Bilgic[3,4,5], Yulin Chang[6], W. Scott Hoge[7], Evgeniya Kirilina[1], Nikolaus Weiskopf[1,8,9], Shahin Nasr[3,4]; [1]Max Planck Institute for Human Cognitive and Brain Sciences, [2]Max Planck School of Cognition, Leipzig, Germany, [3]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, [4]Harvard Medical School, [5]Harvard/MIT Health Sciences and Technology, [6]Siemens Medical Solutions USA Inc., Malvern, PA, USA, [7]Imaginostics, Inc., Orlando, FL, USA, [8]Felix Bloch Institute for Solid State Physics, Faculty of Physics and Earth System Sciences, Leipzig University, Leipzig, Germany, [9]Wellcome Center for Human Neuroimaging, Institute of Neurology, University College London

Despite their importance in shaping visual perception, functional connectivity between ocular dominance columns (ODCs), the building blocks of neuronal processing within the human primary visual cortex (V1), remains mostly unknown. In this study, we used high-resolution fMRI (7T; voxel size = 1mm isotropic) to localize ODCs and assess their resting-state (eyes closed) functional connectivity (rs-FC) in 11 individuals (3 females) with intact vision (age=30.9±5.9 years). Consistent with studies in animals, we found stronger rs-FC between ODCs with alike rather than unalike ocular preference ($p<0.01$). The level of rs-FC was generally higher at mid-cortical depths, while selectivity was more pronounced at superficial and deep cortical depths. Surpassing expectations from anatomical studies of ODC connectivity, we found the following. First, the selective rs-FC between ODCs was preserved for distances of up to 4cm, indicating that connectivity between ODCs remains selective across multiple synapses. Second, rs-FC selectivity was significantly higher between ODCs that exhibited stronger (compared to weaker) ocular preference (p<10-3), even though ODC mapping and rs-FC measurements were conducted in separate scan sessions. Third, the extent of selectivity appeared to vary between foveal vs. peripheral and to a lesser extent between dorsal vs. ventral regions, suggesting a heterogeneity in the distribution of rs-FC within V1. We further tested whether the ODC map was predictable from the rs-FC pattern. Our preliminary results showed a significant correlation between rs-FC and ODC maps (p<10-5). The level of this correlation declined when the size of regions of interest increased from 20.21mm2 ($r=0.20$) to 286.47mm2 ($r=0.08$). This result indicates a promising opportunity for ODC mapping in individuals with disrupted binocular vision (e.g., monocular blindness). In conclusion, our results demonstrate the utility of high-resolution fMRI for studying mesoscale rs-FC within V1, successfully replicating the findings based on animal models and highlighting promising new opportunities for future exploration.

Talk 3, 4:00 pm, 14.23

FEATURE-SPECIFIC AND FEATURE-INDEPENDENT ENSEMBLE REPRESENTATIONS IN THE HUMAN BRAIN

Patxi Elosegi[1,2], Yaoda Xu[2], David Soto[1]; [1]Basque Center on Cognition, Brain and Language, Donostia, Spain, [2]Psychology Department, Yale University, New Haven, CT, US

The human brain overcomes processing limitations by compressing redundant visual input into ensemble representations. While psychophysical studies demonstrate that ensembles are efficiently extracted across low-, mid-, and high-level features, the domain-generality of ensemble perception remains unclear. Neuroimaging holds the key to addressing this, yet prior findings have been inconsistent. This fMRI study aims to (i) test whether ensembles composed of visual features of increasing complexity—orientation, shape, and animacy—are processed locally in feature-selective areas or by a shared neural substrate, and (ii) assess whether different summary statistical descriptors such as mean and ratio involve common brain representations. We collected fMRI data from 24 participants (two scanning sessions) using a mini-block design. In each mini-block, participants saw a sequential presentation of five ensemble displays, consisting of the same twelve objects that shuffled positions across successive displays. Stimuli were carefully generated to vary both in the ratio of items from each class (e.g., living vs. nonliving items) and along a continuous distribution of the mean ensemble feature (e.g., average lifelikeness). Results from a searchlight MPVAs revealed a clear dissociation: mean ensemble features are encoded locally in a feature-specific manner along the medial Ventral-Visual-Pathway, following the texture-sensitive collateral sulcus. In contrast, ratio information—defined by the proportion of items from each class—is encoded in a feature-independent manner in the Dorsal-Visual-Pathway, particularly along the intraparietal sulcus, as demonstrated by cross-decoding analyses. Thus, ensemble representations are neither completely distributed nor centralized but involve the interplay between sensory and PPC areas to encode both stimulus-specific and stimulus-independent information. To test the generality of these results, we will next assess if similar representations emerge on CNN architectures trained to perform the same tasks as participants. This work bridges gaps in prior neuroimaging research and provides an open-source fMRI dataset, fostering computational models of ensemble processing across diverse visual features.

P.E.: Basque Government PREDOC grant; D.S.: Basque Government BERC 2022-2025 program and Spanish Ministry 'Severo Ochoa' Programme for Centres/Units of Excellence in R & D (CEX2020-001010-S) and project grants PID2019-105494GB-I00. Y.X.: is supported by US NIH grant R01EY030854

Talk 4, 4:15 pm, 14.24

LARGER AND EARLIER CATEGORY-SELECTIVE NEURAL ACTIVITY IN THE HUMAN VENTRAL OCCIPITO-TEMPORAL CORTEX THAN IN THE MEDIAL-TEMPORAL-LOBE

Simen Hagen[1], Corentin Jacques[1], Sophie Colnat-Coulbois[1,3], Jacques Jonas[1,2], Bruno Rossion[1,2]; [1]Université de Lorraine, CNRS, F-54000 Nancy, France, [2]Université de Lorraine, CHRU-Nancy, Service de Neurologie, F-54000 Nancy, France, [3]Université de Lorraine, CHRU-Nancy, Service de Neurochirurgie, F-54000 Nancy France

The human ventral-occipito-temporal cortex (VOTC) contains spatially dissociated category-selective neural populations. For example, face-selective (i.e., significantly different responses to faces than non-face objects) neural populations are found in the lateral fusiform gyrus, while place-selective (i.e., significantly different responses to buildings than non-building objects) neural populations are found in and around the more medial parahippocampal gyrus. However, less is known about category-selective populations in the medial temporal lobe (MTL) and their interactions with the VOTC. On the one hand, the socially relevant amygdala (AMG) and the spatially relevant hippocampus (HPC) could relay early activity to face- and place-selective populations in the VOTC, respectively. On the other hand, face- and place-selective neural populations in the VOTC could relay early activity to the AMG and HPC, respectively. Here, we examined the spatio-temporal profiles of face- and place-selective activity in the MTL and the VOTC of a large group of epileptic patients (N=88) implanted with intracerebral electrodes in the grey matter of these regions. Both face- and place(buildings)-selective neural activity was isolated with separate frequency-tagging protocols, providing objective measures of category-selective neural activity, devoid of low-level confounds, with high spatial and temporal resolution. We find both face- and place-selective contacts in the MTL, with larger and earlier face-selective than place-selective activity in AMG. In contrast, places elicit earlier category-selective activity than faces in the HPC. Crucially, category-selective activity is more prominent (~3 times higher) and emerges significantly earlier (~50 ms) in the VOTC than in the MTL. These findings cast doubt on the view that the amygdala can serve as an early "face detector" in the human brain and suggest that face- and place-selective activity follow different circuits from the VOTC to different MTL regions.

Funding declaration: ANR IGBDEV ANR-22-CE28-0028; ERC AdG HUMANFACE 101055175

Talk 5, 4:30 pm, 14.25

THE ORGANIZATION OF HIGH-LEVEL VISUAL CORTEX IS ALIGNED WITH VISUAL RATHER THAN ABSTRACT LINGUISTIC INFORMATION

Adva Shoham[1] (advashoham@mail.tau.ac.il), Rotem Broday-Dvir[2], Rafael Malach[2], Galit Yovel[1]; [1]Tel Aviv University, [2]Weizmann Institute of Science

Recent studies showed that the response of high-level visual cortex to images can be predicted by their linguistic descriptions, suggesting an alignment between visual and linguistic information. We hypothesized that such alignment is limited to textual descriptions of the visual content of the image and does not extend to abstract descriptions. We distinguish between two types of linguistic descriptions of visual images: visual text, which describes the image's purely visual content, and abstract text, which describes conceptual knowledge unrelated to the immediate visual attributes. Accordingly, we tested the hypothesis that visual text, but not abstract text, will predict the neural response to images in high-level visual cortex. To that purpose, we used visual and language deep learning algorithms to predict the iEEG responses in humans to images of familiar faces or places. We generated two types of textual descriptions for the images: visual text, describing the visual content of the image, and abstract text, based on their Wikipedia definitions. We then extracted the relational-structure representations from a large language model (GPT-2) for the text descriptions and from a deep neural network (VGG16) for the images. Using these visual and linguistic representations, we predicted the iEEG responses to the images. Our findings showed that neural responses in high-level visual cortex were similarly predicted by the visual representation of the images and linguistic representations of the visual text, but not by abstract text. Frontal-parietal electrodes showed a reverse pattern. These results are in line with recent findings showing that textual descriptions of the content of the image predict the response to images also in the macaque's brain. These findings demonstrate that visual-language alignment in high-level visual cortex is limited to visually grounded language.

Talk Session: Saturday, May 17, 2025, 8:15 – 9:45 am, Talk Room 1

Attention: Neural, objects, models

Moderator: Adam Greenberg, Medical College of Wisconsin and Marquette University

Talk 1, 8:15 am, 21.11

REVISITING VISUAL AWARENESS: NO EVIDENCE FOR LEVELS OF PROCESSING

Aytac Karabay[1] (ak10268@nyu.edu), Daryl Fougnie[1]; [1]Department of Psychology, New York University Abu Dhabi

There is debate about whether awareness during visual perception occurs abruptly (all-or-none) or gradually. One influential view is the levels of processing (LOP) theory, which states that the nature of visual awareness depends on the stimulus processing level. According to LOP, low-level stimuli (e.g., color) evoke gradual awareness, while high-level stimuli, such as object identity (e.g., letter), elicit abrupt, all-or-none perception. A critical source of evidence supporting LOP is that self-reported perceptual clarity measures reveal more intermediate values of perceptual clarity for low than high-level stimuli. Here, we provide several pieces of evidence inconsistent with this theory. First, previous studies confound stimulus levels with 'category-flatness.' Does increased perceptual clarity of X versus blue reflect that a noisy perception of X is often perceived as X due to large category priors for letter stimuli? Experiment 1 showed that when the perceptual clarity of a high-level stimulus set without meaningful category boundaries (morphed faces) was tested, perceptual clarity was more gradual than that of low-level stimuli. Second, by varying foil-target similarity, we show that the assumption underlying perceptual clarity measures—that they measure perceptual clarity rather than the difficulty of perceptual judgments—is incorrect. Finally, we address a significant issue in the existing literature: performance across different stimuli is often not equated. This lack of equivalence can skew perceptual clarity ratings, as high-level stimuli typically yield better performance, potentially leading to fewer intermediate ratings in categorical stimuli due to high confidence. To rectify this, we equated task performance across stimuli using a staircase method (Experiment 2). We found that the gradualness of perceptual clarity was consistent across all stimulus types, rejecting the notion of distinct awareness pathways for high- versus low-level stimuli. Ultimately, our results suggest that differences between gradual and all-or-none perception arise largely from methodological properties rather than levels of processing.

TALK 2, 8:30 AM, 21.12

PURELY VOLUNTARY SHIFTS OF OBJECT-BASED ATTENTION CAN BE FUNCTIONALLY IDENTIFIED AND CHARACTERIZED VIA FMRI AND MVPA

David H Hughes[1], Adam S Greenberg[1]; [1]Medical College of Wisconsin and Marquette University

We have previously used multivariate pattern analysis (MVPA) and fMRI to compare cued and non-cued (purely voluntary) shifts of spatial attention (Gmeindl et al., 2016). Here, we hypothesized that these methods could be applied to object-based attention (OBA), and that differential activation for cued and non-cued shifts would be observed in attentional control regions. To test this, 17 healthy adults viewed a series of overlapping faces and houses while detecting an infrequently-appearing target face and target house. During fMRI, participants completed six runs, each comprising both a cued and a non-cued block. A thin, colored frame provided shift/hold instructions during cued blocks but was uninformative during non-cued blocks. We used leave-one-run-out cross-validation to train (cued blocks only) and test (cued and non-cued blocks) a support vector machine to determine participants' attentional locus. The output was scaled to a probability (i.e., probability attending house) which allowed us to index shifts at the onset of rapid probability changes. We then compared activations time-locked to these shift indices within published ROI coordinates (Gmeindl et al., 2016). "False" shifts were identified during cued blocks that occurred in the absence of a cue. Comparing cued, false, and purely voluntary (i.e., during non-cued blocks) shifts, we observed differences in right supramarginal gyrus (rSMG) and left precuneus (lPreC). In rSMG, activation was significantly reduced for cued shifts compared to false and non-cued shifts from -3 s to 0 s (p's < .027). In lPreC, activation for false shifts was significantly elevated compared to cued and non-cued shifts from -4.5 s to -1.5 s (p's < .047). Thus, rSMG reflects successful interpretation of an external shift cue while lPreC reflects top-down reorienting of attention. Our results demonstrate that purely voluntary shifts of OBA can be identified and tracked within parietal cortex in the absence of external shift cues.

This work was supported by grants BCS-2122866 from NSF and T32EY014536 from NIH NEI (the contents of this project are solely the responsibility of the authors and do not necessarily represent the official views of the NEI or NIH).

TALK 3, 8:45 AM, 21.13

DISTRACTOR INTRUSIONS: A BRIEF AND HIGHLY RELIABLE MEASURE OF INDIVIDUAL DIFFERENCES IN THE SPEED OF ATTENTION

Alon Zivony[1] (a.zivony@sheffield.ac.uk), Claudia von Bastien[1], Rachel Pye[2]; [1]University of Sheffield, [2]University of Reading

How quickly we attend to objects plays an important role in navigating the world, especially in dynamic and rapidly changing environments (e.g., a busy street). Reaction times (RTs) in visual search tasks have often been used as an intuitive proxy of this ability. However, such measures are limited by inconsistent levels of reliability and the multitude of non-attentional factors that affect an individual's RT. Here, we present an alternative method of studying individual differences in the speed of attention. Specifically, we employ rapid serial visual presentation (RSVP) tasks, where a target is presented for brief durations and embedded among multiple distractors. Previous research showed that the distractor intrusions, that is, reports of an adjacent distractor instead of the target, is associated with the speed of attention. Here, we explored the validity and reliability of individual differences in people's rate of distractor intrusions. In three experiments, we found that intrusion rates predict overall RTs in simple visual search tasks, but emerge independently from measures of attentional control, reading speed, and another well-known limitation in temporal attention (the attentional blink). Moreover, our findings (N=100) show that an individual's intrusion rate can be measured with very high reliability (>.90) within a very short (5-minute) session, both within a single session and between two sessions a week apart. These findings show that the distractor intrusion paradigm is a useful tool for research into individual differences in the temporal dynamics of attention. Links to a downloadable an easily executable distractor intrusion experiment are provided to facilitate such future research.

TALK 4, 9:00 AM, 21.14

EYE TRACKING REVEALS THE EFFICACY OF OBJECT-BASED ATTENTION AT FILTERING OUT DISPROPORTIONATELY SALIENT FOVEAL DISTRACTORS

Lasyapriya Pidaparthi[1] (lasyapriya.pidaparthi@vanderbilt.edu), Frank Tong[1]; [1]Vanderbilt University

Visual attention helps people prioritize task-relevant information. Covert spatial attention subtly increases the perceived contrast of an attended stimulus by ~4%, an effect also observed before goal-directed eye movements via pre-saccadic attention. Attention can thus modestly boost signal strength and improve perceptual performance (Li, Hanning, & Carrasco, 2021). However, it remains unclear how attention interacts with contrast when distracting information dynamically overlaps with an attended target. We have previously shown that object-based attention (OBA), as assessed by eye movements, effectively filters out the presence of an overlapping distractor object (Pidaparthi & Tong, VSS 2024). How then might OBA be impacted as a function of relative target-distractor contrast? We examined this question across two experiments. In Experiment 1, participants attended to one of two naturalistic objects (face, flower) that followed pseudorandom, minimally correlated trajectories while remaining largely overlapping, while monitoring for brief spatial distortions of the attended object. We tested five target-distractor contrast levels: 50% vs. 50%, 33:67, 25:75, 17:83, and 10:90. To measure the efficacy of OBA, we used a sliding-window correlation analysis and evaluated gaze-following of the attended object. We observed nearly complete filtering of the distractor up to extreme target-distractor contrast ratios of 17:83, beyond which filtering efficacy dropped (mean r: 0.571, 0.564, 0.545, 0.490, 0.288). Detection task performance, in comparison, deteriorated more rapidly, so attention-based gaze-following was more robust. In Experiment 2, we substituted the irrelevant object with a Gabor stimulus that underwent brief bursts of drifting motion (4 Hz for 500ms). Attentional filtering was unperturbed by the extraneous motion of the Gabor up to contrast ratios of 10:90. Overall, we show via eye movements that OBA effectively filters out complex motion signals from a distractor across a range of contrast levels, weakening only when the distractor salience exceeds the target salience by a factor of 5-fold.

This research was supported by NEI grants R01EY035157 to FT and P30EY008126 to the Vanderbilt Vision Research Center.

Talk 5, 9:15 AM, 21.15

STIMULUS-INDEPENDENT MODULATION OF CEREBELLAR NEURAL ACTIVITY BY VISUAL SPATIAL ATTENTION

Brenda (Siyue) Qiu[1] (siyueq@uw.edu), Kristin M. Woodard[1], Bridget Leonard[1], Scott O. Murray[1]; [1]University of Washington

The cerebellum has traditionally been associated with motor control, but recent studies also suggest its involvement in cognitive functions, including attention (Brissenden et al., 2016; King et al., 2019). This study investigates how visual spatial attention influences cerebellar neural activity. Functional MRI data were collected from 50 participants completing a motion direction discrimination task. Each trial began with a cue directing participants to attend to the left or right visual hemifield, followed by the bilateral presentation of two Gabor gratings at one of three contrast levels (0%, 6%, or 50%). Participants then reported the motion direction of the attended grating. Cerebellar voxels of interest (VOIs) were identified based on significant differences between peak and baseline activation in the 0% contrast condition, where no grating was shown after the cue, isolating the stimulus-independent attention effects. The identified VOIs revealed an ipsilateral representation of visual spatial attention in the cerebellum, consistent with previous findings (Brissenden et al., 2018). Comparisons of percent fMRI signal change between attended and unattended conditions indicated attentional modulation in lobules VIIb, VIIIa, and VIIIb, as well as in lobules V, VI, Crus I, and Crus II. These findings align with prior research on the cerebellum's role in visual perception and attention (Brissenden et al., 2018; van Es et al., 2019; Brissenden et al., 2021). In cortical visual areas V1 and MT+, we found an additive effect of attention, where the increased neural activity was observed across different contrast levels, similar to previous findings (Murray, 2008). In contrast, cerebellar neural activity was predominantly modulated in the 0% contrast condition. Qualitatively, the gap between attended and unattended contrast response functions diminished as contrast increased. These findings differ from effects typically observed in cortical visual areas and suggest a unique, stimulus-independent role for the cerebellum in spatial attention.

This work was supported by funding from the National Institutes of Health (R01MH131595 to S.O.M.)

Talk 6, 9:30 AM, 21.16

VISUAL AWARENESS POSITIVITY: A NOVEL NEURAL CORRELATE OF CONSCIOUSNESS

Ugo Bruzadin Nunes[1] (bruzadinnunes@chapman.edu), Angelica Nicolacoudis[2], Adi Sarig[3], Nicholas Fish[1], Liad Mudrik[2], Michael Pitts[2], Aaron Schurger[1]; [1]Chapman University, [2]Reed College, [3]Tel Aviv University

Inattentional blindness (IB), the failure to perceive salient stimuli when attention is directed elsewhere, challenges assumptions about visual awareness. Despite extensive research on neural correlates of consciousness (NCCs), mechanisms underlying IB for complex stimuli remain poorly understood. Here, we leveraged a three-phase no-report paradigm to investigate neural signatures of conscious awareness in IB (using EEG), minimizing confounds like motor and decision-making processes. Participants performed a challenging peripheral attention task while simultaneously being presented centrally with faces, houses, or visual noise. At each phase's end, they were probed about various presented objects, including faces and houses. Approximately 45% of participants exhibited IB in phase 1. In phase 2, participants were informed about the presence of faces and houses but continued the same task. In phase 3, they were instructed to ignore the peripheral task and instead identify the central stimuli (faces, houses, noise) in a three-alternative forced-choice (3AFC) task. Non-parametric cluster analysis of event-related potentials (ERPs) contrasting phase 2 with phase 1, controlling for noise trials, identified two distinct neural components: Visual Awareness Negativity (VAN, 180–220 ms) and a novel Visual Awareness Positivity (VAP, 250–400 ms) characterized by bilateral-posterior positive and frontal-central negative differences. The P3b/P300, a traditional NCC marker, was absent during phases 1 and 2 but present during phase 3. Multivariate pattern analysis (MVPA) assessing temporal generalization of decoders showed stable above-chance decoding of seen vs. unseen trials, during the 250–400 ms post-stimulus window. Time-frequency cluster analysis revealed significant differences in theta, alpha, and beta ranges, implicating these rhythms in visual consciousness. These findings replicate previous results regarding the VAN and identify novel markers of conscious perception, including the VAP, meta-stable decoding, and differential theta power, in no-report conditions. This study deepens our understanding of visual awareness and highlights novel neural markers as potential NCCs in no-report paradigms.

Grant 30266 from the Templeton World Charity Foundation

Talk Session: Saturday, May 17, 2025, 8:15 – 9:45 AM, Talk Room 2

3D Processing

Moderator: Roland Fleming, Justus Liebig University Giessen

Talk 1, 8:15 AM, 21.21

TRANSFORMER MODELS BETTER ACCOUNT FOR HUMAN 3D SHAPE RECOGNITION THAN CONVOLUTION-BASED MODELS

Shuhao Fu[1] (fushuhao@g.ucla.edu), Philip Kellman[1], Hongjing Lu[1]; [1]University of California, Los Angeles

Both humans and deep learning models can recognize objects from 3D shapes depicted with sparse visual information, such as a set of points randomly sampled from the surfaces of 3D objects (termed a point cloud). Although deep learning models achieve human-like performance in recognizing objects from point-cloud displays, it is unclear whether these models acquire 3D shape representations aligned with those used by humans. We conducted three human experiments by varying point density and object orientation (Experiment 1), perturbing the local geometric structure of 3D shapes (Experiment 2), and manipulating the global configuration by scrambling parts (Experiment 3). Results showed that humans rely on global 3D shape representations for object recognition, showing robust performance even when local features were disrupted. We tested two deep learning models with different architectures: a convolution-based model (DGCNN) and a vision transformer model (Point transformer), comparing their performance with humans across the three experiments. The convolution-based DGCNN model relied heavily on local geometric features and was highly susceptible to adversarial perturbations of local geometry. In contrast, the transformer-based model exhibited human-like reliance on global shape representations. Ablation simulations revealed that the global shape representations in the transformer model originated primarily from its downsampling mechanism, which explicitly operationalizes a fine-to-coarse and local-to-global process, analogous to the increasing receptive field sizes of neurons in the human visual system. Secondary contributions stem from position encoding, which maintains spatial information across layers, and attention mechanisms which adaptively weigh context information based on similarity. These findings highlight the key computational mechanisms to bridge the gap between human and machine representations in 3D object recognition.

NSF BCS 2142269

Talk 2, 8:30 am, 21.22

COMPUTATIONAL MODELS EXHIBIT INVARIANCE AND MULTISTABILITY IN SHAPE FROM SHADING

Xinran Han[1], Ko Nishino[2], Todd Zickler[1]; [1]Harvard University, [2]Kyoto University

Shape perception from a single shaded image is both unique and inherently ambiguous. It is unique in that a single geometric representation – the curvature field – explains the underlying shape. At the level of surface orientation, however, it is not unique – multiple valid shape solutions exist. Humans experience this ambiguity as multistable perception, where shape from a single image can be alternately interpreted as either convex or concave. We develop two computational methods to model these phenomena, each focusing on a different output representation. Our first method derives a unique representation underlying the observed shading—specifically, the log-Casorati curvature field—and demonstrates that it can be inferred from shading and remains robust under variations in lighting and surface albedo. This invariant representation provides a unified geometric representation of the physically ambiguous observation, laying the foundation for shape perception under varying conditions. Our second method employs a diffusion model—a generative neural network to recover multiple explanations of the surface shape by aggregating putative interpretations from local patches. This model offers a bottom-up mechanism for generating diverse surface orientation proposals, which can be later integrated with top-down queries for further refinement, and may be closer to biological realizations. Both methods are designed to operate without explicit knowledge of lighting direction, inspired by hypotheses suggesting that shape perception often precedes lighting inference. Their reliance on a shift-invariant, bottom-up architecture allows for efficient training on small synthetic datasets while generalizing well to novel scenarios. Our experiments show that these models can mimic ambiguities in shape from shading, including multistable phenomena like the crater illusion. Moreover, our studies highlight shortcomings in prior computational models, which typically produce only a single "best" interpretation. Together, these models provide insight into human perception by demonstrating how invariant and multistable shape representations can emerge from ambiguous inputs.

Talk 3, 8:45 am, 21.23

PERCEIVED 3D SHAPE OF MIRROR-LIKE OBJECTS: INTERACTIONS OF MONOCULAR AND BINOCULAR CUES

Celine Aubuchon[1] (celine.d.aubuchon@gmail.com), Emily A-Izzeddin[1], Fulvio Domini[1,2], Alexander Muryy, Roland W. Fleming[1,3]; [1]Justus Liebig University Giessen, Germany, [2]Brown University, [3]Center for Mind, Brain and Behavior, Universities of Marburg and Giessen, Germany

When viewed binocularly, purely specular (mirror-like) objects are bizarre and quite fascinating. They present a naturally-occurring case of major cue conflict, with monocular and binocular shape cues decoupled almost everywhere in the image. This occurs because reflections are viewpoint dependent. With viewpoint changes (e.g., from left- to right-eye view), specular reflections can slide across the surface, change shape, or disappear altogether. As a result, while 'normal' features—like texture markings, 3D corners or diffuse shading—obey epipolar geometry, and create accurate depth signals, for specular reflections, the 2D vector fields describing the interocular shifts of corresponding features can have arbitrary sizes and directions. This leads to both unfusible regions and spurious depth signals that are incompatible with monocular shape cues. We created 3D "blob" objects, reconstructed their interocular vector fields, and used them to investigate how the brain deals with these peculiar conflicts. Specifically, we had participants make judgments about the 3D depths and orientations of points on objects with different combinations of cues derived from mirror reflections. We teased apart monocular and binocular cues by shifting image content along the vector field to recreate the binocular characteristics of mirrors while independently varying the information provided by monocular cues. In a control condition, we also 'painted' the reflections onto the surface in depth to remove the naturally-occurring cue conflict. Our results show that observers' perception of 3D shape for mirrored objects, when monocular cues are available, correlate with true object shape despite significant biases in perceived shape when mirror disparities are presented alone. We find (1) that the visual system discriminates between "good" and "bad" disparities in determining object shape and (2) that monocular cues are used to constrain perceived shape from specular disparity-fields.

Supported by DFG, project number 222641018–SFB/TRR 135 TP C1, the ERC Advanced Grant "STUFF" (project number ERC-2022-

AdG-101098225); and the Research Cluster "The Adaptive Mind" funded by the Hessian Ministry for Higher Education, Research, Science and the Arts.

TALK 4, 9:00 AM, 21.24

3D SLANT DISCRIMINATION IN MOTION PARALLAX DEPENDS ON RETINAL GRADIENTS AND NOT DISTAL SLANT

Jovan Kemp[1], James Todd[2], Fulvio Domini[1]; [1]Brown University, [2]The Ohio State University

The problem of slant-from-motion in human perception is often thought to be solved by inverting the retinal velocities using estimates of ancillary properties such as the speed and direction of the object in motion. However, previous evidence suggests that shape estimates emerge directly from retinal velocity gradients through a simple heuristic, rather than solving for parameters to perform inverse geometry. Consequently, slant metamers, which denote when two distinct surfaces are perceived to have the same slant, can be produced by ensuring that the retinal projections of the two surfaces produce the same retinal velocity gradients. Here we tested this hypothesis by requiring participants to perform (1) a probe adjustment task, which allowed us to measure absolute perceived slant, and (2) a 2-interval forced choice (2-IFC) task, which allowed us to measure discrimination rate performance when viewing rotating planar surfaces. In the probe adjustment task, participants adjusted a 2-dimensional probe to report their perceived slant relative to surfaces that could have one of three slants and rotate to produce one of three retinal velocity gradients. We find that the perceived slant reflects changes in retinal velocity gradients, rather than the surface slant. In the 2-IFC task, participants judged which two surfaces presented at separate intervals were the most slanted. The fixed standard surface could have one of two slants and rotate to produce one of two retinal velocity gradients. Critically, the variable comparison was chosen to produce a specified retinal velocity gradient either by changing the slant or changing the rotational speed. We find that discrimination thresholds, similarly to perceived slant, are determined by the retinal gradients, rather than by the slant or rotational velocity alone. Taken together, we show compelling evidence that perceived slant-from-motion is largely determined by simple heuristics using retinal information, rather than by more complex computations.

This research was funded by NSF #2120610 to FD.

TALK 5, 9:15 AM, 21.25

UNCERTAINTY IN PERCEPTUAL PRIORS OF SPORTS BALLS' SIZES SCALE INDEPENDENTLY OF EXPERIENCE

Constantin Rothkopf[1], Nils Neupaertl[1]; [1]Technical University of Darmstadt

Human perception resolves ambiguities and uncertainties by combining sensory measurements with prior knowledge. Bayesian models of perception have posited that these priors are well-calibrated to the natural environment through experience, so their accuracy and precision should increase with experience. Here, we measured participants' prior beliefs about the size of three sports balls with which participants had different degrees of prior experience: tennis ball, baseball, and soccer ball. None of our European participants had ever interacted with a baseball. 16 participants viewed pairs of sports balls through an Oculus VR headset in a carefully calibrated scene of the actual laboratory. The experimental procedure used the so-called Markov Chain Monte Carlo with humans paradigm to elicit participants' priors over the balls' sizes. Participants repeatedly selected which of the two balls displayed on a trial looked more realistic in a two-alternative forced-choice decision. The simultaneously shown balls of the same type only differed in size. Balls' sizes in subsequent trials were then sampled probabilistically relative to the preceding choice. This procedure has been shown to elicit and accurately measure individuals' perceptual priors. Participants' priors showed different degrees of bias and uncertainty. However, the individual differences in biases across balls were highly correlated within participants. Moreover, the range of biases across subjects scaled linearly with their mean. Surprisingly, the uncertainty in the prior belief, as quantified by its standard deviation, was not the largest for the balls with which participants had the least experience. Instead, the standard deviation scaled linearly with the mean of the size's prior belief. These results show that perceptual priors for sports ball sizes are not determined only by prior experience. Instead, these results suggest that people have structured size priors across different objects and that the priors' uncertainty is lawfully related to the prior mean.

TALK 6, 9:30 AM, 21.26

NAVIGATING IN A REALISTIC VR ENVIRONMENT WITH CENTRAL FIELD LOSS

Jade Guénot[1], Preeti Verghese[1]; [1]Smith-Kettlewell Eye Research Institute

Impaired depth perception due to central field loss (CFL) significantly impacts mobility, obstacle avoidance, and increases fall risks. We explored obstacle avoidance strategies in CFL patients and control participants under various visual conditions, using a virtual reality environment (HTC Vive Pro Eye headset in conjunction with the PTVR Python toolbox). Participants navigated through a realistic living room to retrieve keys while avoiding or stepping over boxes of varying heights (5-40 cm). Two conditions were tested: an easy condition with six easily avoidable boxes (>15 cm high) and a challenging condition with twelve boxes of mixed heights, including small boxes that were easier to step over than to go around. Participants performed the task binocularly and monocularly. Control participants completed additional trials with and without an artificial gaze-contingent scotoma. We recorded head, hand, and feet positions/rotations, measuring velocity, path tortuosity, foot clearance, and collision frequency, along with gaze behavior through integrated eye tracking. Preliminary results revealed distinct navigation strategies between patients with CFL and controls. CFL patients exhibited significantly slower walking velocities (especially in the monocular condition) and higher path tortuosity ratios, indicating a preference for obstacle avoidance over stepping over obstacles. When stepping over boxes, they showed reduced foot-to-obstacle clearance, resulting in higher collision rates and suggesting difficulties in height estimation. These impairments were even more pronounced than for controls with artificial scotomas. No difference was found between monocular and binocular conditions in patients except for the reduced velocity. Our findings suggest that VR environments can effectively reveal navigation strategies and

challenges for patients with CFL, contributing to our understanding of how CFL affects their behavior in the real world, and offering insights into compensatory strategies and risks associated with vision loss. It may help the development of rehabilitation strategies and assistive technologies for individuals with CFL.

NIH funding R01 EY27390

Talk Session: Saturday, May 17, 2025, 10:45 am – 12:30 pm, Talk Room 1

Object Recognition: Categories and neural mechanisms

Moderator: Chaz Firestone, Johns Hopkins University

Talk 1, 10:45 am, 22.11

BEYOND BUTTON PRESSES: LARGER MOTOR ACTIONS FACILITATE VISUAL CATEGORY LEARNING

Luke Rosedahl[1], Takeo Watanabe[1]; [1]Brown University

Visual category learning has traditionally been viewed as a primarily perceptual-cognitive process, with popular models like the generalized context model, ALCOVE, and SUSTAIN treating it as largely independent from motor systems. Our research challenges this perspective by demonstrating that the magnitude of motor response in visual categorization tasks significantly impacts learning. Participants categorized Gabor discs in a virtual reality paradigm using either controller triggers or lightsaber swings. We tested two category structures: Rule-Based (RB), requiring binary decisions along visual feature dimensions, and Information-Integration (II), requiring integration of multiple feature dimensions (Rosedahl, Eckstein, and Ashby, 2018). Based on evidence that II structures engage basal ganglia systems (Ashby and Ennis, 2006), we hypothesized that larger motor movements might enhance II category learning. In our initial experiment, participants (N=72) were evenly divided across four conditions: RB-Button, RB-Lightsaber, II-Button, and II-Lightsaber. While RB performance remained consistent across response types ($t = -.12$, $p = .91$), II category learning significantly improved with lightsaber swings ($t = 2.7$, $p = .006$). To determine whether the benefits came from larger movement, interacting with the stimulus, or longer stimulus presentation times for the swing condition, we conducted a follow-up experiment where participants (N=30) punched a response box or used the controller triggers to learn the II categories with set stimulus presentation time. The punch group showed significantly faster learning ($t = 3.19$, $p = .002$) and higher final block performance (90% vs. 80%; $t = 2.80$, $p = .01$), indicating that large movements enhance category learning. These findings challenge traditional views on visual category learning and highlight how motor engagement shapes visual pattern classification, particularly for complex categories. This suggests the need to revise existing theories of visual category learning to account for motor system involvement, possibly through enhanced engagement of basal ganglia circuits.

This work was supported by the National Eye Institute of the National Institutes of Health under award numbers [K99EY034891, R01EY019466, R01EY027841, and R01EY031705] and NSF-BSF under award number [BCS2241417].

Talk 2, 11:00 am, 22.12

OBJECT BLINDSIGHT: FUNCTIONAL AND NEURAL PROCESSING OF UNSEEN VISUAL OBJECTS IN THE CORTICALLY BLIND FIELD

Jessica M. Smith[1], Bradford Z. Mahon[1]; [1]Carnegie Mellon University

Cortical blindness refers to the loss of the phenomenal experience of seeing that is caused by lesions along the geniculostriate pathway. In blindsight, visual inputs can continue to be processed by extra-geniculostriate and geniculo-extrastriate pathways bypassing the lesion, and can affect behavior in the absence of awareness. Previous studies have investigated blindsight principally using simple stimuli, such as gabors or simple shapes defined by chromatic or luminance contrast. Here, we set out to test functional and neural signatures of visual object processing in the cortically blind field. Hemianopic participants ($n = 2$) and controls ($n = 54$) completed a series of vision psychophysics and fMRI tasks in which neutral faces, fearful faces, and familiar graspable objects (e.g., pencil) were presented in seeing and blind regions of the visual field. The psychophysics studies used the redundant target paradigm, in which participants indicate (button push) every time they see anything appear; it is classically observed that participants are faster to respond to redundant visual targets compared to a single visual onset. While healthy controls exhibited classic redundancy gains, a pattern of redundancy loss was observed when redundant stimuli were presented in the hemianopic field. This redundancy loss indicates that visual stimuli presented in the hemianopic field are processed. Both hemianopic participants had lesions affecting V1, but one participant’s lesion involved the dorsal occipital cortex, while the other participant’s lesion spared this region. We found that graspable objects (tool images) presented in the hemianopic field significantly activated posterior intraparietal sulcus (IPS) only in the patient with spared dorsal occipital cortex. These observations, together with other findings in the field, indicate that nongeniculostriate pathways into the dorsal stream automatically compute the “graspability” of visual objects, even in the absence of an explicit goal to act.

Talk 3, 11:15 am, 22.13

SPATIOTEMPORAL DYNAMICS INDUCED BY RAPID PERCEPTUAL LEARNING IN THE HUMAN BRAIN AT SINGLE-NEURON RESOLUTION

Marcelo Armendariz[1,2] (marcelo.armendarizgil@childrens.harvard.edu), Julie Blumberg[3], Jed Singer[1], Franz Aiple[3], Jiye Kim[1], Peter Reinacher[3], Andreas Schulze-Bonhage[3], Grabriel Kreiman[1,2]; [1]Boston Children’s Hospital, Harvard Medical School, Boston, MA, USA., [2]Center for Brains, Minds and Machines, Cambridge, MA, USA., [3]University Medical Center Freiburg, University of Freiburg, Germany.

Humans can swiftly learn to recognize visual objects after just one or a few exposures. A striking example of rapid learning is the sudden recognition of a degraded black-and-white image of an object (Mooney image). These degraded Mooney images are initially unrecognizable. However, Mooney images become easily interpretable after a brief exposure to the original intact version of the image. This rapid learning process necessitates the formation of enduring neural signatures to enable subsequent recognition. Despite extensive behavioral characterization, the neuronal mechanisms underlying perceptual changes induced by rapid learning in the human brain are not well understood. Here, we recorded the spiking activity of neurons in medial occipital and temporal regions of the human brain in patients performing an image recognition task that involved rapid learning of degraded two-tone Mooney images. Neurons in the occipital cortex (OC) and medial temporal lobe (MTL) modulated their firing patterns to encode the identity of recently learned images. Population decoding revealed that occipital neurons resolved the identity of learned images at the cost of additional processing time, with delayed responses observed in MTL neurons. Our findings suggest that OC may not rely on feedback from MTL to support recognition following rapid perceptual learning. Instead, learning-induced dynamics observed in OC may reflect extensive recurrent processing, potentially involving top-down feedback from higher-order cortical areas, before signals reach the MTL. These results highlight the need for further computation beyond bottom-up visual input representations to facilitate recognition after learning and provide spatiotemporal constraints for computational models incorporating such recurrent mechanisms.

Talk 4, 11:30 am, 22.14

TRIPLE-N DATASET: NON-HUMAN PRIMATE NEURAL RESPONSES TO NATURAL SCENES

Yipeng Li[1] (moonl@pku.edu.cn), Jia Yang[1], Wei Jin[1], Wanru Li[1], Baoqi Gong[1], Xieyi Liu[1], Kesheng Wang[1], Jingqiu Luo[1], Pinglei Bao[1]; [1]Peking University

Understanding the neural mechanisms of visual perception requires data that encompass both large-scale cortical activity and the finer details of single-neuron dynamics. The Natural Scenes Dataset (NSD) has provided substantial insights into visual processing in humans (Allen et al., 2022), yet its reliance on functional magnetic resonance imaging (fMRI) limits the exploration of individual neuron contributions. To bridge this gap, we present a new dataset: the triple-N dataset, that extends the NSD framework to non-human primates, incorporating single-neuron activity and local field potential recorded from the inferotemporal (IT) cortex. We recorded the neuronal response while macaques passively viewed 1000 NSD shared images with Neuropixels. Over 60 sessions across 15 sub-regions within the IT cortex were recorded from 5 macaques, capturing over 14,000 neural units with good reliability (split-half correlation > 0.4), including approximately 2,000 single neurons. Many recordings were obtained from fMRI-defined category-selective regions, such as face, body, scene, and color-selective areas. Our dataset enables in-depth exploration of neural responses at multiple levels, from population dynamics to single-neuron activity, offering new insights into the spatiotemporal aspects of visual processing. First, most neurons within category-selective regions exhibit similar tuning properties, but a subset of neurons shows responses that cannot be explained by the population response, reshaping our understanding of the neuronal composition within category-selective areas. Second, with neuron-voxel mapping, our dataset provides a foundation for cross-species comparisons and alignment between human and macaque visual processing. Furthermore, given the diversity in the dynamic changes of neuronal responses and encoding features in the IT cortex, our dataset facilitates the development of computational models of the high-level visual system that emphasize the temporal characteristics of visual processing. Overall, our dataset serves as a valuable resource for advancing our understanding of visual perception and bridging the gap between large-scale neuroimaging and fine-grained electrophysiological signal.

Talk 5, 11:45 am, 22.15

CONTROLLING FOR EVERYTHING: CANONICAL SIZE EFFECTS WITH IDENTICAL STIMULI

Chaz Firestone[1], Tal Boger[1]; [1]Johns Hopkins University

Among the most impressive effects in recent vision science are those associated with “canonical size”. When a building and a rubber duck occupy the same number of pixels on a display, the mind nevertheless encodes the real-world size difference between them. Such encoding occurs automatically, organizes neural representations, and drives higher-order judgments. However, objects that differ in canonical size also differ in many mid- and low-level visual properties; this makes it difficult—and seemingly impossible—to isolate canonical size from its covariates (which are known to produce similar effects on their own). Can this challenge be overcome? Here, we leverage a new technique called “visual anagrams”, which uses diffusion models to generate static images whose interpretations change with image orientation. For example, such an image may look like a rabbit in one orientation and an elephant when upside-down. We created a stimulus set of visual anagrams whose interpretations differed in canonical size; each image depicted a canonically large object in one orientation but a canonically small object when rotated, while being pixel-wise identical in every other respect. Six experiments show that most (though not all) canonical size effects survive such maximal control. Experiments 1–2 tested Stroop effects probing the automaticity of canonical size encoding; consistent with previous findings, subjects were faster to correctly judge the onscreen size of an object when its canonical size was congruent with its onscreen size. Experiments 3–4 tested effects on viewing-size preferences; consistent with previous findings, subjects chose larger views for canonically larger objects. Experiments 5–6 tested efficient visual search when targets differed from distractors in canonical size; departing from previous findings, we found no such search advantage. This work not only applies a long-awaited control to classic experiments on canonical size, but also presents a case study of the usefulness of visual anagrams for vision science.

NSF BCS 2021053, NSF GRFP

Talk 6, 12:00 pm, 22.16

COMPARING THE MULTIDIMENSIONAL MENTAL AND NEURAL REPRESENTATIONS OF OBJECT WORDS AND OBJECT IMAGES

Tonghe Zhuang[1,2], Laura M. Stoinski[2,3,4*], Ming Zhou[5], Martin N. Hebart[1,2,6]; [1]Department of Medicine, Justus Liebig University, Giessen 35390, Germany, [2]Max Planck Institute for Human*

*Cognitive & Brain Sciences, Leipzig 04103, Germany, [3]University of Leipzig, Leipzig 04103, Germany, [4]International Max Planck Research School on Cognitive NeuroImaging (IMPRS CoNI), [5]Beijing Normal University, China, [6]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen and Darmstadt, Germany, * equal contribution*

Understanding the dimensions underlying mental and neural representations of objects is crucial for uncovering the mechanisms that link perception to semantic knowledge. However, vision and semantics are strongly interrelated, and the degree to which seemingly semantic dimensions (e.g., animacy) can also be explained by visual features (e.g., curvature) has made it challenging to tease apart the effects of vision and semantics in behavior and brain activations. To address this challenge, we examined the mental and neural representations of 1,388 object words and compared them to the representation of matched object images, thereby disentangling the contribution of visual-perceptual features evoked by images from visual-semantic features evoked by words. To this end, we first collected 1.3 million perceived similarity judgments of 1,388 object nouns in a large online sample, which allowed us to identify 50 object dimensions specifically related to words. These word dimensions showed strong overlap with 49 dimensions previously identified on object images (Hebart et al., 2020), but were restricted to high-level semantic information and object shape and did not include color or texture information. This highlights the importance of using images for evoking relevant visual dimensions in similarity judgment. Next, to examine the neural representations of these dimensions, we collected a densely sampled fMRI dataset of 480 object images and 960 matched image pairs in five participants across 15 sessions (4 word-based sessions, 8 image-based sessions, 3 localizer and structural sessions). By mapping the dimensions of words and images to brain activity patterns, we were able to identify the cortical regions related to the mental representation of object images, object words, and their overlap. Together, this work highlights the interplay of vision and semantics in mental and neural object representations and establishes a large, multimodal dataset to support future research on the intersection of vision, semantics, and neural representation.

This work was supported by SFB-TRR135 “Sonderforschungsbereich SFB/Transregio TRR”, ERC Starting Grant project COREDIM (ERC-StG-2021-101039712) and the Hessian Ministry of Higher Education, Science, Research and Art (LOEWE Start Professorship to M.N.H. and Excellence Program ‘The Adaptive Mind’).

Talk 7, 12:15 pm, 22.17

OBJECT DIMENSIONS UNDERLYING FOOD SELECTIVITY IN VISUAL CORTEX

Davide Cortinovis[1] (davide.cortinovis@unitn.it), Giulia Orlandi[1], Lotte Van Campenhout[1,2], Stefania Bracci[1]; [1]University of Trento, Italy, [2]KU Leuven, Belgium

The occipitotemporal cortex (OTC) has traditionally been viewed as functionally organized into category-selective areas, such as those responding to faces, body parts, and scenes. More recent studies using the Natural Scenes Dataset identified food-selective areas adjacent to face-selective areas (in both lateral and medial OTC), independent of basic visual features like shape, texture or color. However, other evidence found overlapping activations between food and tool responses, suggesting that food-selectivity could be better understood through a dimensional framework that emphasizes shared properties like manipulability. Our study explored the dimensions underlying food-selective areas in the OTC using fMRI and a stimulus set including images of faces, bodies, hands, food, tools, manipulable objects, scenes, and spiky meaningless objects. For food and object categories, both grayscale and colored images in different configurations were presented to assess the roles of visual (e.g., color, clutter-complexity) and action-related (e.g., graspability, effector-specificity) properties. Our localizer identified two distinct food-selective clusters in OTC: one medial, localized between regions selective for faces and scenes, and one lateral, partially overlapping with regions selective for tools and manipulable objects. In lateral OTC, no significant overlap was found between hand and food selectivity. However, we replicated the previously known hand-tool overlap, indicating that tools and hands share effector-specific information absent for food. Moreover, visual properties like object clutter and, to a lesser extent, color contributed to the representations in the medial (but not lateral) food cluster. Finally, computational models of visual cortex topography only partially captured the observed organization of food-selective areas, with similar representation of visual properties but no organization based on action information. Overall, our results show that food responses in OTC may be better understood in the light of a dimensional framework that considers both the visual and the action-related properties of food, going beyond a category-centric framework.

Talk Session: Saturday, May 17, 2025, 10:45 am – 12:30 pm, Talk Room 2

Temporal Processing

Moderator: Rachel Denison, Boston University

Talk 1, 10:45 am, 22.21

THE SPEED LIMIT OF VISUAL PERCEPTION: BIDIRECTIONAL INFLUENCE OF IMAGE MEMORABILITY AND PROCESSING SPEED ON PERCEIVED DURATION AND MEMORY RECALL

Martin Wiener[1] (mwiener@gmu.edu); [1]George Mason University

Visual stimuli are known to vary in their perceived duration, with some stimuli engendering so-called “time dilation” and others “time compression” effects. Previous theories have suggested these effects rely on the level of attention devoted to stimuli, magnitude of the stimulus dimension, or intensity of the population neural response, yet cannot account for the full range of experimental effects. Recently, we demonstrated that perceived time is affected by the image properties of scene clutter, size, and memorability (Ma, et al. 2024), with the former compressing and latter two dilating duration. Further, perceived duration also predicted recall of images 24h later, on top of memorability. To explain the memorability effect, we found that a recurrent convolutional neural network (rCNN) could recapitulate the time dilation effect by indexing the rate of entropy collapse, or “speed”, across successive timesteps, with more memorable stimuli associated with faster speeds. Here, we replicate and extend these findings via three experiments (n=20ea.) where subjects performed a sub-second temporal bisection task using memorability stimuli with increasing memorability, but a constant speed (exp 1), increasing speed, but constant memorability (exp 2), or increasing in both (exp 3), each followed by a surprise memory test 24hr later. We found that either increasing memorability or speed alone led to time dilation effects, with faster/slower speeds shifting memory recall by 10% in either direction. However, when both metrics increased, memorability dilated time while speed compressed it, while still improving recall overall. These findings can be explained by a model wherein the visual system is tuned to a preferred speed for processing stimuli that scales with the magnitude of visual response, such that stimuli closer to this speed are dilated in time. Overall, these findings provide a new lens for interpreting time dilation/compression effects and how visual stimuli are prioritized at temporal scales.

TALK 2, 11:00 AM, 22.22

AN ILLUSION OF TIME CAUSED BY REPEATED VISUAL EXPERIENCE

Brynn E. Sherman[1] (brynns@sas.upenn.edu), Sami R. Yousif[2]; [1]University of Pennsylvania, [2]University of North Carolina, Chapel Hill

The feeling that something happened "only yesterday" makes us feel attached to it — as if it is as much a part of us as the present moment. The feeling that an event occurred long ago enhances our sense of nostalgia (and heightens our awareness that time is always passing, whether we like it or not). But how do we remember when we saw something? One obvious possibility is that, in the absence of explicit cues, we infer elapsed time based on memory strength. If a memory is fuzzy, it likely occurred longer ago than a memory that is vivid. Here, we demonstrate a robust illusion of time that stands in stark contrast with this prediction. In six experiments, we show that experiences which are visually repeated (and, consequently, better remembered) are counterintuitively remembered as having initially occurred earlier in time. This illusion is robust (amounting to as much as a 25% distortion in perceived time), consistent (exhibited by the vast majority of participants tested), stable across a variety of paradigms (e.g., when participants are asked to place seen items on a timeline and also when participants explicitly judge which of two items was seen earlier), immune to various experimental interventions (e.g., encouraging participants to pay attention to each specific presentation of an item), and applicable at the scale of ordinary day-to-day experience (occurring even when participants are tested over one full week). Thus, this "temporal repetition effect" may be one of the key mechanisms underlying why it is that people's sense of time often diverges from reality.

TALK 3, 11:15 AM, 22.23

SYNCHRONIZATION OF VISUAL PERCEPTION WITHIN THE HUMAN FOVEA

Annalisa Bucci[1,2,3] (annalisa.bucci@iob.ch), Marc Büttner[1,2], Niklas Domdei[4], Federica B. Rosselli[1,2], Matej Znidaric[1,3], Julian Bartram[3], Tobias Gänswein[3], Roland Diggelmann[3], Martina De Gennaro[1], Cameron S. Cowan[1], Wolf Harmening[4], Andreas Hierlemann[3], Botond Roska[1,2], Felix Franke[1,2,3]; [1]Institute of Molecular and Clinical Ophthalmology Basel (IOB); 4031 Basel, Switzerland, [2]University of Basel, Faculty of Science, 4031 Basel, Switzerland, [3]Eidgenössische Technische Hochschule Zürich (ETH), Department of Biosystems Science and Engineering (D-BSSE); 4056 Basel, Switzerland, [4]Rheinische Friedrich-Wilhelms-Universität Bonn, Department of Ophthalmology; 53127 Bonn, Germany

Precise timing of action potentials underpins the processing of visual information. Retinal ganglion cells (RGCs), the output neurons of the retina, encode visual information into action potentials that propagate to higher visual areas in the brain. Within the eye, the intraretinal segments of RGC axons remain unmyelinated, resulting in slow action potential propagation. The lengths of these unmyelinated axon segments are determined by their intraretinal trajectories, which are shaped by the anatomical organization of the human eye including the fovea — a specialized retinal region responsible for high-acuity vision. To achieve high-acuity vision in the foveal center ('umbo'), all retinal circuitry, except for the photoreceptors, is displaced into a ring-like structure around the umbo. Consequently, axons originating from RGCs on the temporal side of the fovea must bend around the umbo, whereas nasal RGCs can connect straight to the optic disc. This organization causes neighboring photoreceptors in the center of our vision to connect to RGCs with dramatically different intraretinal axonal lengths. This raises the question: do differences in lengths lead to a temporal dispersion of the arrival times of visual information in the brain? To address this question, we measured human reaction times to single-cone photostimulation in the umbo. Reaction times were uniform across the central visual field. Using high-density microelectrode arrays (HD-MEAs) on human retinal explants, we recorded foveal RGC action potentials and found propagation speeds varied based on the location of RGC somas around the umbo. Axons originating temporal to the umbo exhibited more than 40% higher propagation speeds than those on the nasal side. Transmission electron microscopy revealed these higher speeds were associated with larger axon diameters. A model accurately predicted axonal paths and lengths, which strongly correlated with observed propagation speeds. These findings reveal a compensatory mechanism in the human retina that synchronizes visual perception.

Funded by SNSF (CRSII5_173728, CRSII5_216632; PCEFP3_187001; CRSK-3_220987, CRSK-3_221257; 310030_220209); ERC (neuroXscales, 694829); DFG (Ha5323/5-1; SPP2127 Ha5323/6-1); Carl Zeiss Foundation (HC-AOSLO); Sedinum Foundation. We thank Universitätsspital Basel, donors and their families for support.

TALK 4, 11:30 AM, 22.24

WHEN A BLUNT EVENT IS PERCEIVED DEPENDS ON ITS TEMPORAL PROFILE

Ljubica Jovanovic[1], Pascal Mamassian[1]; [1]Laboratoire des Systèmes Perceptifs, ENS, PSL University, CNRS

When visual objects strike our retinas, they trigger a cascade of activity of different durations and delays along the visual processing hierarchy, making it difficult to predict exactly when they are perceived (Nishida

& Johnston, 2002; Curr. Biol.). Here we investigate which temporal features of the visual stimulus are used to infer when an event is perceived. We presented two events in succession, each event being a pair of Gaussian blobs (1 dva) located at a fixed eccentricity on either side of fixation. We varied the delay between the events (800 or 1200 ms), and participants were instructed to reproduce this delay by pressing a key at the appropriate time after the second event. The orientation of the blobs was random for the first pair, and rotated by 120 degrees for the second. We manipulated the temporal profiles of the events to investigate when they are perceived. Contrast of the first event followed a Gaussian modulation. Contrast of the second event had either a Gaussian, or an asymmetrical temporal profile built from a weighted sum of two Gaussians shifted in time by 100 ms. These stimuli allowed us to independently vary the times of the maximum and mean contrasts. We found a strong effect of the temporal contrast modulation on the perceived time. Depending on the direction of the contrast distribution's skewness, the second pair was perceived earlier (~80 ms) or later (~30 ms) relative to a pair with a symmetrical Gaussian contrast distribution and identical maximum contrast time. From this pattern of results, the information used for estimating when an event occurred cannot be a threshold intensity or the time of maximum contrast. Instead, the time when integrated contrast reaches a threshold appears relevant for when an event occurred (Amano et al., 2006; J. Neurosci.).

ANR grant no. ANR-22-CE28-0025

Talk 5, 11:45 am, 22.25

THE ROLE OF NEURAL OSCILLATIONS IN VISUAL HIERARCHY FOR DURATION PERCEPTION

Yuko Yotsumoto[1], Amirmahmoud Houshmand Chatroudi[1]; [1]The University of Tokyo

'Time' is actualized within a sensory context, making it vulnerable to distortions arising from sensory organization. One such distortion is the tendency to overestimate the duration of visual flickers, a phenomenon known as flicker-induced time dilation (FITD). A decade of research has led to two predominant hypotheses for explaining this temporal illusion: subjective salience (Herbst et al., 2013) and neural entrainment (Hashimoto & Yotsumoto, 2018). However, evidence supporting the neural entrainment hypothesis, particularly through steady-state evoked potentials (SSVEPs)—oscillatory neural responses to regular flickers—has been inconsistent (Li et al., 2020). In this study, we employed the semantic wavelet-induced frequency tagging (SWIFT; Koenig-Robert & VanRullen, 2013) technique to investigate whether the cortical localization of SSVEPs within the visual hierarchy could account for the inconsistency between FITD and the entrainment hypothesis. Using SWIFT, we generated a set of flickers characterized by luminance-based, semantic-based, and combined luminance and semantic properties (hierarchical frequency tagging; Gordon et al., 2017). EEG results revealed that each flicker type elicited distinct SSVEP activation patterns in the occipitotemporal regions, indicating selective engagement of different levels of the visual hierarchy. However, the magnitude of FITD did not differ across flicker conditions. Furthermore, SSVEP amplitude in none of the conditions correlated with FITD. This clear dissociation between neural activation patterns and the extent of the flicker-induced illusion challenges the role of entrainment in explaining FITD. Notably, the FITD magnitude observed in the experimental flicker conditions (luminance, semantic, and combined flickers) was comparable to that of the scrambled control condition. This finding fundamentally challenges existing theories of time perception that seek to explain temporal illusions, suggesting the need to revisit and reevaluate the core mechanisms underlying FITD.

JSPS KAKENHI 23K22372, 23KK0046

Talk 6, 12:00 pm, 22.26

PHASE-DEPENDENT EEG DECODING OF SUSTAINED VISUAL INFORMATION

Michele Deodato[1] (md5050@nyu.edu), David Melcher[1]; [1]New York University Abu Dhabi

Vision research often emphasizes brief stimulus presentations. For example, when it comes to detecting or integrating flashed stimuli, pre-stimulus power and phase of EEG alpha oscillations (8–12 Hz) have been shown to influence neural and behavioral responses. However, in natural viewing, stimuli are often present for extended periods of time, raising the question of how visual representations are maintained and the role of neural oscillations in this maintenance. In this study, we recorded EEG from participants viewing sustained (2-second) Gabor stimuli with varying orientations (left vs. right) and spatial frequencies (low vs. high). Initial analyses, including decoding and event-related potentials, demonstrated significant neural representation for spatial frequency (but not orientation) only for the first 500–1000 ms, despite the stimuli persisting beyond this period and participant not reporting any visual fading. This raises the question of how is visual information stored and maintained into consciousness beyond this initial period. To examine oscillatory contributions to visual maintenance, we implemented a novel decoding approach targeting the 1000–2000 ms time window. Specifically, EEG decoding of stimulus spatial frequency was conducted separately for data points corresponding to different phases of alpha oscillations at each channel location. Strikingly, we found that decoding accuracy varied with the phase of alpha oscillations at frontal and occipito-parietal channels, suggesting that visual information is periodically reactivated during sustained perception. Our phase-specific decoding method underscores the potential of leveraging oscillatory dynamics to study information processing over time in the brain. These findings provide compelling evidence on the role of alpha oscillations in the maintenance of visual information, highlighting their importance in sustained visual processing.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYU Abu Dhabi Research Institute grant CG012, Part of the work was conducted at the Brain Imaging lab within the Core Technology Platforms at NYU Abu Dhabi.

Talk 7, 12:15 pm, 22.27

TESTING TECHNIQUES TO DISCRIMINATE THE MAGNOCELLULAR DIVISION OF THE VISUAL AND AUDITORY THALAMUS.

Josiane Mukahirwa[1] (hijosie@udel.edu), Qianli Meng, Jaeseon Song, Andrew Lisech, Keith Schneider; [1]University of Delaware

Introduction. The magnocellular (M) pathway plays a vital role in the visual and auditory systems, specializing in the processing of transient stimuli. Studies such as those by DeSimone & Schneider (2019) and Meng and Schneider (2022) have investigated transient responses in the lateral geniculate nucleus (LGN) and the magnocellular division of the medial geniculate nucleus (MGN). Yet, this pathway remains challenging to isolate. This gap in understanding may be attributed to the small size of the LGN and MGN, which range from approximately 90–180 mm^3, and the difficulty in generating isolating stimuli for the two pathways. In this study, we focused on the responses to transient stimuli. Methods. Nineteen participants, including individuals with (9) and without dyslexia (11), underwent fMRI scanning using a 3T scanner. Activation maps were generated for transient stimuli, with abrupt onsets and offsets, and sustained stimuli, with smooth transitions. Regions of interest (ROIs) for the LGN and MGN were manually traced on T1 and proton density images. We tested a variety of algorithms, including principal component analysis, clustering algorithms, and multivariate methods, to identify voxels showing a substantial preference for transient stimuli. Results: In the LGN in typical readers, we identified a cluster of voxels showing a preferential response to transient stimuli on the ventral edge, likely corresponding to the magnocellular layers. In the MGN, PCA analysis showed a subset of voxels with a high transient index. These specialized responses to transients were generally absent in subjects with dyslexia. Conclusion: These findings support the existence of functionally distinct M pathways in both the visual and auditory systems. This study shows the potential of transient stimuli and multivariate analysis to explore magnocellular function and its role in sensory disorders like dyslexia. Future work should explore the clinical implications for dyslexia diagnosis and interventions.

Talk Session: Saturday, May 17, 2025, 2:30 – 4:15 pm, Talk Room 1

Moderator: Edward Ester, University of Nevada, Reno

Visual memory: General

Talk 1, 2:30 pm, 24.11

INFORMATION INTEGRATION IN WORKING MEMORY

Qihang Zhou[1], Jinglan Wu[1], Tengfei Wang[1], Yuzheng Hu[1], Fuying Zhu[1], Hui Zhou[1], Mowei Shen[1], Zaifeng Gao[1]; [1]Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China

Understanding how the cognitive system integrates discrete sensory inputs into coherent representations is a central question in psychology. Working memory (WM) plays a key role in this process, yet current WM models (e.g., Baddeley, 2012; Cowan, 2001) primarily focus on storage buffers and executive functions, often overlooking information integration. We propose a novel component, the integration buffer, which is specifically responsible for integrating elemental information into unified representations through a compression mechanism. We conducted a total of five studies. Studies 1-3 provided evidence for the existence of the integration buffer. In Study 1, 176 participants completed two integration tasks alongside ten other WM tasks representing the established components. Confirmatory factor analysis indicated that WM integration could not be attributed to any existing component (e.g., visuospatial sketchpad, central executive, or episodic buffer), supporting the distinctiveness of the integration buffer. Study 2 demonstrated that cognitive load manipulations on the visuospatial sketchpad, central executive, and episodic buffer had no impact on WM integration, further supporting the buffer's independence. In Study 3, using multi-modal magnetic resonance imaging (MRI), we provided complementary evidence from brain activity patterns. Studies 4 and 5 further clarified the functional role of the integration buffer through eye tracking and event-related potentials. We found that participants directed more attention to the center of the integrated representation (a more centralized fixation pattern; study 4) and had lower memory load (reduced contralateral delay activity; study 5) during WM maintenance compared to memorizing discrete items, suggesting a compressive integration process. Together, these findings support the existence of a distinct integration buffer, which integrates discrete elements into a coherent representation via a compression mechanism.

Talk 2, 2:45 pm, 24.12

BEYOND SWAPS: HOW WORKING MEMORY COMBINES GIST AND ITEM INFORMATION

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

Working memory representations of multiple items are not encoded independently, but interact with each other across different hierarchical levels of representation. For example, individual item representations can be biased toward gist-level representations (Chunharas & Brady, 2023). Yet their potential benefits for memory performance remain unclear. While some researchers hypothesize that combining gist and item-based representations leads to more precise memory with increased gist-based bias, quantitative evidence supporting this hypothesis is lacking. To address this gap, we developed a computational model that implements a weighted summation of the familiarity signals that contribute to gist and single-item representations. Our model predicts the full shape of error distributions in tasks where people use gist and item memories. This modeling revealed that previously observed precision-bias trade-off can be explained by increased weighting of gist representations when stimuli are highly similar. The model further predicted that within-trial stimuli would show variable precision depending on their similarity to the gist-level representation. We validated these predictions by fitting our model to existing experimental data (Utochkin & Brady, 2020). As predicted, we found an inverse relationship between gist representation weighting and stimulus similarity. Critically, reanalysis of the data confirmed that stimuli more similar to the gist showed better precision than dissimilar items within the same trial, aligning with our model's predictions. To determine whether chunking occurs through representation combination or swap-like replacement of individual items, we compared our model against two alternatives: a gist-only model and an individual item-only model. Our model outperformed both alternatives, suggesting that chunking in working memory operates through adaptive combination of single-item and gist-level representations rather than through simple replacement. Overall, the model provided a unified account of how working memory balances between precision and compression through dynamic weighting of representation combination. Furthermore, the model demonstrated a connection between chunking and ensemble perception.

TALK 3, 3:00 PM, 24.13

THE SPATIAL, CATEGORICAL, AND VERBAL REPRESENTATIONS UNDERLYING VISUAL WORKING MEMORY IN THE HUMAN BRAIN

Thomas Christophel[1,2], Andreea-Maria Gui[1,2], Carsten Allefeld[3], Vivien Chopurian[1,2], Joana Seabra[1,2]; [1]Humboldt Universität zu Berlin, [2]Berlin Center of Advanced Neuroimaging, [3]City, University of London

The division between visual and non-visual storage is foundational to the conception of visual working memory. Recent work, however, suggests that working memory relies on multiple concurrently held representations across multiple cortical regions. These regions are believed to enact a division of labor where some regions harbor near-veridical representations of memorized stimuli while other representations capture the stimulus in abstract form. Here, we demarcate this division of labor using fMRI and multivariate encoding modelling. In a large sample (N = 40), we measure patterned BOLD activity for cortical representations of memorized orientations, spatial locations, and words that can be used to describe orientations (e.g., "vertical"). We identify markers in patterned cortical activity identifying the spatial, categorical and verbal representations underlying visual working memory. For spatial representations, we find cross-classification between location and orientation working memory explaining a substantial part of stimulus-specific activity even in visual cortical regions. These spatial representations are subject to retinotopic distortions that are most pronounced in anterior regions and later in the delay. Categorical and veridical representations appear to coexist independently in anterior cortical regions, while in visual regions, veridical, sensory-like encoding models outperform categorical encoding models. Finally, we find cross-classification between orientation representations and spatial language, showing that neural activity during visual working memory in part mimics activity during language processing. In this way, we demonstrate the unique contribution of a diverse set of cortical regions to visual working memory storage.

This work was supported by DFG Emmy Noether Research Group Grant CH 1674/2-

TALK 4, 3:15 PM, 24.14

VISUAL WORKING MEMORY CAN REPRESENT ITEMS BEFORE THEY APPEAR

Reut Peled[1] (reutgadot@gmail.com), Roy Luria[1,2]; [1]School of Psychological Science, Tel Aviv University, [2]Sagol School of Neuroscience, Tel Aviv University

One of our main characteristics as humans is to generate predictions. Still, the mechanisms supporting this ability remain poorly understood. This study demonstrates that Visual Working Memory (VWM) plays a crucial role in generating predictions by creating future-representations: actively representing stimuli even before they appear. Across four EEG experiments (N=90), participants performed a change-detection task involving moving objects. In the predictable condition, objects moved toward each other, forming converged objects, while in other control conditions, objects moved independently. We monitored the contralateral delay activity (CDA), a neural marker of VWM's sensitivity to the amount of visual information. Our key evidence for future-representations in VWM was an increased CDA in the predictable condition during objects' movement - before perceiving the converged objects. Control experiments ruled out alternative explanation for the increased CDA activity such as trajectory anticipation or representing the converging event itself. We used machine-learning analyses to provide evidence that the converged objects can be decoded from the EEG signal during the movement and before they converged. First, a supervised SVM decoder classified EEG activity from two perceptually identical conditions: in both conditions the objects moved towards each other, but only in one condition the objects converged while in the other, the objects crossed and continued moving independently ("crossing"). Although the two conditions were perceptually identical until the meeting event, and both meeting-results were predictable, classification accuracy exceeded 70%, suggesting participants represented additional information before the objects converged. Second, we used an unsupervised clustering algorithm, and demonstrated that the converging condition was classified as resembling to an already converged-items condition relative to the crossing condition (with ~70% accuracy), even during the movement before the meeting, when the crossing condition was perceptual identical to the merging condition. These findings provide strong evidence that VWM generate fututre-representations as part of predictive cognition.

TALK 5, 3:30 PM, 24.15

BRIDGING BAYESIAN AND REPRESENTATIONAL THEORIES OF MEMORY TO PREDICT MEMORY BIASES

Anxin Miao[1], Timothy Brady[2], Maria Robinson[3]; [1]University of Illinois of Urbana Champaign, [2]University of California San Diego, [3]University of Warwick

Understanding how people integrate gist and item-specific information in memory is crucial for understanding how memory changes with time and how people make memory-based decisions under conditions of uncertainty. In the current work, we investigated how people balance these two types of information in color memory tasks. We combined a Bayesian approach with prior signal detection-based models to predict gist memory based on performance in an ensemble and individual item memory task. Our model integrates gist and item-specific information by weighting them based on their relative discriminability, making predictions about how participants will rely on each source when asked to recall an item where both gist and item-specific information could be relevant. Participants completed three tasks: an Ensemble Task to estimate discriminability of gist (d'), an Individual Item Task to assess item-specific memory (d'), and a Gist Bias Task to evaluate memory recall for a specific item under different levels of offset between gist and item-specific colors. Our model makes parameter-free predictions of participants' performance in the Gist Bias Task using d' estimates derived from the other two tasks, combined via Bayesian weighting. Despite not being fit to the data from this task at all, this model captured the complete distributions of people's memory errors in the Gist Bias task as well as key qualitative trends in the data, such as the proportional shift of item-specific memory towards memory gist. Model comparisons and a permutation test showed that the model captured meaningful individual differences in memory integration (p<.001). These findings highlight the predictive power of bridging Bayesian models with representational theories of memory, offering a comprehensive framework for understanding how gist and item-specific information combine to create memory biases.

Talk 6, 3:45 pm, 24.16

EARLY VISUAL CORTEX IS RECRUITED TO ACT AS A COMPARISON CIRCUIT BETWEEN MENTAL REPRESENTATIONS AND VISUAL INPUTS

Maria V. Servetnik[1,2], Rosanne L. Rademaker[1]; [1]Ernst Strüngmann Institute for Neuroscience in Cooperation with Max Planck Society, Frankfurt, Germany, [2]Vrije Universiteit Amsterdam, Amsterdam, Netherlands

Imagine briefly losing your friend in a crowd: As you scan one face after another (perception), you hold your friend's face in mind (short-term memory) and compare it against each new input (visual search). This process spans multiple cognitive domains and requires mental representations of visual information at every step in order to find your friend again. Early visual cortex (EVC) has been implicated in maintaining such mental visual representations, both in the presence and absence of concurrent sensory input. It remains an outstanding question why EVC - the primary processing site of visual input - is also used for representing mental content. We hypothesize that EVC is recruited to serve as a comparison circuit, matching mental content to incoming sensory information in a visual format native to EVC. To test this, we collected fMRI data while participants (N=5) remembered the direction of a coherent random dot motion (RDM) stimulus for 8 minutes. During this period, participants saw 48 new RDM stimuli with independent motion directions. Before each of these probes, participants were cued to either compare its motion direction to the one they were holding in mind, or to withhold a comparison. Results show that BOLD responses for comparisons were higher than for non-comparisons throughout EVC. Importantly, a decoder trained on independently collected visual localizer data showed that making a comparison also rendered the motion direction held in mind significantly more decodable. This suggests that EVC is recruited to represent mental content especially when such content needs to be compared to visual input. The tentative role of EVC as a comparison circuit highlights that visual information temporarily held in mind is typically stored for use, and that a broader cognitive context must be considered in order to uncover how and why the brain represents mental content.

Talk 7, 4:00 pm, 24.17

TRAVELING WAVES OF HUMAN NEOCORTICAL ACTIVITY COORDINATE VISUALLY-GUIDED BEHAVIORS

Edward Ester[1], Canhuang Luo[2], Thane Houghton[1]; [1]University of Nevada, Reno, [2]Shenzhen University

Cortical traveling waves, or global patterns of activity that extend over several centimeters of the cortical surface, are a key mechanism for guiding the spatial propagation of neural activity and computational processes across the brain. Recent studies have implicated cortical traveling waves in successful short- and long-term memory encoding, storage, and retrieval. However, human memory systems are fundamentally-action oriented: eventually, the contents of memory must be utilized to produce appropriate behaviors. Cortical traveling waves could contribute to the production and control of memory-guided behaviors by flexibly routing information between brain areas responsible for storing memory content and brain areas responsible for planning and executing actions. Here, using short-term memory as a test case, we report evidence supporting this possibility. By applying image-based analyses to published human EEG studies, we show that the initiation of a memory-guided behavior is accompanied by a low-frequency (2-6 Hz) feedforward (occipital to frontal) traveling wave that predicts intra- and inter-individual differences in response onset, while the termination of a memory-guided behavior is followed by a higher frequency (14-32 Hz) feedback (frontal-to-occipital) traveling wave. Neither waveform could be explained by nuisance factors, including passive volume conduction, feedforward propagation of visually-evoked responses, or eye movements. Moreover, both waveforms required an overt behavior: when participants selected task-relevant memory content and prepared but did not yet execute an action based on this content, neither waveform was observed. Our findings suggest a role for traveling waves in the generation and control of memory-guided behaviors by flexibly organizing the timing and direction of interactions between brain regions involved in memory storage and action.

NSF 2050833

Talk Session: Saturday, May 17, 2025, 2:30 – 4:15 pm, Talk Room 2

Face and Body Perception: Facial expressions, social relationships

Moderator: Suzanne Scherf, Penn State University

Talk 1, 2:30 pm, 24.21

EYES VS. ATTENTIVENESS: PUPIL DILATION WIDENS THE PERCEIVED CONE OF DIRECT GAZE

Clara Colombatto[1] (clara.colombatto@uwaterloo.ca), Sarah D. McCrackin[2], Brian Scholl[3], Jelena Ristic[2]; [1]University of Waterloo, [2]McGill University, [3]Yale University

One of the most important social signals we perceive is the direction in which another person looks, especially when they are looking at us. In fact, humans perceive a range of eye-gaze deviations as directed towards (vs. away from) us, a range known as the Cone of Direct Gaze (CoDG). Does the CoDG reflect perceived eye direction per se, or might it indirectly reflect the degree to which we see another person *attending* to us? We explored this by asking whether the CoDG is affected by another salient property of others' eyes often linked to attention—how dilated the pupils are. We reasoned that perceived pupil dilation (vs. constriction) may lead to increased perception of the gazer attending to the observer, resulting in a wider CoDG. We tested this idea in two preregistered experiments (N=150 each). Observers viewed faces with constricted, normal, or dilated pupils, embedded in eyes looking with varying eccentricities to the left, right, or directly at the observer. The manipulation of pupil dilation was entirely task-irrelevant, as observers' task was simply to report the gaze direction that they saw. In an initial experiment, faces with dilated pupils were more likely to be perceived as gazing directly toward the observers, compared to faces with either normal or constricted pupils. When the faces were inverted in another experiment, however, this effect vanished—suggesting that the impact of pupil dilation on the CoDG is not just a function of the brute physical differences of the pupils themselves (which of course still exist even when inverted). We interpret this effect—that pupil dilation widens the CoDG—in terms of enhanced percepts of attentiveness. Judgments of gazing direction are thus influenced not just by physical properties of the eyes, but also by higher-level percepts of the cognitive states behind the eyes.

This project was funded by grants from NSERC to J.R. and C.C.

Talk 2, 2:45 pm, 24.22

MINIMAL EFFECTS OF STEREOPSIS ON PROCESSING REALISTIC FACES

Camille Proszanski[1], Erez Freud[1], Laurie M. Wilcox[1]; [1]York University

There is some evidence of an upper visual field advantage for face processing that has been taken as support for evolutionary pressures for face detection. However, these effects tend to be small and the outcomes variable; this may be due to the use of 2D images, which lack the volumetric 3D information present in the real world. Here, we evaluate the impact of naturalistic 3D face stimuli (relative to 2D), and their location in the visual field, on face detection and recognition. Stereopairs of photorealistic face stimuli were presented using a mirror stereoscope in 3D and 2D. In all experiments, the target was present in 50% of the trials, and proportion correct was used to compute sensitivity (d'). In Experiment 1 (N=28), we used a visual search paradigm; stimuli were presented in a semi-circular array in either the upper or lower visual field. The distractor faces were tilted 15 deg to the left (or right), and observers indicated if the target face (tilted in the opposite direction) was present. We varied the number of distractors, location and modality (2D vs. 3D). This low-level task showed no effect of modality and a weak effect of location. In subsequent experiments, we used the same stimuli but in high-level recognition-based tasks. We varied task difficulty, modality and also tested both upright and inverted faces (Experiment 2, N=22; Experiment 3, N=26). We found no effect of 3D viewing or location in either of the experiments, nor was there an interaction. The presence of a strong face inversion effect confirmed that observers were processing the faces holistically. Our results suggest that visual field asymmetries may only occur for tasks that rely on low-level properties. Further, the lack of effect of stereopsis implies that 2D images can be reasonable proxies for natural 3D faces.

NSERC grant RGPIN-2019-06694

Talk 3, 3:00 pm, 24.23

CONTINUOUS THETA-BURST STIMULATION OF THE POSTERIOR SUPERIOR TEMPORAL SULCUS AND INTERSUBJECT SYNCHRONY DURING NATURALISTIC VIEWING

James Thompson[1], Peter Kakalec[1], Courtney Marsh[1], Rebecca Roy[1]; [1]George Mason University

Sharing an understanding of commonly experienced events is important for forming and maintaining social connections. Friends share similar patterns of functional MRI activity (intersubject correlation, or ISC) when viewing the same naturalistic stimuli such as movies. The superior temporal sulcus (STS) is one region that might play an important role in forming the shared neural representations that underlie ISC. The STS tracks social interactions during task-based and naturalistic viewing. Here we propose that shared representations rely on the perception of social interactions in the pSTS. We used TMS to examine the contribution of the STS to ISC and encoding of social information during naturalistic viewing. In one session, participants viewed videos consisting of people, places, food, objects, and scrambled videos, to localize pSTS. Resting motor TMS thresholds were also acquired. In a second scanning session, we administered inhibitory (continuous) theta burst TMS (cTBS) to functionally-localized right pSTS or vertex (sham) in a between groups design, before participants viewed a 20min movie during multiband/multiecho fMRI scanning. After scanning, participants were asked to recall details from the movie. ISC and intersubject pattern similarity (ISPS) were calculated from fMRI responses from parcels derived using the Schaeffer 200 parcel atlas. Memory accuracy for social but not non-social details was lower following cTBS to the pSTS, relative to sham. Lower ISC and ISPS during movie watching following cTBS to the pSTS relative to sham TBS was observed in limbic and cortical-limbic regions, while increased in ISC following cTBS, relative to sham, was observed in parahippocampal and frontoparietal cortex. This study provides details about the causal role of the STS to ISC and ISPS across multiple brain networks during the viewing of naturalistic social stimuli, and helps identify the contribution of social perception to the formation of shared neural representations.

Talk 4, 3:15 pm, 24.24

SEMANTIC AND SOCIAL FEATURES DRIVE HUMAN GROUPING OF DYNAMIC, VISUAL EVENTS IN LARGE-SCALE SIMILARITY JUDGEMENTS

Kathy Garcia[1] (kgarci18@jh.edu), Leyla Isik[1]; [1]Johns Hopkins University

How do humans perceive naturalistic social scenes, especially in dynamic contexts? Similarity judgments offer critical insight into the mental representations humans use to perceive actions, form categories, and predict behavior. Real-world perception is complex and dynamic, involving features that static representations fully cannot capture. While prior research has focused on static scenes, less is understood about the features driving human similarity judgments in dynamic settings and their alignment with representations learned by deep neural networks (DNNs). To address this, we collected ~20,000 triplet odd-one-out similarity judgments from ~2.5 million possible unique triplets generated from a curated dataset of 250 three-second videos depicting everyday human actions—a scale far exceeding comparable studies. We then constructed a similarity matrix for all 250 videos, computing p(i, j), the likelihood of participants selecting stimuli i and j as similar when paired with different third videos. Finally, these similarity judgments were compared to human ratings across visual and social scene features, fMRI responses, video DNN extractions, and word embeddings derived from human-written captions, using representational similarity analysis (RSA). The results reveal striking patterns in human perception: word embeddings of video captions, reflecting how humans choose to describe the videos, showed the strongest correlation with overall similarity judgments, followed by ratings of "intimacy" (relational closeness) and video DNN embeddings. Similarity judgements correlated significantly with neural responses in regions, including EBA, LOC, STS (lateral stream) and FFA (ventral stream), but not early visual (EVC) or scene-specific (PPA) areas, underscoring the dominance of social over scene features. Together, these findings highlight the alignment between linguistic descriptions and cognitive strategies, revealing how verbal encodings of social features capture key mental processes shaping our understanding of dynamic social scenes.

TALK 5, 3:30 PM, 24.25

DEVELOPING A NON-HUMAN PRIMATE MODEL TO DISSECT NEURAL MECHANISMS OF HUMAN FACIAL EXPRESSION PROCESSING

Maren Wehrheim[1,2,3], Shirin Taghian[3], Hamidreza Ramezanpour[3], Kohitij Kar[3]; [1]Department of Computer Science, Goethe University Frankfurt, Frankfurt, Germany, [2]Frankfurt Institute for Advanced Studies (FIAS), Frankfurt, Germany, [3]Department of Biology, York University, Toronto, Canada

Understanding how the human brain processes facial expressions requires quantitative models that bridge neural mechanisms with behavior. While many qualitative descriptions exist, the field lacks tight coupling between neural hypotheses and behavioral measurements. To address this, we developed a comprehensive approach combining behavioral measurements, large-scale neural recordings in macaques, and computational modeling to investigate the computations underlying facial emotion discrimination. We first established a rigorous behavioral paradigm, using a binary human facial expression discrimination task across six emotions (360 images), comparing facial emotion recognition between humans and macaques. To probe the neural mechanisms, we conducted chronic multi-electrode recordings in the macaque inferior temporal (IT) cortex during passive viewing of emotional faces. Using 205 logistic regression decoders, we tested how different transformations of the IT population activity predicted behavioral error patterns. We evaluated a suite of artificial neural networks (ANNs) to identify computational models that mirror these neural processes. Significant image-by-image correlations (r=0.69, p<0.001) validated macaques as a suitable model for studying the neural basis of facial emotion processing. We found that macaque IT activity significantly predicted image-level behavioral responses in both humans (70-170 ms, R=0.49, p<0.001) and monkeys (70-110 ms, R=0.69, p<0.001). Additionally, we observed that traditional action-unit models for facial expression analysis are significantly less aligned with human behavior than other ANNs (e.g., ImageNet-trained models, CLIP, simCLR). In addition, ANN-IT representations most closely matched monkey IT, as assessed by representational similarity analyses, also predicted human behavior more accurately. These findings provide critical insights into the neural computations underlying facial emotion discrimination and establish macaques as a robust model for studying these processes. By integrating neural, behavioral, and computational insights, this work provides a critical step toward developing biologically plausible models of facial expression recognition.

KK has been supported by funds from the Canada Research Chair Program, the Simons Foundation Autism Research Initiative (SFARI, 967073), Brain Canada Foundation (2023-0259), the Canada First Research Excellence Funds (VISTA Program), and Google Research Award and MW by DFG – 414985841.

TALK 6, 3:45 PM, 24.26

INTIMATE RELATIONSHIP EXPERIENCE PREDICTS SENSITIVITY TO FACIAL EMOTIONAL EXPRESSIONS

Katherine A. Billetdeaux[1], Brittany E. Woodruff[1], K. Suzanne Scherf[1]; [1]The Pennsylvania State University

Facial expressions provide critical non-verbal communicative signals in social interactions. We have shown that emerging adulthood is an important period for developing sensitivity to socially complex expressions, like those that provide important signals about the status of romantic/sexual partnerships (e.g., sexual interest; Motta-Mena & Scherf, 2017). Here, we tested the hypothesis that ongoing experience in intimate partnerships during emerging adulthood is associated with increased sensitivity to perceive these complex facial expressions. Emerging adult participants (N = 410, ages 18-25) completed a relationship questionnaire and emotional expression perception task. The relationship questionnaire asked participants to self-report about the presence of ongoing romantic and/or sexual relationships. We derived three scores: relationship duration (in months), relationship commitment, and relationship intensity (duration x commitment). Facial emotional expression stimuli consisted of 4 basic (angry, fearful, happy, sad) and 4 socially complex (betrayed, brokenhearted, contempt, sexual interest) expressions taken from the CEED database (Benda & Scherf, 2020). Each expression was morphed with a neutral expression from the same actor to generate 12 images. On each trial, participants observed the neutral expression and one of the morphed stimuli. They picked the image that displayed "more expression." Perceptual thresholds were computed separately for each expression. Approximately half (58.3%) of the participants reported being in a romantic and/or sexual relationship. While there were no differences at the group level in perceptual thresholds for basic or complex expressions between those in a relationship and those not, we were primarily interested in whether ongoing experience in an intimate relationship predicted expression sensitivity. Among individuals in a current relationship, we found both the duration and intensity of the relationship negatively predicted thresholds to detect to both basic and complex expressions. These results suggest that as people gain experience in intimate partnerships, they become increasingly sensitive to both basic and complex facial expressions.

Talk 7, 4:00 pm, 24.27

PERCEPTION OF SOCIAL INTERACTIONS: COMPOSITIONAL FORCES AS UNDERLYING MENTAL REPRESENTATIONS

Yiling Yun[1] (yiling.yun@g.ucla.edu), Yi-Chia Chen[1], Shuhao Fu[1], Hongjing Lu[1]; [1]University of California, Los Angeles

When one shape moves closer to another on a screen, we naturally see social interactions, such as greetings. While this type of animation can be described using low-level visual features (e.g., location, speed, distance) or high-level semantic labels, we proposed that mid-level representations of forces are how humans effectively represent social interactions. The computational advantage lies in the compositional nature of forces: Multiple forces can combine through vector addition to jointly act on an entity. We tested such a force model that represents social interactions through two types of compositional forces: interactive forces, driven by interactions between agents; and self-propelled forces, driven by individual intentions. These forces are quantified using parameters that capture the strength of attraction and repulsion as well as the distance at which one transitions to the other. In Experiment 1, we used an odd-one-out task to collect human similarity judgments for 27 animations, each generated with distinct semantic labels. With a noise ceiling of .811, calculated by correlating split-half human responses over 50 iterations, the force dynamics model provided a strong correlation with human similarity judgments (r = .520). Ablation analyses revealed that interactive force features contributed more significantly than self-propelled force features. Human judgments were less aligned with other control models, including a model based on low-level visual features (r = .411), a deep learning model (LSTM) trained to discriminate social interactions (r = .329), and language-based models using semantic labels (r = .146). In Experiment 2, we tested a different set of animations and replicated the same pattern of results. These findings suggest that people interpret social dynamics through compositional forces driven by distinct goals. Furthermore, this study sheds light on the development of social perception, which may build upon perceptual processes underlying intuitive physics.

NSF BCS 2142269

Talk Session: Saturday, May 17, 2025, 5:15 – 7:00 pm, Talk Room 1

Decision Making

Moderator: Constantin Rothkopf, Centre for Cognitive Science, TU Darmstadt

Talk 1, 5:15 pm, 25.11

THE THEORY OF SUBJECTIVE INFLATION IN PERIPHERAL VISION: ACCEPT, REJECT, OR REVISE?

Brian Odegaard[1], Joseph Pruitt[1], Doyeon Lee[1], Angus Macgregor[1], Trevor Caruso[1]; [1]University of Florida

Several years ago, the theory of subjective inflation was presented as a possible explanation for why observers may think they perceive the visual periphery better than they actually do (Odegaard, Chang, Lau, & Cheung, 2018). Based on findings that observers have (1) liberal detection biases and (2) overconfidence for decisions in peripheral vision, the theory posited that observers have an inflated sense of perceptual capacities outside the fovea. Here we ask: how well do predictions made by this theory hold up when put to the test in wide-field visual displays? Using a monitor which ranged from -40 to +40 degrees, we presented observers with either Gabor patches or noise and asked them to judge the presence (or absence) of the stimulus, and rate confidence in their responses. Across two experiments, when comparing attended and unattended parts of the visual field, observers exhibited higher numbers of false alarms for unattended (compared to attended) peripheral locations, providing evidence of liberal detection biases in the periphery. Further, two experiments with feedback showed that these detection biases were resistant to different forms of feedback, including not only trial-by-trial feedback, but also explicitly telling observers about the liberal detection bias before they performed the task. However, in a final experiment which matched performance across peripheral locations, confidence was shown to decline with increasing eccentricity, challenging the idea that observers may overestimate their perceptual capacities in the periphery. In light of these new findings, we review the viability of subjective inflation, as some predictions appear to be confirmed (i.e., a liberal detection criterion in the unattended periphery), but others appear to be refuted (e.g., overconfidence in the periphery). We conclude by discussing the importance of using different tasks and subjective measures to further probe perceptual decisions in the visual periphery.

This work was supported by an Office of Naval Research Young Investigator Award (#N00014-22-1-2534)

Talk 2, 5:30 pm, 25.12

ACCUMULATING EVIDENCE OVER SPACE AND TIME REQUIRES PRIMARY AND SECONDARY VISUAL CORTICES

Joao Couto[1], Lillian Wilkins[1], Anup Khanal[1], Anne K Churchland[1]; [1]University of California Los Angeles, Department of Neurobiology

Accumulation of evidence relies on neural circuits that use time flexibly and strategically. Genetic dissection tools for mice make them an appealing model to understand such circuits, but encouraging mice to accumulate evidence over time and space has proved challenging. Further, not all visual tasks in mice recruit cortical structures (and instead rely on subcortical pathways), making it difficult to draw parallels between mammalian species. Here, we ask if the visual cortex is required for spatio-temporal integration of cues in a novel decision-making task for head-restrained mice. Our task uses localized visual patches (25-30˚), drawn stochastically at different azimuths/elevations, inviting accumulation of evidence from distinct locations in the visual field. After a delay period (0.5-0.7s), a correct choice is reported by licking to the spout corresponding to the side with the highest patch rate. We performed optogenetic suppression of primary visual cortex (V1) or the anteromedial visual area (AM) using an inhibitory opsin (stGtACR2) selectively expressed in excitatory neurons. Unilaterally suppressing V1 during the stimulus period (1 second) strongly biased responses to the ipsilateral side, consistent with a role for V1 in spatio-temporal integration (2 mice/13 sessions). When the manipulation window was restricted to the initial or late stimulus periods, we observed a similar impairment (2 mice/8 sessions), albeit smaller in magnitude. This result, together with psychophysical regression, suggests that mice utilize the majority of the stimulus period. To test whether V1 activity was required after the stimulus period, we suppressed activity in the delay period, and no performance impairment was seen. Similar to V1, suppressing AM during the stimulus period evoked a strong bias towards ipsilateral choices. These results demonstrate that mice can accumulate visual evidence and that this computation relies on visual cortical structures.

This work was supported by NIH (award R01EY022979).

TALK 3, 5:45 PM, 25.13

OVERCONFIDENCE FOR FILLED-IN INFORMATION IN THE FOVEAL ROD SCOTOMA IN MESOPIC VISION

Hui Men[1] (hui.men@uni-marburg.de), Alexander C. Schütz[1]; [1]University of Marburg

Different types of retinal photoreceptors contribute to vision, depending on the level of illumination: cones function in photopic vision in bright light, while rods function in scotopic vision in dim light. In intermediate mesopic vision, rods and cones are active simultaneously. Due to the lack of rods in the fovea, a foveal scotoma occurs under scotopic vision. It has been shown that this scotoma can be filled in with surrounding information, and that humans tend to trust such inferred central information more than veridical peripheral information. However, it remains unclear whether these effects also occur in mesopic vision where cones are active in the fovea. We investigated filling-in of the rod scotoma and the confidence preference in photopic, scotopic, and mesopic vision. In the mesopic condition, rods and cones were independently stimulated using a tetrachromatic projector. Our stimuli comprised two concentric circles of sine wave gratings, with the center smaller than the size of rod scotoma to allow foveal filling-in by surrounding information. The orientation of the center and the surround was parallel (continuous stimulus) or orthogonal (discontinuous stimulus). Participants had to determine the continuity of two successively presented stimuli and to choose for which stimulus they felt more confident. For rod-isolating stimuli in the fovea, participants were unable to discriminate continuous and discontinuous stimuli and were biased to perceive stimuli as continuous, both under scotopic and mesopic conditions. This suggests that the absence of rod information in the fovea was filled-in by the immediate surround, even when cones could supply information at the fovea in mesopic vision. Moreover, participants tended to trust the centrally inferred information more than veridical information in the periphery, both under scotopic and mesopic conditions. This suggests that the preference for inferred information in the foveal rod scotoma persists even under mesopic conditions.

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).

TALK 4, 6:00 PM, 25.14

PROBING METACOGNITION WITH THE "PSYCHOPHYSICIST'S MICROELECTRODE"

Tarryn Balsdon[1], Paolo Bartolomeo[2], Vincent de Gardelle[3], Pascal Mamassian[1], Marion Rouault[2]; [1]Ecole Normale Superieure and CNRS, [2]Institut du Cerveau and INSERM, [3]Paris School of Economics and CNRS

Self-evaluations of perceptual decision accuracy are thought to rely on a general metacognitive mechanism that computes confidence as a common currency across tasks. As a subjective variable, it is a challenge to directly manipulate confidence. Rather, the study of confidence has relied on correlational approaches, indirectly manipulating confidence by varying the available decision evidence. Here, we tested whether we can directly manipulate confidence by appropriating the classical psychophysical method of sensory adaptation to confidence adaptation. We tested 100 observers' metacognitive sensitivity in a forced-choice confidence paradigm: Participants were asked to choose which of two orientation discrimination decisions was more likely to be correct. This followed a long sequence of numerosity and colour discrimination decisions, in which the task difficulty was manipulated to either be varied (baseline), relatively difficult (low-confidence adaptation), or relatively easy (high-confidence adaptation). The presented stimuli, orientation discrimination sensitivity, and thus the underlying decision evidence relevant for evaluating confidence, was unchanged following each of these prolonged exposure sequences of numerosity/colour discrimination decisions. Yet, we found a systematic change in metacognitive sensitivity following prolonged exposure to high/low confidence decisions compared to baseline, consistent with cross-task confidence adaptation. The pattern of this change in metacognitive sensitivity disambiguates models of the neural coding scheme of confidence, suggesting a dual-channel coding scheme involving tuning to both increasing confidence and increasing uncertainty. Greater channel responsiveness and more channel overlap predicted better metacognitive sensitivity. This cross-task adaptation provides direct evidence for general metacognitive computations in human observers, and demonstrates adaptive resource allocation for metacognitive processes.

This work was supported by the Fyssen Foundation

TALK 5, 6:15 PM, 25.15

CONFIDENCE ACCURATELY TRACKS INCREASING INTERNAL NOISE IN PERIPHERAL VISION

Luhe Li[1], Michael S. Landy[1,2]; [1]Department of Psychology, New York University, [2]Center for Neural Science, New York University

Visual sensitivity declines with increasing eccentricity, yet it remains unclear whether we are aware of this. Previous studies report both overconfidence and underconfidence in peripheral vision, with mixed results complicated by conflicting definitions of confidence. Here, we used the most common definition of confidence, the perceived probability that a decision is accurate. We tested the Bayesian-confidence hypothesis that confidence is derived from the posterior probability distribution of the feature given noisy sensory measurements. The first-order task was visual localization and confidence was measured using post-decision wagering. Participants fixated on a central crosshair monitored by an eye tracker. A 33 ms Gaussian blob appeared at a random horizontal location (uniformly from –31.5 to 31.5°). Participants adjusted a cursor to localize the target and reported confidence by setting a symmetrical range around the localization response. They earned higher points for narrower confidence ranges but received zero points if the range did not encompass the target. Points decreased linearly with range length, which incentivized both accurate localization and confidence reports. Feedback was only given during practice to ensure participants learned the cost function, with no trial-wise feedback in the main experiment. Both localization variability and confidence range increased with eccentricity. We estimated the sensory noise from the localization responses assuming that sensory noise scales with eccentricity. We then compared two confidence models. The ideal observer uses accurate estimates of sensory noise and combines the posterior distributions of location with a subjective cost function to maximize expected gain. The heuristic observer determines the confidence range proportionally to the distance of location estimates from fixation. Model comparison shows that confidence reports were better fit by the ideal-observer model. Humans can accurately monitor the increasing internal noise in peripheral vision and use it to make optimal confidence judgments.

NIH EY08266

Talk 6, 6:30 pm, 25.16

THE INFLUENCE OF RESPONSE BIAS ON CONFIDENCE AND ACCURACY IN MULTI-ALTERNATIVE TASKS FOR HUMANS AND ARTIFICIAL NEURAL NETWORKS

Bogeng Song[1], Dobromir Rahnev[1]; [1]Georgia Institute of Technology

How does response bias influence confidence judgments in perceptual decision making? Most previous research has used 2-choice tasks, which makes it difficult to address this question. Indeed, in 2-choice tasks, bias for one category is equivalent to a bias against the other category, meaning that one cannot independently assess the bias for each category. Therefore, to determine how response bias influences confidence judgments, here we used two datasets with multi-alternative tasks. In the first dataset, 60 subjects performed an 8-choice digit categorization task using noisy MNIST images. In the second dataset, 37 subjects performed a 16-choice image classification task using blurred ImageNet images. We quantified the bias for each stimulus category as the rate at which a category was chosen when the stimulus was selected from all other categories. Unlike for 2-choice tasks, this procedure allowed us to estimate the bias for each category independently from the remaining categories. We then used the response bias for each category to predict the confidence and accuracy for that category. We found that a bias towards a given category predicted higher accuracy. Counterintuitively, however, being biased towards a category predicted lower confidence ratings for that category. These results were robust to variations in task difficulty and condition. We further observed that artificial neural networks (ANNs) trained on object recognition replicated the positive effect of response bias on accuracy, but, in contrast to humans, also showed a positive effect of response bias on confidence. These findings suggest that when rating confidence, humans – but not ANNs – actively discount evidence for stimulus categories that they are biased towards. More broadly, our results demonstrate the utility of complex, multi-alternative tasks in discovering the mechanisms driving visual confidence.

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

Talk 7, 6:45 pm, 25.17

RESPONSE DURATION: A UBIQUITOUS IMPLICIT MEASURE OF CONFIDENCE

Hanbei Zhou[1] (hzhou43@jhu.edu), Rui Zhe Goh[2], Ian Phillips[3], Chaz Firestone[4]; [1]Johns Hopkins University

Among the most reliable connections between internal mental processing and external behavior is *response time*, with easier, more accurate, and more confident judgments typically made faster. But which aspects of response time are relevant? Whereas psychophysical studies traditionally focus on the time taken to initiate a response, an underexplored measure is the duration of the response itself—not just the amount of time between stimulus onset and keypress (reaction time), but also how long one holds down the key before releasing it (response duration). Response duration is a ubiquitous and freely available data source, yet almost no studies report or analyze it (Pfister et al., 2023). Here, 3 varied experiments demonstrate that response duration reliably predicts subjective confidence, independent of reaction time. In Experiment 1, subjects detected faces within white noise, with difficulty manipulated by varying face opacity. Subjects responded with a keypress (with both keyUp and keyDown events recorded separately), followed by a confidence judgment. Remarkably, subjects held down the response key longer during trials in which they subsequently reported lower confidence, as if making these face-detection judgments in a tentative fashion. The same pattern held in another visual task (judging the coherence of random-dot motion; Experiment 2), and a cognitive task (classifying American cities as geographically Eastern or Western; Experiment 3). In all cases, response duration accounted for variance in confidence that was not predicted by reaction time. Response duration has distinct advantages as a measure of confidence: It taps confidence at the time of judgment (rather than retrospectively), it can be used when traditional confidence judgments are difficult to elicit (e.g., in animals or infants), and it may be less affected by biases associated with explicit reports. Our results suggest that response duration is a valuable and untapped source of information, raising many avenues for future investigation.

Talk Session: Saturday, May 17, 2025, 5:15 – 7:00 pm, Talk Room 2

Visual Search

Moderator: Freek van Ede, Vrije Universiteit Amsterdam

Talk 1, 5:15 pm, 25.21

PREDICTIVE AND REACTIVE DISTRACTOR SUPPRESSION RELIES ON INTEGRATED ATTENTIONAL MECHANISMS

Oscar Ferrante[1], Ole Jensen[2,3], Clayton Hickey[1]; [1]Centre for Human Brain Health, University of Birmingham (UK), [2]Oxford Centre for Human Brain Activity, University of Oxford (UK), [3]Department of Experimental Psychology, University of Oxford (UK)

Visual attention is significantly influenced by statistical regularities in the environment, with spatially predictable distractors being proactively suppressed. The neural mechanisms underlying this proactive suppression remain poorly understood. In this study, we employed magnetoencephalography (MEG) and multivariate decoding analysis to investigate how predicted distractor locations are proactively represented in the human brain. Participants engaged in a additional-singleton visual search task, identifying a target stimulus while ignoring a colour-singleton distractor when present. Crucially, the distractor appeared more frequently on one side of the visual field, creating a statistical learning spatial prediction. Our results revealed that distractor locations were encoded in temporo-parietal brain regions prior to stimulus presentation, supporting the hypothesis that proactive suppression guides visual attention away from predictable distractors. The neural activity patterns corresponding to this pre-stimulus distractor suppression extended to post-stimulus activity during late attentional stages (~200 ms), indicating an integrated suppressive mechanism. This generalization from pre-stimulus to post-stimulus was absent in the early sensory stages (~100 ms), suggesting that post-stimulus suppression is not merely a continuation of sustained proactive suppression. Instead, the same suppressive mechanism is activated at two distinct stages. These findings establish a mechanistic link between proactive and reactive suppression of predictable distractors, elucidating their shared and unique contributions to attentional processes.

TALK 2, 5:30 PM, 25.22

BALANCING EXPLORATION AND EXPLOITATION IN VISUAL SEARCH: INSIGHTS FROM BEHAVIORAL AND COMPUTATIONAL MODELS

Haokui Xu[1,2] (haokuixu.psy@gmail.com), Xutao Zheng[2], Jianzhe Xu[2], Jingjing Hu[3], Jifan Zhou[2], Mowei Zhen[2]; [1]Zhejiang University of Technology, [2]Zhejiang Universisty, [3]Zhejiang International Studies University

Goal-directed attention in visual search tasks has been extensively studied, but the role of exploration—the complementary aspect of attention—remains less understood. In complex and unfamiliar scenarios encountered in daily life, individuals often gather information to simplify the situation and reduce the search space before locating a target. Our study investigated this process through three experiments involving a digit search task, where participants searched for a target digit within sequences arranged either regularly or randomly. Behavioral results indicated that search times were shorter for regular sequences compared to random sequences and increased logarithmically with set size. Eye-movement analysis revealed a similar pattern, as the number of fixations aligned with search times, suggesting that participants leveraged regularities to narrow the search space efficiently, thereby locating the target with fewer fixations. To further investigate attention selection strategies, we developed computational models. The optimal model demonstrated that attention flexibly balances two strategies based on scenario complexity. When the search space is large, attention prioritizes locations that were most helpful in narrowing the search space (with the highest expected information gain), even if these locations are less likely to contain the target. Conversely, when the search space is reduced, attention focuses on the most probable target locations. This dynamic selection mechanism highlights how the visual system balances exploration and exploitation to achieve optimal outcomes, enabling humans to navigate complex environments despite limited cognitive resources.

Zhejiang Provincial Philosophy and Social Sciences Planning Project (23NDJC037Z)

TALK 3, 5:45 PM, 25.23

EFFICIENT HEURISTIC DECISION PROCESSES IN OVERT VISUAL SEARCH

Anqi Zhang[1,2] (anqizhang@utexas.edu), Wilson S. Geisler[1]; [1]The University of Texas at Austin, [2]University of California, Santa Barbara

Simple heuristic decision rules that perform close to the Bayes optimal rule are likely candidates for decision processes in biological systems because they are the ones most likely to be found by natural selection and learning over a life span. In covert search, the Bayes-optimal decision rule takes into account the prior probability at each potential target location, weighs the response at that location by the local detectability (d'), and then picks the location with the maximum posterior value. Recently, we showed that in covert search a wide range of simple decision heuristics closely approach optimal accuracy, even though these heuristics largely ignore the actual variation in prior probability and detectability across the visual field. For example, even for targets where d' falls rapidly with retinal eccentricity, assuming a constant d' over the search area has a negligible effect on overall search accuracy. We extended this analysis to overt search, where the Bayes-optimal searcher uses each target's specific d' map for both updating the posterior probability map and selecting fixations. We found that (1) Changes in heuristic parameters and stopping criteria cause substantial tradeoffs between the overall search accuracy and mean number of fixations. (2) Many heuristics with a fixed, foveated d' map are highly efficient, but few heuristics with a constant-valued d' map are highly efficient. Specifically, we defined the efficiency of a heuristic rule as the ratio of the overall search accuracy of that heuristic searcher to that of the ideal searcher, when the heuristic searcher is required to make the same number of fixations for each stimulus as the ideal searcher. Overall, our findings uncover several biologically plausible and testable near-optimal heuristics for overt visual search.

Supported by NIH grants EY11747 and EY024662.

TALK 4, 6:00 PM, 25.24

PREDICTING HUMAN FORAGING BEHAVIOUR IN 3D: A COMPUTATIONAL APPROACH

Manjiri Bhat[1] (mb23365@essex.ac.uk), Russell A Cohen Hoffing[2], Anna Hughes[1], Alasdair Clarke[1]; [1]University of Essex, [2]DEVCOM Army Research Laboratory

Spatial exploration is a key cognitive ability for humans and other animals, allowing them to find food and other resources. Visual foraging paradigms, where participants search for multiple items in a two-dimensionally bound environment, allow us to explore the strategies used in spatial exploration. In many previous studies, the key outcome measurement has been differences in item selection patterns based on run behaviour. In our paradigm, we adopt a novel approach, where participants (in an avatar, first person view) forage items by moving in a semi-naturalistic three-dimensional environment that accounts for depth, body orientation, time (pauses), and rotation. We also test participants under different conditions to investigate how factors such as stamina (an "energy bar" that reduces with movement at a rate based on low, medium or high stamina), item scarcity (high or low item availability), and knowledge of item locations and availability (birds-eye-view map or no-map) influences exploration behaviour and foraging strategy. With an aim to better characterise underlying cognitive processes of exploration behaviour and foraging strategy in this unconstrained foraging task, we have developed and fit a generative model which formally predicts sequences of item selections using latent parameters such as proximity (selecting items closest to the previous selection) and momentum (selecting items in a forward motion or doubling-back). Our results suggest that proximity is highly predictive of behaviour, even when stamina and item scarcity change. However, proximity is less predictive when items are scarce. In addition, momentum (i.e., forward motion) is a greater predictor when participants do not have a map or have a map with high stamina. We conclude that our model is able to accurately predict behaviour on a target-by-target basis in an unconstrained foraging paradigm and discuss future directions for improving the model by incorporating pauses and rotation.

Economic and Social Research Council (ESRC), US DEVCOM Army Research Laboratory

Talk 5, 6:15 pm, 25.25

COMPUTER CURSOR TRAJECTORIES ARE PREDICTIVE OF UPCOMING SUCCESS IN VISUAL SEARCH

Audrey Siqi-Liu[1] (audrey.liu@gwu.edu), Sarah Malykke[1], Kelvin Oie[2], Dwight Kravitz[1], Stephen Mitroff[1]; [1]The George Washington University, [2]DEVCOM Army Research Laboratory

Performance on visual search tasks—finding targets amongst distractors—is typically assessed with aggregate behavioral measures (e.g., response times, accuracy). Such measures are effective in their simplicity, but can obscure subtle dynamics in the visual decision-making processes that underlie search through a complex visual array. The current project explored the utility of cursor tracking data to reveal nuanced patterns of search behavior. Using a web-based implementation of a standard 'Ts among Ls' task and crowdsourced data collection, we recorded a large number of participant's cursor behaviors during visual search (either through the movements of a computer mouse or trackpad). From the recorded cursor traces, we calculated time-resolved features such as trajectory length, speed, and dwell times on stimuli and non-stimuli areas, characterizing how these behaviors change over the course of search within each trial. Machine-learning models trained on these cursor features accurately predicted errors before they occurred (i.e., up to ~500-1000 ms prior to response). Additional analyses delineated which cursor features were particularly important to error prediction, with the rank order of importance changing as the trial progressed. These and other analyses (e.g., a detailed description of how each time-resolved feature varied in hit versus miss trials) will be discussed. Combining cursor tracking data with advanced computational methods, supported by crowdsourced data collection, we provide a richer narrative of behavior during visual search. This novel approach opens new research possibilities by demonstrating the resolution of behavioral data that can be collected via online platforms and how such fineness of measurement can provide more insight into the basic mechanisms of search and real-time error prediction. Practical applications range from understanding search strategy differences between individuals to characterizing optimal windows for computer-assisted detection in artificial intelligence applications.

W911NF-23-2-0210, W911NF-23-2-0097, W911NF-24-2-0188

Talk 6, 6:30 pm, 25.26

LOOKING INTO WORKING MEMORY TO VERIFY POTENTIAL TARGETS DURING VISUAL SEARCH

Sisi Wang[1], Freek van Ede[1]; [1]Vrije Universiteit Amsterdam

Finding what you are looking for is a ubiquitous task in everyday life that relies on a two-way comparison between what is currently viewed and internal search goals held in memory. Yet, despite a wealth of studies tracking visual verification behavior among the external contents of perception, complementary processes associated with visual verification among internal contents of memory remain elusive. Building on a recently established gaze marker of internal visual focusing in working memory, we tracked the internal-inspection process associated with confirming or dismissing potential targets during search. We show how we look back into memory when faced with external stimuli that are perceived as potential targets and link such internal inspection to the time required for visual verification. A direct comparison between visual verification among the contents of working memory or perception further revealed how verification in both domains engages frontal theta activity in scalp EEG, but also how mnemonic verification is slower to deploy than perceptual verification. This establishes internal verification behavior as an integral component of visual search, and provides new ways to look into this underexplored component of human search behavior.

This work was supported by an NWO Vidi Grant from the Dutch Research Council (14721) and an ERC Starting Grant from the European Research Council (MEMTICIPATION, 850636) to F.v.E. We thank Baiwei Liu and Anna van Harmelen for their useful discussions.

Talk 7, 6:45 pm, 25.27

INDIVIDUAL WORKING MEMORY CAPACITY PREDICTS SEARCH PERFORMANCE IN MULTIPLE COLOUR SEARCH

Anna Grubert[1] (anna.k.grubert@durham.ac.uk), Ziyi Wang[1], Courtney Turner[1]; [1]Durham University

Visual search is guided by attentional templates, i.e., target representations that are assumed to be held in visual working memory (vWM). vWM capacity varies between individuals but is typically limited to 3-4 items. Similar capacity limitations have also been observed in visual search, but it is yet unclear whether they are directly linked to the individually limited resources of vWM. We tested this by correlating behavioural and electrophysiological markers of individual vWM capacity (Cowan's k, CDA amplitudes) with behavioural and electrophysiological markers of search efficiency (accuracy rates, N2pc amplitudes) measured in change detection and visual search tasks, respectively. In each trial of the change detection task, participants were presented with a memory display containing one, two, or three coloured squares. After a retention period, they were shown a test display and had to decide whether it was identical to the memory display or contained a colour change. In the visual search task, participants searched for a target-colour bar amongst five differently coloured nontargets and had to indicate whether it had a horizontal or vertical orientation. One, two, or three possible target colours were cued at the beginning of each block. Results revealed significant load effects both in the change detection and search tasks with larger k values and CDA amplitudes and lower accuracy rates and N2pc amplitudes in higher- as compared to lower-load trials, respectively. More importantly, vWM indices predicted search performance at the individual level: Individuals with higher k values and larger CDA amplitudes produced greater search accuracy, larger N2pc amplitudes, and smaller load costs both in terms of search accuracy and N2pc amplitudes. These results suggests that individual search performance in multiple colour search directly depends on individual vWM limitations.

This work was funded by a research grant of the Leverhulme Trust (RPG-2020-319) awarded to AG.

Talk Session: Sunday, May 18, 2025, 8:15 – 9:45 am, Talk Room 1

Spatial Vision: Crowding and eccentricity, clinical, models

Moderator: Michael Herzog, EPFL

Talk 1, 8:15 am, 31.11

MECHANISMS OF FOVEAL CROWDING: INSIGHTS FROM RETINAL IMAGING AND RETINAL-CONTINGENT PSYCHOPHYSICS

Krishnamachari 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, Berkley, [3]University of Bonn

Visual crowding, the interference in recognition of a stimulus caused by surrounding similar objects, occurs not only in the periphery but also at the center of gaze, where visual resolution is highest. Yet, the mechanisms underlying foveal crowding are still unclear, due to the difficulty in controlling for optical factors and fixational eye motion. To address these challenges, we used an Adaptive Optics Scanning Light Ophthalmoscope (AOSLO) to stimulate the retina while reducing the impact of optical aberrations and maintaining the stimulus at a fixed retinal location despite the presence of fixational eye movements. Subjects (N=8) performed a 4AFC digit identification task using Pelli's font, designed specifically to probe foveal crowding. Stimulus sizes were set at three times the acuity threshold. Flanker distances varied using the method of constant stimuli. Stimuli were presented under retinal stabilization at the Preferred Retinal Locus (PRL) and 15 arcmin away. By assessing cone spacing at the stimulated location, our results show that the spatial extent of crowding at the PRL approximates a single cone diameter (1.15 ± 0.48 times the cone diameter). However, just 15 arcmin away the extent of crowding exceeds cone spacing (3.08 ± 2.08 times the cone diameter), revealing that additional cortical pooling mechanisms dominate with increasing eccentricity even within the foveola. Further, at the PRL, incorrect responses showed no bias in reporting either flanker, but 15 arcmin away, subjects were more likely to report the inner flanker on incorrect trials. These findings reveal that photoreceptors spacing predicts the spatial extent of crowding only at the PRL. They indicate that further signal pooling takes place just a few arcmin away from this location suggesting the presence of a cortical magnification gradient within the fovea.

R01 EY029788, R01 EY018363, EY001319, R01 EY023591, Ha5323/6-1 and Ha5323/8-1

Talk 2, 8:30 am, 31.12

PERCEPTOGRAMS MEASURED FOR AMBLYOPIC FORM DISTORTIONS, MODELED WITH CORTICAL DEFICITS

Akihito Maruya[1] (user3098@sunyopt.edu), Farzaneh Olianezhad[1], Jingyun Wang[1], Jose-Manuel Alonso[1], Qasim Zaidi[1]; [1]State University of New York, College of Optometry

Striking form distortions seen by amblyopes have been documented by showing high contrast sinusoidal gratings to the amblyopic eye (AE) and drawing the percepts through the fellow eye (FE). The distortions fall into 7 classes, each resembling sums-of-grating-pairs (Barrett et al 2003). These distortions of orientation processing are keys to understanding neural deficits in amblyopic cortex. Drawings may not accurately represent distortions of contrast, shading and frequency, so we measured perceptograms for 4 amblyopes ages 22-45. In a dichoptic display, AE was shown a 3° test grating in the center (6, 9, 12 cyc/deg, 4 orientations, ON or OFF). FE was shown 8 surrounding patterns: the test grating and the 7 types of distortion plaids. The observer picked the pattern most like the central image. Then the AE was shown the test grating and the FE the chosen match while contrast, frequency, orientation and phase of the 2 constituent gratings were adjusted until the two percepts matched perfectly (3 repeats). In the model, images were convolved with the optical point-spread-function, passed through the ON/OFF nonlinearity, and convolved with center-surround retino-thalamic filters with multiplicative internal noise. Then a formal equation that the signals generated in visual cortex by the test grating seen through AE match the signals generated by the matched perceptogram seen through FE, was used to analytically derive the modified cortical filters processing AE signals for 24 perceptograms jointly for each observer, using standard orientation tuned filters for normal cortex. The modified filters recreated the distortions seen through AE and were appreciably broader in tuning, slightly shifted in orientation preference, and varied in response magnitude. The orientation-based changes in neuronal tuning-width and response-magnitude provide a target for neural development models of amblyopia and could predict distortions in other orientation-based percepts such as contours, 3D shape-from-texture, mirror-symmetry and object poses.

NEI grants EY035085 & EY035838

Talk 3, 8:45 AM, 31.13

LOCALIZATION BIASES IN THE PERIPHERY ARE IDIOSYNCRATIC: EVIDENCE FROM OVER 9000 OBSERVERS

Anna Kosovicheva[1] (a.kosovicheva@utoronto.ca), Ido Ziv Li[1], Jihahm Yoo[2], Jeremy M. Wolfe[3,4], Jiali Song[1]; [1]University of Toronto Mississauga, [2]Korea Science Academy of KAIST, [3]Brigham & Women's Hospital, [4]Harvard Medical School

Accurately registering the location of an object is a fundamental visual process. Previous studies have emphasized commonalities across individuals in the effect of the polar angle of the target relative to fixation in visual localization. However, there is also evidence that individuals exhibit consistent, idiosyncratic patterns of directional, angular error when reporting target locations in the periphery, and such patterns of error are weakly correlated between observers. This evidence comes from small-scale laboratory studies, involving dozens of participants, which may be underpowered to detect subtle consistencies across the population. We examined the consistency of individuals' localization errors using a large-scale dataset from an online game (over 9,400 observers and 4.5 million trials across 639,000 sessions). On each trial, participants were instructed to identify a symbol in the center of the screen and were simultaneously shown a peripheral target. The target could appear with 0-10 distractor items. The eccentricities of the peripheral targets and distractors varied randomly and independently across trials. Participants clicked on the location of the peripheral target, and if correct, were asked to identify the central symbol among 5 alternatives. We analyzed trials where participants correctly identified the symbol and localized the target. We divided trials into bins based on their polar angle. We then calculated pairwise correlations in the angular (directional) click error between participants relative to display center (clockwise vs. counterclockwise). We found that directional localization errors were, on average, uncorrelated between all possible pairs of participants. Between-subject correlations of localization errors were normally distributed and centered around 0. However, within individuals, errors were non-random. Split-half correlation yielded a reliable, positive correlation ($r = 0.41$, $t(9411)= 185.82$, $p< 0.001$). These results align with the findings of small-scale laboratory studies and suggest that consistent idiosyncratic localization errors in the visual periphery are uncorrelated at the population level.

This work was supported by an NSERC Discovery Grant to AK, and NIH EY017001 and NSF 2146617 to JMW.

Talk 4, 9:00 AM, 31.14

WHERE INTERNAL NOISE AND EFFICIENCY UNDERLIE VISUAL FIELD ASYMMETRIES

Shutian Xue[1] (shutian.xue@nyu.edu), Antoine Barbot[1], Rachel Chen[1], Marisa Carrasco[1]; [1]New York University

[Background] Visual performance typically peaks at the fovea and declines with eccentricity, and exhibits polar angle asymmetries: better performance along the horizontal than the vertical meridian and better performance at the lower than the upper vertical meridian. Do these performance differences reflect differential ability to process task-relevant information from noise? Here, we investigate how two factors limiting performance–internal noise (amount of internal variability in the system) and efficiency (the ability to extract information from the target)–underlie performance differences throughout the visual field. [Method] At each location, observers discriminated the orientation (±45º off the vertical axis) of Gabor patches (4, 5 and 6 cpd) embedded in varying contrast levels of dynamic white noise. Using an equivalent noise protocol, we mapped contrast threshold as a function of noise contrast to estimate additive and multiplicative internal noise, as well as efficiency, at three eccentricities (fovea, parafovea: 4°, perifovea: 8°) and four polar angles (left and right horizontal meridian, upper and lower vertical meridian). [Results] (1) Additive (but not multiplicative) internal noise increased with eccentricity. Efficiency was higher at the fovea and 4º than 8º eccentricity, higher at the horizontal than vertical meridian at 8º eccentricity, and at the lower- than upper- vertical meridian at both 4º and 8º eccentricities. (2) Multiplicative (but not additive) internal noise was similar at 4 and 5 cpd and lowest at 6 cpd. Efficiency was higher at 4 cpd than at 5 and 6 cpd. These differences were similar across locations. [Conclusion] Distinct computations limit performance throughout the visual field: Additive internal noise primarily underlies eccentricity differences, consistent with variations in cortical surface area. Efficiency primarily underlies polar angle asymmetries, particularly in the perifovea, reflecting variation in neural tuning properties.

Funding: NIH R01-EY027401 to M.C.; NIH training grant to NYU, NIH1F31EY036732-01 to S.X.

Talk 5, 9:15 AM, 31.15

THE NEURAL SIGNATURES OF REDUNDANCY MASKING INVESTIGATED BY EEG FREQUENCY TAGGING

Nihan Alp[1], Dogukan Nami Oztas[1], Li L-Miao[2], Bilge Sayim[2]; [1]Sabanci University

Redundancy masking (RM) –the reduction of the number of perceived items in repeating patterns– occurs with as few as three items. For example, when three identical, closely spaced lines are presented in the periphery, individuals often perceive only two lines. The underlying neural mechanisms of these substantial detection-like errors remain elusive. Here, we use steady-state visual evoked potential to examine the neural correlates of RM. Three arcs (quarter-circles; 0.44° line width) were presented for 10s either on the right or left side of fixation at 17.3°, 19.5° and 21.7° eccentricity. Each arc was tagged with a different frequency. Participants were instructed to maintain fixation, and the stimulus disappeared if participants' gaze shifted using a gaze-contingent paradigm. After the stimulus offset, participants reported the number of arcs they mostly perceived during a trial. We quantified baseline-corrected amplitudes for each arc (tagged frequency and harmonics) and calculated signal-to-noise ratios (SNRs) for intermodulations, separating both by the behavioral responses (RM: 1 or 2, non-RM: 3). The middle arc elicited a weaker amplitude compared to the inner one, with no significant differences between the middle and outer or the inner and outer arcs. Across intermodulations, SNRs for response '2' were significantly higher than '3', indicating greater neural integration when perceiving 2 arcs (RM) than 3 (non-RM). Specifically, the integration of inner and middle arcs, as well as middle and outer arcs, indicated by corresponding intermodulations, was significantly stronger when participants perceived 2 arcs compared to 3, further supporting stronger integration when RM occurred. These results indicate that while RM involves a loss of access to visual information, the lost signals may nevertheless be integrated by the visual system across space and time. We suggest that the effects of redundancy-masked items –although unavailable for conscious report– are still observed in the neural signatures of RM.

ANR-19-FRAL-0004; TUBITAK 122N748

Talk 6, 9:30 am, 31.16

PROBABILITY SUMMATION MODEL OF SPATIAL POOLING WITHIN THE CARDINAL AXES OF DKL COLOR SPACE

Christopher S Wu[1], Daniel R Coates[1]; [1]University of Houston

The critical area of spatial summation is thought to reflect the spatial extent of some pooling mechanism in the visual system, although, the exact underlying processes are unknown. Previously, we have characterized the spatial summation of chromatic contrasts defined by the cardinal axes of DKL color space. Here, we describe a simple model to bridge physiological properties of the visual system and functional findings. Spatial summation functions were obtained for contrasts along the cardinal directions of DKL color space in four subjects. Functions were measured in the oblique meridians at 5, 10, 15, and 20 degrees eccentricity with stimuli consisting of solid circular spots of multiple sizes. Contrast detection thresholds to images of these stimuli were simulated through the model. The model simulates the pooling of information through the visual system using a cascade of convolutional filtering and downsampling layers. Filtering kernels representing receptive fields were generated based on eccentricity dependent sizes and the spatial overlap of specific retinal ganglion cell type dendritic fields. Relative detection thresholds were calculated through probability summation. The model was able to reproduce the bilinear shape characteristic of human spatial summation functions, matching our psychophysical dataset across eccentricities and within theorized visual processing pathways (average r2 across all conditions = 0.71). The resultant critical areas matched the size of simulated receptive fields, while the slope of partial summation could be controlled by the exponent of probability summation. The average exponent of summation was 4.0 in the achromatic condition, 1.6 in the +(L-M), 1.4 in the -(L-M), 3.1 in the +S-(L+M), and 2.7 in the -S-(L+M) conditions. This fitted exponent likely reflects the processing characteristics of pathway-dependent higher level cortical mechanisms. Thus, a combination of probability summation and physiologically inspired filtering can account for the distinct psychophysical spatial summation functions obtained across eccentricities and pathways.

Talk Session: Sunday, May 18, 2025, 8:15 – 9:45 am, Talk Room 2

Development

Moderator: Tessa Dekker, UCL

Talk 1, 8:15 am, 31.21

OBJECTIVELY MEASURING SIGHT RESCUE IN SEVERELY VISION-IMPAIRED YOUNG CHILDREN FOLLOWING GENE THERAPY

Marc Pabst[1,2,3], Kim Stäubli[1,2,3], Yannik Laich[3], Roni Maimon-Mor[1,2,3], Steven Scholte[4], Peter Jones[5], Michel Michaelides[1,3], James Bainbridge[1,3], Tessa Dekker[1,2,3]; [1]Institute of Ophthalmology, University College London, [2]Experimental Psychology, University College London, [3]Moorfields Eye Hospital NHS Foundation Trust, [4]Faculty of Social and Behavioural Sciences, University of Amsterdam, [5]Department of Optometry and Visual Sciences, City St George's, University of London

Recent breakthroughs in ocular gene therapy hold significant promise for treating inherited retinal diseases (IRDs), the most prevalent cause of blindness in children and young people. IRDs compromise the retina's structure and function, with severe forms leading to complete loss of light sensitivity in early childhood. However, significant challenges remain in objectively characterising the effects of new therapies, particularly for very young children. Vision-impaired toddlers and children typically struggle to keep their gaze focused on visual stimuli or provide consistent responses, and traditional assessments typically are heavily reliant on subjective evaluations by highly trained clinical specialists. To develop objective measures of therapeutic effects complementing existing assessment, we used child-friendly neuroimaging and gamified testing approaches. Using steady-state visual evoked potentials (ssVEP) measured with EEG, we non-invasively recorded cortical responses to flickering sinusoidal gratings with varying spatial frequencies. To ensure engagement we integrated gratings within personalised age-appropriate videos. Additionally, we used a novel reaching-behaviour test embedded in a child-friendly iPad game that involved searching and tapping moving Gabor patches of varying spatial frequencies. These protocols were applied to eight young children (ages 3–6) diagnosed with a severe form of Leber Congenital Amaurosis (LCA) who had received novel gene therapy.

Visual function was assessed either by comparing treated and untreated eyes or by conducting pre- and post-treatment evaluations, depending on the patient. Our tasks revealed substantially stronger visual cortex responses and better behavioural task performance for treated than untreated eyes. The improvements observed were remarkable, especially when considering the typical progression of the disease and the benefits of available ocular gene therapies. This is likely facilitated by the exceptionally early time of intervention, minimising retinal degeneration and maximising neural plasticity. Establishing sensitive, child-friendly, and objective measures for evaluating early treatment effects is critical for advancing the field of ocular therapy.

Funding was partially provided through the NIHR Moorfields Biomedical Research Centre.

TALK 2, 8:30 AM, 31.22

CORTICAL FEEDFORWARD-RECURRENT CIRCUIT ALIGNMENT MATURES FOLLOWING EXPERIENCE

Augusto Abel Lempel[1], Sigrid Trägenap[2], Clara Tepohl, Matthias Kaschube[2], David Fitzpatrick[1]; [1]Max Planck Florida Institute for Neuroscience, [2]Frankfurt Institute for Advanced Studies

Sensory cortical areas guide behavior by transforming stimulus-driven inputs into selective responses representing relevant features. A classic example is the representation of edge orientations in the visual cortex, which displays a functional connectivity alignment where layer 4 (L4) neurons co-activated by an orientation provide feedforward inputs to specific functional modules in layer 2/3 (L2/3) that share strong recurrent connections. Such aligned state of feedforward-recurrent interactions is critical for amplifying selective cortical responses, but how it develops remains unclear. Using simultaneous electrophysiology and calcium imaging, we probed the trial-to-trial correlation (coactivity) between single-unit spiking responses to oriented gratings in L4 and L2/3 with millimeter-scale modular responses in the L2/3 network before and after experience. We then compared each unit's orientation preference with that of coactive L2/3 modules. In experienced animals, units in both layers display coactivity with modules matching their preferred orientation. In naïve animals, despite high trial-to-trial response variability, L2/3 units are coactive with modules displaying similar orientation preference, consistent with a well-structured recurrent network. In contrast, L4 units are coactive with L2/3 modules corresponding poorly with their orientation preference, and this is consistent with a poor alignment between feedforward inputs from L4 elicited by oriented stimuli and activity patterns amplified by L2/3 recurrent interactions. One factor that could contribute such lack of functionally specific coactivity is high variability in naïve L4 neuron responses that decreases significantly following experience, but a computational model of feedforward-recurrent interactions suggests that high variability alone is insufficient to explain the naïve state. This model also provides a biological signature of misalignment that we have confirmed with in-vivo whole-cell recordings: dynamic changes in the orientation preference of L2/3 subthreshold responses. In conclusion, we provide diverse evidence for a realignment of feedforward-recurrent interactions following experience that is critical for building reliable and temporally consistent sensory representations.

National eye institute funding: K99EY034936-01A, 2R01EY011488

TALK 3, 8:45 AM, 31.23

THE DEVELOPMENTAL VISUAL INPUT ACROSS ENVIRONMENTS

Philip McAdams[1], Alexis Colwell[1], Linda B. Smith[1]; [1]Indiana University, [2]Indiana University, [3]Indiana University

Visual development is experience dependent; however, little is known about the low-level visual experience of infants. One might assume that at the scale of daily life, the statistics of visual experience are similar for all perceivers. However, the visual input depends on the perceivers' perspective, looking behavior, and their environment. For example, younger and older infants show different looking biases and have different egocentric perspectives and motor behaviors, and perceivers live in different environments with different visual properties. Using head-mounted cameras, we captured 100,000 egocentric images from the daily life experiences of infants aged 1-3, 6-8, and 10-12 months, from a small town in the USA (N=24) and from a dense urban fishing village in India (N=24). We extracted a range of spatial image statistics relating to early vision and complexity, e.g. edge density, edge orientations and their predictive relations, and fractal dimension, to characterize and compare infants' visual input across development and environment. Overall, the youngest infants' visual input, across locations, was characterized by greater simplicity and predictive edge patterns than older infants, with both locations showing similar developmental trends. However, complexity and predictive properties among edges differed across locations. For example, US infants' input had overall lower fractal complexity than Indian infants, unrelated to amount of time spent outdoors. By 10-12 months, fractal complexity decreased for Indian infants but increased for US infants. Our results suggest that young infants' visual systems are biased toward simplicity, and that as the visual and motor systems develop, infants can select their own egocentric perspectives to create a curriculum for learning. Our findings are consistent with developmental changes in looking biases found in laboratory studies, showing these same changing biases at the scale of everyday input. The cross-environmental differences suggest both universal and context dependent regularities in the visual input.

TALK 4, 9:00 AM, 31.24

VISUAL SHAPE PROCESSING FOR ACTION DOES NOT DEPEND ON VISUAL EXPERIENCE: EVIDENCE FROM LATE-SIGHTED CHILDREN

Shlomit Ben-Ami[1,2,3] (shlomit@mit.edu), Roy Mukamel[2], Sana Khan[7], Chetan Ralekar[1], Mrinalini Yadav[4], Pragya Shah[4], Suma Ganesh[5], Priti Gupta[4], Abhinav Gandhi[6], Pawan Sinha[1]; [1]MIT Department of Brain and Cognitive Sciences, MA, USA, [2]Sagol School of Neuroscience, School of Psychological Sciences, Tel-Aviv University, Israel, [3]Minducate science of learning research and innovation center, Tel-Aviv, Israel, [4]The Project Prakash Center, Delhi, India, [5]Department of Ophthalmology, Dr. Shroff's Charity Eye Hospital, Delhi, India, [6]Worcester Polytechnic Institute, MA, USA, [7]Wellesley College, MA, USA

Does our capacity to process the visual properties of an object and manipulate it accordingly depend on visual experience? Efficient object interaction requires visual estimation of to-be-grasped object properties such as size, weight, and shape to guide pre-programming of grip aperture, force, and positioning. While prior visual experience is widely assumed to be critical for such estimates, empirical evidence remains sparse. We investigated this assumption in visual shape processing among 14 late-sighted children who were born with bilateral dense cataracts and gained vision only after surgery in late-childhood (provided by Project Prakash, India). Patients were tested pre- and post-surgery, alongside 10 sighted controls assessed under normal vision and acuity-matched simulated visual loss. Participants performed a visually guided pincer grasping task (vision-for-action) and a delayed-match-to-sample visual discrimination task (vision-for-perception), as well as visual acuity testing. Grip efficiency, measured by object stability scores based on grasp positioning, was impaired in patients relative to normally sighted controls but comparable to acuity-matched controls as early as the first post-surgery assessment. Conversely, visual discrimination accuracy remained suboptimal even years after surgery, beyond limitations explained by visual acuity reduction. These results reveal a developmental dissociation of visual shape processing for action and perception. Action-oriented visual shape processing depends on acuity but not on early-life or post-surgical visual exposure, whereas shape perception requires early-life visual experience. This distinction aligns with theories and evidence differentiating dorsal (action-oriented) and ventral (perception-oriented) visual streams. Unlike previous findings on size and weight processing, where grasp performance did not recover after cataract-removal, but perception was preserved, these results suggest unique mechanisms for shape processing. Our findings underscore the developmental trajectories and limitations of late-acquired vision, highlighting implications for tailored rehabilitation strategies. Further studies on processing specific object properties (e.g., size, weight, shape) will be essential to unravel distinct developmental pathways and mechanisms.

(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.

Talk 5, 9:15 am, 31.25

EARLY DEVELOPMENT OF NAVIGATIONALLY RELEVANT LOCATION INFORMATION IN RETROSPLENIAL COMPLEX

Yaelan Jung[1] (jung.yaelan@gmail.com), Daniel Dilks[1]; [1]Emory University

Representing the locations of places is critical for finding our way from a specific place to some distant, out-of-sight place (e.g., from our house to our favorite restaurant in another part of town) – a process referred to as map-based navigation. Neuroimaging work in adults reveals that this ability involves the retrosplenial complex (RSC) – a scene-selective region in the medial parietal cortex. Despite understanding the neural basis of map-based navigation in adults, however, nothing is known about how this system develops. So, does map-based navigation only develop after a decade or more of experience, as generally assumed? Or is it, perhaps counterintuitively, present even in the first few years of life? To directly test this question, using fMRI multivoxel pattern analysis and a virtual town paradigm, we investigated the representation of location information in the RSC of 5-year-olds. We found that i) the RSC in 5-year-olds already represents the locations of particular places and ii) this neural representation is correlated with their performance on a location task. We also found that RSC not only represents the locations of particular places but also the distance between them – another kind of information necessary for map-based navigation. Finally, the parahippocampal place area (PPA) – a scene-selective region hypothesized to be involved in scene categorization, not map-based navigation – did not represent location information, but instead category information, the exact opposite of RSC. Taken together, these findings reveal the early development of navigationally relevant location information in RSC and thus the early development of map-based navigation.

The work was supported by grants from the National Eye Institute (R01 EY29724 to DDD).

Talk 6, 9:30 am, 31.26

THE HOMOLOGY AND DEVELOPMENT OF THE PROTO-WORD AREA IN MACAQUES AND THE VISUAL WORD FORM AREA IN HUMANS

Jia Yang[1] (jiayang@pku.edu.cn), Yipeng Li[1], Wenfang Zhang[3], Haoxuan Yao[1], Jingqiu Luo[1], Hongyu Li[1], Xiaoya Chen[2], Shiming Tang[1], Pinglei Bao[1]; [1]Peking University, [2]Vanderbilt university, [3]China Women's University

The Visual Word Form Area (VWFA) is believed to develop by repurposing a pre-existing area through literacy, but which specific area is repurposed and why it is chosen remain unclear. Given the presence of similar category-selective regions (e.g., face, body, scene) in both macaques and humans, could the human VWFA develop from a proto-word area that can also be identified in macaques? Using fMRI, we identified word-selective regions spanning from the posterior to anterior inferotemporal (IT) cortex in word-naïve macaques. Widefield calcium imaging and high-density electrophysiological recordings confirmed a high concentration of word-selective neurons in these regions through measuring responses to thousands of words and non-word objects. Additionally, objects similar to words in the object space elicited stronger activity, suggesting that proto-word areas in primates may develop through exposure to such objects. This idea is further supported by simulations using deep neural networks. To examine the homology between the macaque word patch and the human VWFA, we conducted human fMRI experiments, showing that the same object space model could explain human VWFA responses. Additionally, a strong correlation between macaque word patch responses and human VWFA responses to 1000 NSD images further supported the homology between these two areas. While both species' word-selective areas follow the object space model, notable differences persist. By measuring responses to nearby objects, faces, and words in human adults and macaques using fMRI, we observed that the macaque word area favors nearby objects over words, whereas the human VWFA shows the opposite preference. Furthermore, responses from preschool (N = 13) and primary school children (N = 17) revealed a shift from a preference for nearby objects to a preference for words as reading experience increased. This study highlights how the human brain repurposes a homologous word-selective area, identified in macaques, to specifically represent words through literacy.

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, 2022T150021to J.Y.

Talk Session: Sunday, May 18, 2025, 10:45 am – 12:30 pm, Talk Room 1

Visual Memory: Neural mechanisms of working memory

Moderator: Geoffrey Woodman, Vanderbilt university

Talk 1, 10:45 am, 32.11

INVERSION OF FEATURE PREFERENCE IN VISUAL CORTICAL NEURONS DURING WORKING MEMORY AND ATTENTION

Diego Mendoza-Halliday[1] (diegomendoza@pitt.edu), Andrii Zahorodnii[2], Haoran Xu[2], Christopher Cueva[2], Julio Martinez-Trujillo[3], Robert Desimone[2]; [1]University of Pittsburgh, [2]Massachusetts Institute of Technology, [3]Western University, London, Ontario, Canada

Dominant theories and models of visual working memory are based on the principle that the maintenance of feature representations in working memory is subserved by the persistence of activity of neurons whose preferred feature closely matches the memorized feature. Similarly, dominant theories and models of attention are based on a feature-similarity gain principle—that the strength of attentional modulation of a given feature-selective neuron is determined by the similarity between its preferred feature and the attended feature. Here, we show experimental evidence that challenges these two principles. In macaque monkeys performing a working memory-guided feature attention task or delayed match-to-sample task for motion direction, we found that in nearly half of the neurons in visual cortical area medial superior temporal (MST), the preferred motion direction during the working memory delay period was opposite to the preferred direction during the stimulus presentation period. A similar inversion in feature preference was observed during feature attention in neuron from both MST and area middle temporal (MT), contrary to that predicted by the feature-similarity gain principle. Interestingly, along the visual processing stream, the percentage of neurons with inversions of feature preference during working memory or attention was highest in earlier processing stages, decreased downstream, and was lowest in later processing cortical areas, including the lateral intraparietal (LIP) and lateral prefrontal cortex (LPFC). These results challenge the generality of theories and models of working memory and attention that assume the predominant recruitment of neurons with preferred features that match the attended or memorized features. Last, dimensionality reduction analyses of neuronal population activity patterns during working memory in MST, and a comparison with competing computational models of working memory, suggest that the observed inversions of feature preference during working memory in visual cortical neurons represent a putative mechanism to protect working memory representations from sensory interference.

Talk 2, 11:00 am, 32.12

FUNCTIONAL RELEVANCE OF THE MEDIAL TEMPORAL LOBE IN VISUAL WORKING MEMORY QUALITY REVEALED BY WITHIN-SUBJECT LESION-SYMPTOM MAPPING

Weizhen Xie[1] (weizhen.xie@nih.gov), Sanikaa Thakurdesai[1], Oceane Fruchet[2], Samantha Jackson[2], Radhika Chatterjee[2], Evalyn Johnson-Ramsay[1], Sara Inati[2], Zaghloul Kareem[2]; [1]University of Maryland, College Park, [2]National Institute of Neurological Disorders and Stroke, National Institutes of Health

Classic lesion case-control studies suggest that the medial temporal lobe (MTL) has minimal involvement in visual working memory (VWM), particularly for simple features like colors and orientation gratings. However, recent findings from direct MTL recordings suggest its crucial role in VWM, potentially through pattern separation that reduces mnemonic interference during short retention intervals. The absence of significant findings in earlier studies may reflect less sensitive task measures or variability across individuals. To address these, we investigated the effects of MTL lesions on VWM representation using a color recall task in 40 neurological cases, assessing performance before and after neurosurgery through lesion-symptom mapping. Of these, 18 cases had MTL lesions, including the hippocampus, while 22 had no lesions or lesions outside the MTL, such as in the insula or prefrontal cortex. Measuring participants' VWM recall variability and overall recall likelihood using a mixture model, we found that MTL lesions led to a significant increase in recall variability, indicating reduced VWM precision post-surgery within individuals. Finer lesion-symptom mapping revealed a strong correlation between hippocampal damage and recall variability, even after controlling for overall lesion size across the whole brain. However, while overall lesion size influenced recall likelihood, reflecting a smaller amount or quantity of remembered VWM content following brain resection in general, hippocampal lesion size could not account for this effect. These findings underscore the MTL's specific role in supporting VWM quality, distinct from VWM quantity supported by distributed neocortical mechanisms. This dissociation challenges unitary models of VWM constraints that overlook the distinction between memory quality and quantity, emphasizing the need for updated frameworks incorporating critical neuropsychological evidence.

W.X. is supported by NIH grant R00NS126492

Talk 3, 11:15 am, 32.13

COMPUTATIONAL METHODS FOR EXTRACTING NEURAL CORRELATES OF WORKING MEMORY AND MENTAL IMAGERY FROM INTRACORTICAL RECORDINGS IN HUMAN VISUAL CORTEX

Jacob Granley[1] (jgranley@ucsb.edu), Lily M. Turkstra[1], Galen Pogoncheff[1], Fabrizio Grani[2], Leili Soo[2], Alfonso Rodil[2], Cristina Soto[2], Thomas C. Sprague[1], Eduardo Fernandez[2], Michael

Beyeler[1]; [1]University of California, Santa Barbara, [2]University of Miguel Hernandez, Elche, Spain

Neural activity in early visual cortex (EVC) is known to contribute to visual working memory (WM) and mental imagery (MI), but the role of spiking activity in humans remains unclear. This study investigates computational techniques for extracting spiking activity and their ability to reveal correlates of WM and MI. Intracortical recordings were collected from two awake blind humans implanted with a 96-channel visual prosthesis in EVC during a delayed-match-to-sample (DMTS) WM task and a MI visualization task. In 465 trials of the WM task, participants encoded visual perceptions (phosphenes) elicited by stimulation of one of three electrodes, maintained them over a 5-second delay, and recalled whether a subsequent phosphene was the same or different. The MI task followed a similar structure, with recall replaced by vivid mental visualization. Neural activity during stimulation, delay, recall, and spontaneous periods was analyzed using methods to extract multi-unit activity (MUA), entire spiking activity (ESA), and local field potential (LFP) signals. Significant differences were observed in MUA, ESA, and LFP (theta, alpha, and beta bands) across trial periods (t-tests, $p < 0.05$). ESA and MUA exhibited electrode-specific neural signatures during delay and recall periods, with over 90% classification accuracy in leave-one-trial-out cross-validation (LOOCV). Stimulus-specific ESA changes remained decodable throughout delay and recall (random forest classifier sliding window, LOOCV, 70% of windows above chance), indicating sustained stimulus-selective activity for both tasks despite day-to-day variability. These findings reveal sustained stimulus-selective spiking activity in human EVC during WM and MI tasks, underscoring its critical role in retaining and recalling information and providing new insights into the neural mechanisms underlying perception and cognition.

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

TALK 4, 11:30 AM, 32.14

HOW LONG-TERM LEARNING ALTERS VISUAL WORKING MEMORY REPRESENTATIONS: EVIDENCE FROM EEG

Philipp Musfeld[1], William X. Q. Ngiam[2], Kirsten C. S. Adam[3], Olga Kozlova[3], Olya Bulatova[4], Keisuke Fukuda[4,5]; [1]University of Zurich, [2]The University of Adelaide, [3]Rice University, [4]University of Toronto, [5]University of Toronto Mississauga

Visual Working memory (VWM), our central system for temporarily holding visual information in mind for further thought and action, is limited in capacity. To overcome such limitations, we frequently leverage prior knowledge from visual long-term memory (VLTM), allowing us to integrate and represent information more efficiently. Yet, it is not well understood how prior knowledge affects the representation of information in VWM. Here, we assessed how prior learning affects the load of VWM representations by using multivariate load classification from EEG. Participants (N=30) learned a 6-color visual array to criterion, and then completed a VWM task including both new and pre-learned arrays. Crucially, new arrays differed in set size (0, 1, 2 or 6), which we used to train a classifier to identify the load in VWM from the multivariate EEG signal. After establishing strong classification accuracy (~55%; chance = 25%), we asked the classifier to predict the load elicited by pre-learned arrays. We find evidence that the availability of VLTM for pre-learned arrays reduced load in VWM, as the classifier predicts a load of 1 or 2, instead of 6 – the actual set size of pre-learned arrays. However, further exploration revealed that representations of pre-learned arrays were still dissociable from pure VWM representations of lower set sizes. This was further supported by representational similarity analyses, which suggested that the obtained representational pattern for pre-learned arrays was best explained by a model assuming 1) a reduction in load in VWM together with 2) a distinct contribution from VLTM to the representation. We conclude that the availability of prior knowledge reduces the load in VWM but leads to qualitative changes in multivariate neural signals, potentially rendering memory representations more distinctive.

This research was supported by the Natural Sciences and Engineering Research Council (5009170).

TALK 5, 11:45 AM, 32.15

BEHAVIORAL AND NEURAL EVIDENCE FOR DISSOCIABLE SUBPROCESSES WITHIN VISUAL WORKING MEMORY

Gayathri Satheesh[1] (gayathri.satheesh@nyu.edu), A. J. Abdujabborov[1], Kartik K. Sreenivasan[1,2]; [1]Division of Science and Mathematics, NYU Abu Dhabi, [2]Center for Brain and Health, NYU Abu Dhabi

Is working memory (WM) a unitary construct or is it composed of separable sub-processes with distinct neuroanatomical bases? In line with this latter view, WM engages a wide network of brain regions (Brissenden et al., 2018; Rahmati et al., 2020) and neurological patients display deficits in specific aspects of WM rather than global impairments (Cañas et al., 2018; Lee et al., 2010). We tested whether WM is composed of dissociable processes using a task requiring participants to maintain the locations of multiple discs over a memory delay and report the location of the cued disc. The task selectively engaged putative subcomponents of WM (storage, selection, distractor resistance, updating, and manipulation) by varying the set size, presence of task-irrelevant information during encoding or maintenance, and requirements to reformat or manipulate memory content. We collected behavioral data (n = 200) across two sessions and compared performance within and across subcomponents over sessions. Employing dimensionality reduction techniques and model fitting, we found that participants' behavior was best described by a model that includes at least four subcomponents. These results argue against a unitary WM construct. Next, we collected fMRI data (n = 30) to determine whether these putative subcomponents map onto distinct neural substrates. We found heterogeneous activation clusters that were unique to individual subcomponents as well as regions of overlapping activation. To further tease apart the neural basis of WM subcomponents, we used representational similarity analysis to compare activity within and between subcomponents across regions of interest. The resulting dissimilarity matrix was compared to theoretical models of possible subcomponent combinations. The best-fitting models indicated distinct activity patterns across subcomponents, particularly in frontal and parietal regions. Overall, our neural data was also best explained by a multi-subcomponent model. Together, these findings indicate that WM involves multiple subcomponents with distinct patterns of neural activation.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYUAD Research Institute grant CG012, and the ASPIRE Award for Research Excellence (AARE-19-230).

Talk 6, 12:00 pm, 32.16

NEURAL OSCILLATIONS ENABLE CONCURRENT VISUAL PERCEPTION AND VISUAL WORKING MEMORY PROCESSING

Khayla Santiago[1], Chunyue Teng[1]; [1]Lawrence University, Appleton, WI

Successful goal-directed behavior requires balancing task-relevant information stored in working memory and the continuous processing of incoming sensory input, yet the neural mechanisms underlying this coordination remain unclear. The current study investigates the role of neural oscillations in supporting concurrent visual perception and visual working memory. We recorded electroencephalogram (EEG) while healthy human participants performed a dual task paradigm 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 stimulus onset asynchrony (SOA) , they were prompted to compare a test probe against either the memorized orientation or the visually monitored orientation. Critically, we manipulated the duration of the SOA: 500-1500 ms with a 20 ms step, resulting in a total of 50 SOAs. 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, and the analysis identified increased power within the theta and low-alpha frequencies 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 external and internal visual representations. Additionally, Inverted Encoding Models (IEM) were applied on EEG data, and successfully reconstructed the orientations of both visual working memory and perceptual representations during periods of concurrent task relevance. 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 segregating visual representations of different sources.

Talk 7, 12:15 pm, 32.17

ELECTRICAL STIMULATION OF VISUAL CORTEX ENHANCES VISUAL WORKING MEMORY FIDELITY

Xinger Yu[1] (xinger.yu@vanderbilt.edu), Gengshi Hu[1], Geoffrey Woodman[1]; [1]Vanderbilt University

Theories of visual memory suggest that the same neurons responsible for object perception also store the memory representations of these objects. This study tests the sensory recruitment hypothesis of visual working memory, which predicts that enhancing perceptual precision improves the fidelity of memory representations. We employed a combination of noninvasive transcranial direct current stimulation (tDCS), behavioral testing, and electrophysiological measures to causally manipulate neural activity in the visual cortex and evaluate its effect on memory performance. Experiment 1 involved a continuous report visual working memory task in which participants recalled the colors of items presented in varying set sizes (one, two, four, or six items). Anodal tDCS over the visual cortex (P1 or P2 site, International 10-20 System) enhanced memory precision, particularly by reducing the variability in recall errors for items located in the hemifield contralateral to the stimulation site. Experiment 2 used a change detection paradigm with retrocues, demonstrating increased accuracy for items in the contralateral hemifield following anodal tDCS, whereas no improvement was observed during sham stimulation. Simultaneous EEG recordings revealed greater alpha power suppression during the maintenance phase of the working memory tasks with anodal stimulation, suggesting a role for attention in enhancing memory performance. These findings provide causal evidence that tDCS can modulate visual cortex activity, thereby improving the fidelity of visual working memory representations. The results underscore the functional overlap between perception and memory, supporting the sensory recruitment hypothesis and highlighting the potential of noninvasive brain stimulation to improve cognitive performance.

This work was supported by grants from the National Science Foundation (BCS-2147064) and the National Eye Institute (T32-EY007135), as well as funding by the G. Forrest Woodman Foundation for the Advancement of Neuroscience.

Talk Session: Sunday, May 18, 2025, 10:45 am – 12:30 pm, Talk Room 2

Object Recognition: Models

Moderator: Hans Op de Beeck, University of Leuven

Talk 1, 10:45 am, 32.21

CONNECTING THE DOTS: SIMILAR NEURAL AND BEHAVIOURAL REPRESENTATIONS FOR VISUAL BRAILLE AND LINE-BASED SCRIPTS

Filippo Cerpelloni[1,2], Olivier Collignon[2,3], Hans Op de Beeck[1*]; [1]Department of Brain and Cognition, Leuven Brain Institute, KU Leuven, Belgium, [2]Institute of Psychology (IPSY) & Institute of Neuroscience (IoNS), UCLouvain, Belgium, [3]HES-SO Valais-Wallis, The Sense Innovation and Research Center, Lausanne & Sion, Switzerland*

Visual object recognition relies on shape processing and progressive integration of line-junctions. Reading has been thought to co-opt this central mechanism in object perception since most scripts share similar visual features like line-junctions. We present a series of studies on visual Braille, a script developed for touch that possesses no explicit shape cues. Behaviourally, we show similar script acquisition in novice participants who learn visual Braille or a custom-made script with Braille dots joined into lines (Line Braille). Across four days of visual training, we do not observe differences in the accuracy and speed of transcription of words from the novel to the native script. Only a limited advantage is present in the initial training on single letters, quickly levelled by training on full words. In the brain, the Visual Word Form Area (VWFA) preferentially responds to Braille over scrambled dots in expert visual Braille readers, and not in naïve controls. Moreover, we can decode the linguistic properties of letter strings (e.g. words vs. pseudo-words) in both groups for their known scripts. The representational similarity between conditions in visual Braille for experts correlates to the similarity structure for the active line-based script. We observe this pattern in several key regions of the visual stream (V1, LO), although without being able to decode linguistic properties across scripts, and in linguistic regions (l-PosTemp), where we also found cross-script decoding. Lastly, we replicated in computational models the designs tested in human, to separate visual and linguistic influences. Overall, converging evidence show that the linguistic properties of a visual script, rather than its low-level line-junctions, play a major role both in how an individual approaches reading, and how the visual system, and VWFA in particular, processes scripts.

Talk 2, 11:00 am, 32.22

A BAYESIAN MODEL OF CAMOUFLAGE DETECTION BY HUMANS

Abhranil Das[1], Wilson Geisler; [1]University of Texas at Austin

Camouflage is an impressive feat of biological evolution, but so is its detection by predators and prey. Well-camouflaged animals copy the luminance, contrast, colour and texture of their natural backgrounds, leaving only the animals' boundary available for detection. This is one of the hardest cases of visual detection and reveals many detection strategies and their limits. We conducted experiments where humans detect synthetic camouflaged targets across varying conditions, including different textures. To explain this data, we developed a principled detection model that follows human optics and biologically plausible computations and is informed by the statistics of the relevant features in natural images. The model filters an image with human optics, then computes edge gradients and groups them into edge contours. It then computes several contour features: the fraction of area they cover, their lengths, position and orientation alignments with the true target boundary, curvatures, and edge powers at 5 scales. Additionally, it computes histograms of edge gradient magnitude, orientation, and their product (a proxy for their correlation) across all pixels. In parallel, we model the statistical distribution of each of these features over natural images, by computing them on our database of optics-filtered natural image patches. Using these known feature distribution families, we then construct optimal Bayesian decision variables that measure whether the features in the camouflage image are the same over the target boundary region, as compared to outside it. We combine these feature decision variables using a multivariate Gaussian model, which outputs a final detection response, as well as the relative contribution of each feature to detection. We fit these to our experimental data so that the single principled model can predict human camouflage detection performance across our entire array of diverse stimuli, and account for many of our parametric experimental observations.

Talk 3, 11:15 am, 32.23

DOMAIN-GENERAL OBJECT RECOGNITION ABILITY HAS PERCEPTION AND MEMORY SUBFACTORS

Conor J. R. Smithson[1] (conor.smithson@vanderbilt.edu), Isabel Gauthier[1]; [1]Vanderbilt University

The ability to make subordinate-level object identity judgements tends to be general across object categories. This general object recognition ability (o) is measured using tasks in which errors result from pressure on either perceptual or memory mechanisms. For example, perception can be taxed by high target-distractor similarity, short presentation times, and visual noise, while memory can be taxed by having multiple targets and delays between study and test. Both task types have been assumed to measure a single o construct. Using structural equation modeling we asked whether object memory and perception abilities are separable, whether they both reflect a general o factor, and whether they are fully explained by established visual, memory, and cognitive abilities. Participants completed eight object recognition tests, four challenging perception and four challenging memory, alongside tests of visual, memory, and intelligence abilities. A model with separate object perception and memory factors fit significantly better than a model with a single object recognition factor. Low-level visual discrimination ability uniquely predicted object perception, but not object memory, while working memory uniquely predicted object memory, but not object perception. General intelligence uniquely predicted both object memory and perception to a similar extent. After regressing out these constructs, the residual correlation between object perception and object memory was very high, with the majority of remaining variance being shared between them. In a model where object memory and object perception were subfactors of a higher-order o, both loaded strongly onto o, while each subfactor's non-o variance was completely explained by either low-level visual discrimination ability or working memory. The higher-order o factor was strongly related to general intelligence. Object perception and memory abilities are differentiable, but both substantially reflect a more general o factor. o accounts for variance in object recognition tests that is not already explained by established cognitive abilities.

This work was supported by the David K. Wilson Chair Research Fund from Vanderbilt University and NSF BCS Award 2316474.

Talk 4, 11:30 am, 32.24

EMBEDDING OBJECT-SCENE RELATIONSHIPS: INSIGHTS FROM HUMAN BEHAVIOR AND VISION-LANGUAGE MODELS

Karim Rajaei[1] (rajaei.k@ipm.ir), Hamid Soltanian-Zadeh[2,1]; [1]School of Cognitive Science, IPM, Tehran, Iran, [2]University of Tehran

In real-world environments, objects are embedded within scenes, where semantic and syntactic relationships guide perception. The human brain efficiently encodes these contextual regularities, embedding object-scene and object-object relationships into neural systems. However, the mechanisms by which context influences object recognition and the extent to which computational models replicate these effects remain poorly understood. To address this, we generated a novel, parametrically controlled dataset using an embodied AI platform (OmniGibson). The dataset included 48 objects presented within either scene or phase-scrambled backgrounds, with parameters such as object distance and viewing angle, occlusion, scene lighting, and crowding varied to create both "simple" and "challenging" recognition tasks. Behavioral experiments assessed human recognition accuracy, while computational experiments evaluated conventional CNNs (e.g., AlexNet, ResNet), transformers (e.g., ViT), vision-language models (e.g., CLIP), and multimodal models. Behavioral results demonstrated that humans consistently recognized objects more accurately in scene contexts than in scrambled backgrounds; advantage most pronounced under challenging conditions (70% vs. 54% accuracy for scene and scrambled conditions, respectively; p<0.001, two-sided sign-rank-test). Computational results revealed that vision-only models failed to achieve human-level performance, even in simpler tasks (83.5% vs. 54% accuracy for humans and ResNet, respectively; p<0.001, two-sided sign-rank-test). In contrast, models incorporating language supervision, such as CLIP (87.5% accuracy), or multimodal training approached human-level performance but still fell short under most challenging conditions. These findings highlight that language-aligned models may embed object-scene relationships into their visual representations and utilize semantic relationships through language-aligned readout mechanisms. This underscores the importance of integrating contextual regularities into computational frameworks, suggesting that multimodal training paradigms incorporating language may better capture the dependencies inherent in real-world perception. Unique contributions of this study include the creation of a parametrically controlled naturalistic object-scene dataset and a direct comparison of human and state-of-the-art computational models in their ability to recognize objects within scene contexts.

TALK 5, 11:45 AM, 32.25

MOSAIC: AN AGGREGATED FMRI DATASET FOR ROBUST AND GENERALIZABLE VISION RESEARCH

Benjamin Lahner[1], N. Apurva Ratan Murty[2], Aude Oliva[1]; [1]Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA., [2]School of Psychology, Georgia Institute of Technology, Atlanta, GA, USA.

Large-scale fMRI datasets are revolutionizing our understanding of the neural processes underlying human perception, driving new breakthroughs in neuroscience and computational modeling. Yet individual fMRI data collection efforts remain constrained by practical limitations in scan time, creating an inherent tradeoff between subjects, stimuli, and stimulus repetitions. This tradeoff often compromises stimuli diversity, data quality, and generalizability of findings such that even the largest fMRI datasets cannot fully leverage the power of high-parameter artificial neural network models and high-dimensional feature spaces. To overcome these challenges, we introduce MOSAIC (Meta-Organized Stimuli And fMRI Imaging data for Computational modeling): a scalable framework for aggregating fMRI responses across multiple subjects and datasets. We preprocessed and registered eight event-related fMRI vision datasets (Natural Scenes Dataset, Natural Object Dataset, BOLD Moments Dataset, BOLD5000, Human Actions Dataset, Deeprecon, Generic Object Decoding, and THINGS) to the fsLR32k cortical surface space with fMRIPrep to obtain 426,245 fMRI-stimulus pairs over 93 subjects and 163,202 unique stimuli. We estimated single-trial beta values with GLMsingle (Prince et al., 2022), obtaining parameter estimates of similar or higher quality than the originally published datasets. Critically, we curated the dataset by eliminating stimuli with perceptual similarity above a defined threshold to prevent test-train leakage. This rigorous pipeline resulted in a well-defined stimulus-response dataset with 145,007 training stimuli, 18,145 test stimuli, and 50 synthetic stimuli well-suited for building and evaluating robust models of human vision. By unifying datasets under an identical preprocessing and registration pipeline, MOSAIC allows researchers to circumvent the limitations of individual datasets and address complex research questions with unprecedented scope. This framework is also extensible: new datasets can be seamlessly incorporated post hoc, enabling this meta-dataset to evolve alongside advances in experimental design and methodology. This adaptability empowers the neuroscience community to collaboratively generate a scalable, generalizable foundation for studying human vision.

TALK 6, 12:00 PM, 32.26

WEIGHT-SIMILARITY TOPOGRAPHIC NETWORKS IMPROVE RETINOTOPY AND NOISE ROBUSTNESS

Nhut Truong[1] (leminhnhut.truong@unitn.it), Uri Hasson[1]; [1]University of Trento

Typical deep neural networks (DNNs) lack spatial organization and a concept of unit adjacency. In contrast, topographic DNNs (TDNNs) spatially organize units, and are therefore potential spatio-functional models of cortical organization. In previous work, this spatial organization was achieved by adding a loss term that encourages adjacent neurons to exhibit similar activation patterns (activation-similarity, AS-TDNN). However, this optimization is not biologically grounded, and ideally, these correlations should arise naturally as a consequence of biologically motivated constraints. This led us to develop a new type of TDNN, whose training is grounded in the biologically-inspired principle that spatially adjacent units should have similar afferent (incoming) synaptic strength, modeled by similar incoming weight profiles (weight-similarity, WS-TDNN). Using hand-written digit classification (MNIST) as a test domain, we compared the properties of AS-TDNNs, WS-TDNNs, and a control (non-topographic) DNN. Both AS-TDNNs and WS-TDNNs were tested under six different weighting levels for the spatial loss term. While all models achieved nearly identical classification accuracy, WS-TDNNS showed several positive advantages, including greater robustness to several types of noise, greater resistance to node ablation, and higher unit-level activation variance. Interestingly, WS-TDNNs produced higher correlations between adjacent units than AS-TDNNs, even though the latter were explicitly trained on this objective. Importantly, when tested using standard retinotopy protocols (i.e., rotating wedge and eccentric ring stimuli), WS-TDNNs, but not AS-TDNNs, naturally produced angular and eccentricity-based spatial tuning. This was evident in the smooth transitions in units' preferred angles and spatial grouping by preferred eccentricity. Moreover, these properties naturally emerged through end-to-end training, without requiring separate pre-optimization steps required in recent studies. These results were also replicated using the CIFAR-10 dataset for object recognition. Overall, our results suggest that TDNNs trained with weight-similarity constraints are viable computational models for visual cortical organization.

Talk 7, 12:15 pm, 32.27

RE-EVALUATING THE ABILITY OF OBJECT TRAINED CONVOLUTIONAL NEURAL NETWORKS FOR CLASSIFYING 'OUT OF DISTRIBUTION' IMAGES

Connor J. Parde[1], Hojin Jang[2], Frank Tong[1,3]; [1]Psychology Department, Vanderbilt University, [2]Department of Brain and Cognitive Engineering, Korea University, South Korea, [3]Vanderbilt Vision Research Center, Vanderbilt University

Convolutional neural networks (CNNs) trained for object classification are severely impaired by almost any form of image degradation (e.g., visual noise, blur, phase-scrambling, etc.) unless they receive direct training (Geirhos et al., 2018; Jang, McCormack & Tong, 2021). Although the effect of direct training is not believed to generalize to un-trained distortions, recent work has demonstrated that blur-trained CNNs achieve better human accord and exhibit higher general performance than clear-trained CNNs (Jang & Tong, 2024). Here, we evaluated systematically the performance of CNNs when trained and tested on object images that underwent different types of manipulation. We trained ResNet-50 and VGG-19 models on images from the ILSVRC12 dataset using either no manipulation, low-pass filters, high-pass filters, uniform noise, salt-and-pepper noise, or phase-scrambled noise. In addition, separate models were trained using either 1000-category or 16-category classification tasks and then tested with the full set of image manipulations. In all cases, 1000-category trained models exhibited higher performance than their 16-category trained counterparts. In addition, all models performed highest on images with the same type of manipulation that was present during training. There was no difference in the generalizability of the 1000-class trained models and the 16-class models. However, the low-pass filter trained models outperformed their clear-trained counterparts for images with uniform noise and salt-and-pepper noise, and maintained similar performance to the clear-trained models for all other types of manipulation. Further, uniform-noise trained and salt-and-pepper-noise trained CNNs outperformed clear-trained CNNs on images with low-pass filters, uniform noise, and salt-and-pepper noise. In addition, high-pass trained CNNs outperformed clear-trained CNNs on phase-scrambled images. Our results demonstrate that CNNs trained on degraded images exhibit some ability to generalize performance to images outside of their training distribution. This underscores the importance of challenging or degraded stimuli for learning robust representations of visual categories.

This research was supported by NEI grants R01EY035157 to FT and P30EY008126 to the Vanderbilt Vision Research Center

Talk Session: Sunday, May 18, 2025, 2:30 – 4:15 pm, Talk Room 1

Perceptual Organization: Objects, events, ensembles

Moderator: Samuel McDougle, Yale University

Talk 1, 2:30 pm, 34.11

THE CROWD SIZE ILLUSION

Gabriel Waterhouse[1], Sami Yousif[1]; [1]University of North Carolina at Chapel Hill

After several relaxing days on the beach, you step up to the podium to give your talk. Looking out on the audience, you find that some, but not all, of the seats are occupied. The crowd looks empty. But what if your impression is an illusion? What if the number you perceive is influenced not just by the number of filled seats, but the number of empty ones? We tested this putative “crowd size illusion” by having participants compare and estimate the numbers of dots in displays with and without visible grids: Some displays contained random arrangements of dots, whereas others contained dots arranged within “cells” in a grid (like people, in seats). When only about 15-30% of the “seats” were occupied, people tended to find the display with the grid to have fewer dots (consistent with the intuition described above). In a second experiment, we replicated this finding in a direct estimation task. But does the presence of a grid always result in underestimation? Imagine the same scenario as before, except the audience is full to the brim. In that case, might you perceive more people? In a third experiment, we found that when occupancy of the grids was high, those displays were perceived as more numerous. This effect is continuous: Observers underestimated the number of dots when more “seats” were empty and overestimated when more “seats” were full. Both illusions are powerful enough that they are readily appreciated in simple demonstrations. Furthermore, the fact that the direction of the illusion depends on the percentage of occupied cells indicates that this illusion cannot be explained by confounds with continuous spatial properties. In this way, the crowd size illusion is more than a curiosity: It points to a number system that represents number in a part-whole format.

Talk 2, 2:45 pm, 34.12

AFTEREFFECTS OF NUMEROSITY ARE CAUSED BY DENSITY AND SIZE ADAPTATION

Frank H. Durgin[1], Abigail Love[1], Grace Taylor[1]; [1]Swarthmore College

DeSimone et al. (2021) sought to show that number aftereffects could be distinguished from density aftereffects by using an adapter that was more numerous, but lower in density than the target display. (It was also larger.) They interpreted downward aftereffects in the perception of numerosity as evidence that number was directly adapted, independent of density. However, size adaptation is also known to affect perceived number (Zimmermann & Fink, 2016), and the size ratio used by DeSimone et al. was nearly 3 to 1, while the density ratio of their adapter was only 0.7 to 1, a ratio which produces essentially no upward density adaptation (Sun et al., 2017). Thus, the downward numerosity aftereffect observed may have been due to size adaptation alone. We used the same test stimulus that DeSimone et al. used (30 dots in a circle of 10 deg2, centered 3.5° from fixation), but we varied the adapters more systematically and directly measured matches for perceived size and perceived density as well as for perceived number after adaptation. In the critical conditions in which adapter size was 3x that of the test stimulus, and adapter density was 1/3 that of the test stimulus perceived area was decreased by about 15%, while perceived density was increased about 20%, and perceived number was increased slightly, but reliably, as well (by about 5%, as predicted by combining size and density effects). Thus, DeSimone et al.'s (2021) evidence of downward "number" adaptation was likely due to downward size adaptation (to the 3:1 size ratio) in the absence of any upward density adaptation. This new preregistered observation shows that number adaptation is not easily dissociated from aftereffects of adaptation to patch size and to dot density.

Talk 3, 3:00 pm, 34.13

WHEN ARE DOT ARRAYS PERCEIVED AS SHAPES? EVIDENCE FROM CONFIGURAL SUPERIORITY PARADIGMS

Nicholas Baker[1] (nbaker1@luc.edu), Wiliam Friebel[1], Alexa Vushaj[1], Peyton Daly[1], Madeline Geittmann[1], Mihika Tewari[1]; [1]Loyola University of Chicago

The visual system is so attuned to shape that even sparse arrays of disconnected dot elements (such as star constellations) are sometimes perceived as contours. Because they contain so little information, arrays of dots may be the minimum signal necessary to give rise to a shape percept. Understanding what relations between dots result in a perceived shape offers insight into the process by which the visual system forms shape representations from sensory elements. One of the challenges of this research is that it is phenomenological: Whether an array of dots appears to be a shape is a question of an observer's subjective experience. We made use of a long history of perceptual organization research showing configural advantage effects to quantify the degree to which dot arrays are perceived as shapes using objective tasks. We used two well-known paradigms: object-attracted attention in visual search (Kimchi et al., 2016) and configural superiority effects (CSE) (Pomerantz et al., 1977). We began by testing these methods on a manipulation known to affect shape perception: a shape's angularity (Baker & Kellman, 2024). In Experiment 1, participants identified the orientation of a target within or outside an array of dots sampled from a smooth vs. angular shape. In Experiment 2, participants identified a target with either 12 dots alone or with a noninformative context added that completed the shape of the target. Both paradigms revealed greater configural advantages for smooth contours. We then tested another possible factor in arrays' perceived shapehood: the ratio of a contour's area to its perimeter. We repeated both experiments, this time comparing dots sampled from shapes with high area: perimeter ratios with dots sampled from shapes with low area: perimeter ratios. Both experiments showed a greater configural advantage for shapes with larger ratios of area to perimeter.

Talk 4, 3:15 pm, 34.14

'VISUAL VERBS' DRIVE ADAPTIVE PREDICTIONS: PERCEPTION OF DYNAMIC EVENT TYPES SPONTANEOUSLY CHANGES VISUAL WORKING MEMORY ENCODING

Huichao Ji[1] (huichao.ji@yale.edu), Brian Scholl[1]; [1]Yale University

We see the world not only in terms of specific features (such as the color or shape of a ball), but also in terms of a foundational set of abstract/categorical 'event types' (such as a ball bouncing vs. rolling). Recent work has demonstrated that such categorical perception occurs spontaneously during passive viewing of visual scenes, even when verbal encoding is discouraged or disrupted: observers are better able to detect changes across different event types, even when the magnitudes of within-type changes (e.g. across two different animations of bouncing) are objectively greater. Why might this occur? Here we explored the possibility that such spontaneous categorical encoding is adaptive, insofar as it enables differential predictions about likely future states, and so changes what is encoded into memory. This was inspired by the idea that the purpose of perception is not only to characterize the present ("What's out there?") but also to predict the future ("What's about to happen?"). We studied this in a single-trial memory task, e.g. when contrasting bouncing vs. rolling animations: observers viewed a single animation of a ball moving, and then simply reported its final position (after the video had ended and the display had disappeared). This placement was systematically biased by the underlying event type: rolling balls tended to be localized as further back in their actual trajectories horizontally (but not vertically), compared to bouncing balls -- presumably because a bouncing ball can only move forward in the coming moments, while a rolling ball could roll backwards down a ramp. And careful controls showed that this depended on the event-type itself, rather than any lower-level properties (such as the details of the trajectories). This shows how representations of 'visual verbs' might drive adaptive predictions about how a dynamic world is likely to unfold.

Talk 5, 3:30 pm, 34.15

DISSOCIATING EXTERNAL FEATURES FROM INTERNAL STRUCTURES IN VISUAL SEGMENTATION OF ACTIONS

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

The human mind tends to represent continuous experience as discrete events, imposing "event boundaries" on incoming streams of sensory data. This phenomenon, known as event segmentation, is not just a function of top-down decisions about where events begin and end – event boundaries appear to structure attention and perception as well. How are event boundaries represented perceptually? Previous studies on event structure typically employ remarkable changes of physical features at event boundaries, e.g., walking through a doorway, large shifts in objects, figures and scenes, salient motion cues, and disruptions of visual statistics. This raises an important question: Is the perceptual representation of an event boundary exclusively driven by processing salient, lower-level physical changes and motion dynamics? Or might higher-level semantic structure also shape how we perceive event boundaries? Here we attempt to disentangle low-level spatiotemporal features of continuous visual input versus high-level representations of natural action structure. Across six pre-registered experiments, we asked participants to detect subtle disruptions in 20 short, individual actions (e.g., kicking a ball, stepping over an obstacle, throwing a frisbee, etc., generated as motion-capture-based simple animations, static images, and point-light biological motion displays), which were presented either in a recognizable intact form or in a distorted manner that only preserved low-level spatiotemporal dynamics and visual features. Results consistently demonstrated that visual detection of subtle disruptions was weaker at action boundaries (i.e., the transition between discrete steps within the action) relative to non-boundaries, extending previous findings to naturalistic action perception. Crucially, these perceptual effects were driven by both lower-level visual features and by high-level information of action structure. Thus, automatic and rapid perceptual segmentation of actions is likely structured in time by both external cues inherent to the stimulus and our internal models of the world.

This work is supported by NIH grant R01 NS13292

Talk 6, 3:45 pm, 34.16

DISCRETE VS. CONTINUOUS TIMER BARS: HOW VISUAL SEGMENTATION SHAPES THE PERCEPTION OF TIME "RUNNING OUT"

Jasmindeep Kaur[1], Jiaying Zhao[1], Joan Danielle K. Ongchoco[1]; [1]The University of British Columbia

Our lives are flooded with visual reminders of time slipping away — from ticking clocks to countdowns timers, that all depict a sense of time “running out”. In time perception, the same duration can feel longer or shorter as a function of various factors (e.g., attention, predictability) — but we know less about the factors that influence the perception of how much time is left. In visual processing, a key discovery is that while sensory input may be a continuous wash of light, what we experience — what the mind parses — are discrete objects and events. Here we explored how discreteness structures our sense of time running out. Observers completed a multi-item localization (MILO) task, where they clicked on multiple targets in a sequence. In every trial, there was a black-bordered rectangular ‘timer-bar’ initially filled with a color that emptied over a period (e.g., 3 seconds) to visually depict the passage of time. The color diminished either *continuously*, gradually and evenly depleting throughout, or *discretely*, in which the bar was segmented into discrete chunks that disappeared at regular intervals. To measure perceived urgency of time ‘running out’, we examined inter-click latencies (i.e., the time between clicks). Results revealed longer inter-click latencies for discrete (compared to continuous) timer-bars, suggesting greater urgency in the continuous case. This difference disappeared in a separate experiment, where the bar was instead filled over time continuously or discretely, with a reliable interaction between experiments — suggesting that effects could not simply have been a function of one condition being more distracting than another. Thus, discreteness may have distinct effects on our sense of time running out versus time accumulating. Segmentation in visual depictions of time depletion may make time feel more “manageable,” altering our sense of urgency in time-sensitive tasks.

Talk 7, 4:00 pm, 34.17

PROCESSING FLUENCY MEDIATES TRUST IN DATA VISUALIZATIONS

Hamza Elhamdadi[1], Suyeon Seo[2], Lace Padilla[3], Cindy Xiong Bearfield[2]; [1]University of Massachusetts Amherst, [2]Georgia Institute of Technology, [3]Northeastern University

Trust plays a significant role in how people perceive scientific information and make critical decisions. Information can be discounted or dismissed without mutual trust between the audience and the presenter. Therefore, establishing trust is a critical first step in visual data communication. Drawing on theories of visual perception, we investigate the role of processing fluency—the ease with which visual stimuli are encoded and processed—in shaping trust in visualizations, using scatter plots as our case study. Through two empirical studies, we demonstrate that visualization design can impact processing fluency, leading to altered trust judgments. In Experiment 1, we validated perceptual fluency manipulations using scatterplots through design manipulations based on prior perception and visualization research, such as adding gridlines, introducing blur, or varying data mark transparency. Participants completed a perceptual task estimating the proportion of points in a specific range and rated task difficulty. We found that fluent visualizations yielded higher accuracy and lower difficult ratings, while manipulations to create disfluent visualizations led to worse performances. In Experiment 2, we created a decision task based on trust games adapted from behavioral economics. Participants allocated resources between two hypothetical companies, each presenting their investment strategies using a scatter plot. We manipulated the relative processing fluency of the plots and found that participants tended to allocate fewer resources to the company presenting data with a disfluent plot. These findings highlight the critical role of perceptual processing in trust and suggest that optimizing processing fluency of data visualizations can enhance their perceived trust and their ability to effectively communicate.

NSF IIS-223758 and IIS-2311575

Talk Session: Sunday, May 18, 2025, 2:30 – 4:15 pm, Talk Room 2

Binocular Vision

Moderator: Fulvio Domini, Brown University

Talk 1, 2:30 pm, 34.21

REAL-WORLD STATISTICAL REGULARITY IN BINOCULAR RIVALRY: THE ADVANTAGE OF GOOD EXEMPLARS

Yiwen Wang[1] (yiwenw11@illinois.edu), Ling Lee Chong[1], Diane M. Beck[1]; [1]University of Illinois at Urbana-Champaign

Real-world statistical regularities, unlike regularities introduced in experimental settings, are learned through extensive exposure over a lifetime in the natural visual environment. These natural patterns reflect the consistent features and structure of real-world scenes and objects. By utilizing these predictable patterns, the visual system processes statistically regular stimuli more quickly and effortlessly than less regular ones, enhancing perceptual efficiency (Beck, Center, Shao, 2024). One example of real-world statistical regularity is the distinction between good and bad exemplars images. Good exemplars, which are highly representative of a category, are more easily detected than bad exemplars. Building on this, the current study aimed to investigate whether statistical regularity influences perceptual selection in binocular rivalry, where conflicting images presented to each eye compete for the dominance of perception. In this study, participants were shown two images from the same scene category (i.e., beach, mountain, city, highway). One image was a good exemplar, and the other was a bad exemplar of that category, with each image presented to a different eye. Results revealed statistical regularity biased the perceptual selection to the good exemplar: good exemplars were more likely to be selected as the initial percept and had faster perceptual onset times compared to bad exemplars. Our results align with the predictive coding framework of binocular rivalry, where good exemplars, with higher priors, would be more likely to dominate perception over bad exemplars. These findings extend findings on statistical learning (e.g. Denison, Piazza, Silver, 2011) to real-world statistical regularities that are derived over a lifetime rather than within an experiment.

Talk 2, 2:45 pm, 34.22

IS AVERAGE DOMINANCE PHASE DURATION A RELIABLE MEASURE OF MULTISTABLE STIMULI DYNAMICS?

Alexander Pastukhov[1,2] (pastukhov.alexander@gmail.com), Paula Finkenauer[1], Leonie Littek[1], Lea Voss[1], Claus-Christian Carbon[1,2]; [1]University of Bamberg, [2]EPÆG Research Group, Bamberg, Germany

When participants view stimuli compatible with multiple perceptual interpretations, their perception continuously switches. Multistable perception is often characterized by an average dominance phase duration, particularly, for studies that compare perception of different groups (patients versus healthy controls, old versus young adults, etc.). Here, we asked a question of how reliable this measure is and how variable it is across sessions compared to within session. We tested this by recruiting 31 participants over 3-5 days reporting on five bistable stimuli (2 versions of kinetic-depth effect that were identical in appearance but differed in rotation axis, Necker cube, moving plaid, and auditory streaming), three-minute blocks, twice per session in random order. The design means that for a given stimulus we can pick a pair of sessions, each two blocks long, and make a direct comparison for statistics of our choice within session (consistency of first and second block within each session A and B) versus between (consistency of first blocks between sessions A and B and same for second blocks). Using this approach, we compared block pairs within and between sessions based on average phase duration and state-dominance index using correlation, average phase differences, variance, and consistency in participant and stimuli order measures. First, we found no systematic changes over sessions. Second, we found that for most statistics variance between sessions was comparable to variance within session, so that reliability is not compromised by measuring over multiple days. Third, critically, variability of average dominance phase duration was so great even within session that it did not allow to reliably differentiate between participants or stimuli (within participant). Taken together, our results argue for caution when using average dominance phase duration as a measure of difference between participants.

Talk 3, 3:00 pm, 34.23

FUNCTIONAL PROCESSING ASYMMETRIES BETWEEN NASAL AND TEMPORAL HEMIFIELDS DURING INTEROCULAR CONFLICT

Chris L.E. Paffen[1] (c.l.e.paffen@uu.nl), Surya Gayet[1]; [1]Utrecht University

We recently reported that targets presented to the nasal (i.e., inner) visual hemifield of a single eye have a processing advantage over targets presented to the temporal (i.e., outer) hemifield during continuous flash suppression (bCFS; Sahakian et al., 2022). We speculated that this nasal advantage benefits natural vision, by prioritizing fixated objects in the nasal hemifield of one eye over (nearby) occluders in the temporal hemifield of the other. We investigated this using an interocular grouping paradigm, whereby image A was presented to the nasal hemifields (i.e., right side of the left eye, left side of the right eye), while image B was presented to the temporal hemifields (i.e., left side of the left eye, right side of the right eye). The images were either blurred or sharp, to mimic the difference in focus between near and far objects in real-world vision. Observers continuously reported perceiving image A or B (indicating interocular grouping of either the nasal or temporal hemifields), or a mixture of both (indicating perceptual dominance of one eye's image). We found more grouping (1) for sharp than for blurred images, and (surprisingly) (2) for temporal than for nasal ones. Applying bCFS in the same participants, however, replicated the nasal advantage from our earlier work. Thus, temporal parts of an image were grouped more than nasal ones, while nasal targets broke suppression faster than temporal ones. We suggest that these discrepant results reflect distinct monocular occlusion conditions in natural viewing: we reason that a nasal hemifield advantage is adaptive when an occluder is in the temporal hemifield of a single eye (e.g., when a fixated object passes behind an occluder), while a temporal hemifield advantage is preferred when an occluder is in the nasal hemifields of both eyes (e.g., a person blocking your central view at the cinema).

Talk 4, 3:15 pm, 34.24

INTEGRATING VISUAL INPUT FROM BOTH EYES: BINOCULAR RETINOTOPIC ORGANIZATION BASED ON MONOCULAR INPUT

Abdalla Z Mohamed[1] (azm9155@nyu.edu), Omnia Hassanin[1,2], Rania Ezzo[1], Alessio Fracasso[3], Jonathan A Winawer[4], Bas Rokers[1,4,5,6]; [1]New York University Abu Dhabi, Abu Dhabi, UAE, [2]Vilcek Institute of Graduate Biomedical Sciences, New York University Grossman School of Medicine, NY, USA, [3]School of Psychology and Neuroscience, University of Glasgow, Hillhead Street 62, Glasgow, G12 8QE5, Scotland, UK, [4]Department of Psychology and Center for Neural Science, New York University, NY, USA, [5]NYUAD Research Institute, New York University Abu Dhabi, *Abu Dhabi, UAE, [6]ASPIRE Precision Medicine Research Institute, Abu Dhabi, UAE*

Introduction: Our brain combines visual information from both eyes to create a coherent representation of the visual world. Prior work suggests BOLD response amplitude is increased under binocular input, but population receptive field (pRF) size is unchanged. Here we further investigated pRF properties under monocular and binocular stimulation across the visual hierarchy. Methods: Fifteen healthy participants (9 males, age = 28.7 ± 10.7 years) completed nine 5-minute functional MRI (fMRI) retinotopic mapping scans. Mapping stimuli consisted of dynamic, colorful textures windowed by bar-, wedge-, or ring-shaped apertures, within a circular viewing field (7.5º radius), with participants maintaining fixation. The stimuli were presented either monocularly or binocularly using a VPixx projector and polarizing lenses. PRF models were solved using Vistasoft (https://github.com/vistalab/vistasoft). We compared pRF properties, including amplitude and size, between binocular and monocular conditions. Results: Binocular stimuli elicited greater response amplitude than monocular stimuli, but response amplitude was much smaller than predicted by linear summation of the two monocular responses. The amplitudes were 15.7 ± 8.9% greater for binocular than monocular responses in V1 (mean±std across subjects) (PFDR<0.001), 7.2 ± 6.8% in V2 (PFDR=0.004), 6.7 ± 6.1% in V3 (PFDR=0.003), 7.6 ± 7.5% in hV4 (PFDR=0.004), 9.2 ± 10.3% in TO1 (PFDR=0.004). The pRF size was also greater in V1 (29.9 ± 35.3%, PFDR=0.026) for binocular than for monocular stimuli. No significant differences in pRF size were observed between binocular and monocular stimuli in the extrastriate areas. Conclusion: The binocular responses exhibited subadditive summation, consistent with normalization by a shared pool of neurons. We speculate that the larger pRF size for binocular inputs reflects slight mismatches in the centers of the monocular receptive fields represented in a voxel. These mismatches would have the largest impact on pRF size where pRFs are small, namely V1, consistent with our observations.

This research is funded by the ASPIRE Precision Medicine Research Institute Abu Dhabi award grant number VRI-20-10. NYUAD Center for Brain and Health, funded by Tamkeen under NYU Abu Dhabi Research Institute grant CG012. NIH grants R01EY027401 and NIH R01 EY033628 supported JW.

Talk 5, 3:30 pm, 34.25

THE BEST STEREOACUITY MAY NOT BE AT THE FOVEA

Preeti Verghese[1], Ângela Gomes Tomaz, Adrien Chopin, Dennis Levi; [1]Smith Kettlewell Eye Research Institue, [2]UC Berkeley, [3]Smith Kettlewell Eye Research Institue, [4]UC Berkeley

Classic studies have shown that sensitivity to stereoscopic disparity declines with eccentricity from the fovea (Blakemore, 1970; Cummings & DeAngelis, 2001). We set out to examine whether stereoacuity was indeed highest at the fovea. We measured local stereo sensitivity in 19 controls and 8 amblyopic participants at the fovea and along the horizontal and vertical meridians at 2.5°, 5° and 10° eccentricity. Participants performed a front/back judgment on a square patch whose disparity varied adaptively. Fixation was monitored with eye tracking of the dominant eye. Results showed that the best locus for stereopsis was often not at the fovea. This was true for 13 of the 19 controls, 70% of whom had sensitivities an order of magnitude worse at the fovea compared to the peripheral locus with best sensitivity. The eccentricity of the “best” peripheral locus was: 2.5° (n=6), 5°, (n=4), and 10° (n=3), with a tendency to be in the lower visual field and for near disparities. The 6 controls whose stereo sensitivity was best at the fovea tended to have very good stereoacuity (<12 arcsec); those with best loci outside the fovea tended to have worse stereoacuity (172 ± 60 arc sec). For amblyopic participants, 2 of 3 anisometropic participants had their best stereoacuity at the fovea; the others (1 anisometropic, 3 mixed, and 2 strabismic amblyopes) showed no detectable stereopsis at the fovea, but measurable stereopsis in the periphery. Furthermore, there was a strong correlation between measures of clinical stereoacuity (Randot and Asteroid) and psychophysical stereoacuities at the best locus, whether foveal or peripheral. Taken together, our results indicate that the locus of best stereopsis is often not at the fovea and that clinical measures of stereopsis are correlated with best stereoacuity, whether this occurs at the fovea or in the periphery.

NIH R01 EY034370

Talk 6, 3:45 pm, 34.26

EXOCENTRIC INFORMATION INFLUENCES EGOCENTRIC DISTANCE ESTIMATION IN PERCEPTION AND ACTION

Chaeeun Lim[1], Dhanraj Vishwanath[2], Fulvio Domini[1]; [1]Brown University, [2]University of St Andrews

Previous research suggests that estimating the absolute distance of an object from the observer (egocentric information) and the object depth (exocentric information) relies on different cues, each associated with distinct biases and roles in guiding action. Ocular vergence is a primary cue to absolute distance in near space, while depth is specified by exocentric cues such as relative disparity and texture. Although exocentric cues alone do not specify absolute distance, we found evidence that they integrate with vergence signals, influencing absolute distance estimates. In a series of experiments involving perception and action tasks, participants viewed a stereoscopic paraboloid protruding toward them and either reached for its front or back (reaching), grasped it front to back (grasping), or compared the front and back locations to a reference 2D object (perceptual adjustment). We varied object distance (near vs. far) such that the back of the near object (Near-Back) physically aligned with the front of the far object (Far-Front). If distance perception depended solely on vergence, Near-Back and Far-Front would correctly appear equidistant. However, if depth-from-disparity, known to be overestimated in near space, is integrated with vergence-specified distance then Near-Back would be perceived as farther than Far-Front. In all reaching, grasping, and perceptual adjustment tasks, Near-Back was consistently perceived as farther than Far-Front, despite being equidistant from the observer. This discrepancy intensified when monocular depth cues were added to specify object depth. Remarkably, the discrepancy persisted even when the front and back locations were marked by a pair of isolated stereoscopic dots, suggesting that the relative disparity between the two dots was sufficient to affect their perceived location without a continuous surface. These findings demonstrate that exocentric cues affect distance perception, even when they are not directly relevant, raising intriguing questions about how the visual system selects and integrates information in a scene.

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

Talk 7, 4:00 pm, 34.27

MONOVISION-INDUCED MISPERCEPTION OF MOTION IN GENERAL AND PRESBYOPIC POPULATIONS

Callista Dyer[1], Victor Rodriguez-Lopez[2], Johannes Burge[1,3,4];
[1]Department of Psychology, University of Pennsylvania, PA,
[2]Institute of Optics, Spanish National Research Council, IO-CSIC,
[3]Neuroscience Graduate Group, University of Pennsylvania, PA,
[4]Bioengineering Graduate Group, University of Pennsylvania, PA

Monovision is a common prescription lens correction for presbyopia that focuses one eye at far distances and the other at near distances. A recent study showed that monovision can induce a variant of the classic Pulfrich effect: a visual illusion that causes dramatic misperceptions of the depths of moving objects. The variant—the reverse Pulfrich effect—arises because the discrepant optical lens powers, and consequent interocular differences in blur, cause an interocular processing delay: blurry images are processed milliseconds faster than sharp images. Shockingly, asynchronies of only 2ms can cause multimeter misperceptions of depth, especially when the viewed object is fast-moving (driving). However, the illusion has been demonstrated in only a small number of individuals to date. To determine the scientific generality and potential clinical significance of these findings, it is important to establish the pervasiveness of monovision-induced depth misperceptions. Here, we measure the prevalence of monovision-induced Pulfrich illusions in much larger samples of the general (n=45) and presbyopic (n=17) populations. The stimulus included two sets of horizontally moving bars. One set of bars was positioned directly above the other, and the two sets were moving in opposite directions. Interocular differences in blur, mimicking monovision corrections, were induced with trial lenses (1.0D differences) and with matched onscreen Gaussian blurring. The task was to report which set of bars appeared closer in depth. Psychometric functions were measured. Binocular disparity was the independent variable. The reverse Pulfrich effect occurs in 84% of the general population (lenses: mean=-1.49ms, SD=2.47ms; Gaussian-blur: mean=-1.27ms, SD=1.53ms) and in 100% of presbyopic subjects (lenses: mean=-1.86ms, SD=1.73; Gaussian-blur: mean=-3.64ms, SD=2.86ms). The classic Pulfrich effect, which we also measured, occurred in 89% of the general population and 94% of presbyopic subjects. The reverse Pulfrich effect is reliably induced by optical power differences smaller than the 1.5D difference typically prescribed to presbyopes.

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.

Talk Session: Sunday, May 18, 2025, 5:15 – 7:15 pm, Talk Room 1

Attention: Neural mechanisms

Moderator: James Herman, University of Pittsburgh

Talk 1, 5:15 pm, 35.11

SHARPER SPATIALLY-TUNED NEURAL ACTIVITY IN PREPARATORY OVERT THAN IN COVERT ATTENTION

Damian Koevoet[1], Vicky Voet[1], Edward Awh[2], Henry M. Jones[2], Christoph Strauch[1], Stefan Van der Stigchel[1]; [1]Utrecht University, [2]The University of Chicago

Attention is shifted with or without an accompanying saccade (overtly or covertly, respectively). The neural signatures of overt and covert attention largely overlap, and have even been deemed identical. However, by definition the neural signatures of overt and covert attention must diverge at some point (i.e. saccade initiation), but it remains unclear when and how they diverge. Here, we capitalized on the high temporal resolution of electroencephalography (EEG) in combination with multivariate decoding to investigate when and how overt and covert attention differ neurally. Neural decoding reliably predicted whether overt or covert attention was shifted well before saccade onset (~700ms). We then used an inverted encoding model to compare spatially-tuned neural responses to the attended location between overt and covert shifts. Strikingly, we observed that overt shifts caused sharper spatially-tuned neural responses compared with covert shifts. But why were these spatially-tuned neural responses sharper: did overt attention employ more of the same attention or does imminent saccade execution recruit an additional spatially-tuned process? To address this, we reconstructed spatially-tuned responses when training on only one of the two conditions. We found overt and covert attention to only partly employ similar spatially-tuned responses, arguing against a ‘more-of-the-same attention’ account. Our results instead demonstrate overt attention to recruit an additional spatially-tuned process. We speculate that this additional spatially-tuned process is related to predictive remapping across saccadic eye movements. Together, we demonstrate the neural signatures of overt and covert attention to diverge rapidly because overt attention employs an additional process which sharpens spatially-tuned neural activity.

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).

Talk 2, 5:30 pm, 35.12

CHANGES IN NEURAL TUNING RATHER THAN NOISE STRUCTURE EXPLAIN ATTENTIONAL ENHANCEMENT OF POPULATION REPRESENTATIONS IN THE HUMAN VISUAL CORTEX

Yu-Qi You[1] (yyq117711@gmail.com), Kendrick Kay[2], Ru-Yuan Zhang[1]; [1]Shanghai Jiao Tong University, [2]University of Minnesota

Attention is thought to enhance behavior by refining neural population representations of visual stimuli. Past research has proposed two competing theories to explain these effects: attention either changes the tuning of individual units (i.e., tuning-change theory) or alters trial-by-trial noise correlations between units (i.e., correlation-change theory). However, there is currently no unified framework to accommodate or quantitatively compare these two theories. Leveraging linear Fisher information from computational neuroscience, we first analytically demonstrate that, in addition to tuning and correlation changes, changes in response variability across trials emerge as a third factor that influences population representations. To assess which mechanisms the brain employs, we conducted an fMRI experiment (3T; 2.5mm³) with eight human subjects performing two tasks: attending to digits at center-of-gaze (i.e., a fixation task) or attending to face stimuli randomly presented at one of 16 positions in a 4 × 4 grid (2° spacing) (i.e., a face task). We recorded 80 trials per position for each task and analyzed BOLD responses across regions of the human visual cortex (V1, V2, V3, hV4, OFA, FFA-1, FFA-2). Consistent with previous human imaging studies, compared to the fixation task, the face task systematically altered voxel receptive fields (vRFs) and improved multivariate decoding of face positions. In line with electrophysiological studies, the face task also reduced voxel Fano factors and noise correlations. Most importantly, we developed a neural population geometry approach to quantify the contributions of tuning, noise correlations, and response variability to population representations. Surprisingly, changes in vRFs were the primary driver of improved population representations in both low- and high-level visual areas, while the other two factors had little or even detrimental effects. These findings strongly support the tuning-change theory of visual attention and provide new insights into how attention enhances neural representations to optimize stimulus coding.

Talk 3, 5:45 pm, 35.13

PRONOUNCED MODULATION OF ACTIVITY IN PRIMATE EARLY VISUAL CORTEX BY INTERNAL STATE

Bharath Chandra Talluri[1], Incheol Kang[1], Jacob L. Yates[2], Daniel A. Butts[3], Hendrikje Nienborg[1]; [1]Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, USA, [2]Herbert Wertheim School of Optometry and Vision Science, University of California, Berkeley, CA, USA, [3]Department of Biology and Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD, USA

Internal states of an organism can have a profound influence on visual processing, exemplified by strong modulation due to non-retinal factors like body movements and arousal reported in several species. Recent work found that spontaneous body movements only minimally modulate early visual processing in primates. However, the degree to which general arousal modulates early visual processing in awake macaques is unknown. Here, we recorded extracellular spiking activity from populations of neurons in early visual areas (V1 & V2) of macaques, while manipulating behavioral states in two conditions that required different levels of task engagement: by rewarding the animal contingent on maintaining central fixation in one set of trials (fixation blocks) or rewarding the animal at unpredictable times while free-viewing with no explicit task (free-viewing blocks). Both conditions contained repeated presentations of full-field flashes, such that the visual input to neuronal receptive fields was invariant to the animal's gaze position. The blocks were interleaved, and we monitored physiological correlates of arousal, such as pupil size. The stimulus-response amplitude of the neurons showed a pronounced increase (93±5%) in fixation vs free-viewing blocks, and this modulation could not be explained by differences in retinal input, or saccade statistics between blocks. A computational model that had no knowledge of the conditions but allowed for trial-to-trial variations in shared gain and baseline firing across the neuronal population explained the spiking activity better than a model that allowed separate, fixed stimulus-driven responses in each condition. The shared gain and offset signals were different between conditions and correlated with pupil size suggesting that the observed modulations in neuronal activity reflect modulations in the internal state of the animals. Our findings show that modulation of early visual processing in non-human primates by general arousal can be pronounced, in stark contrast to the minimal modulation by body movements.

National Institutes of Health (grant no. R00EY032179); National Science Foundation (NCS-FO 2123568); National Eye Institute Intramural Research Program at the National Institutes of Health (1ZIAEY000570-01)

Talk 4, 6:00 pm, 35.14

HUMAN PULVINAR STIMULATION ENGAGES SELECT CORTICAL PATHWAYS

Jordan A. Bilderbeek[1], Nicholas M. Gregg[2], Maria Guadalupe Yanez-Ramos[1], Harvey Huang[3], Morgan N. Montoya[1], Peter Brunner[4,5], Jon T. Willie[4,5], Jamie J. Van Gompel[6], Gregory A. Worrel[2], Kai J. Miller[6], Dora Hermes[1]; [1]Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester MN, [2]Department of Neurology, Mayo Clinic, Rochester, MN, [3]Medical Scientist Training Program, Mayo Clinic, Rochester, MN, [4]Department of Neurosurgery, Washington University School of Medicine, St. Louis, MO, [5]National Center for Adaptive Neurotechnologies, St. Louis, MO, [6]Department of Neurosurgery, Mayo Clinic, Rochester, MN

The pulvinar is a large associative thalamic nucleus and plays a role in attentional regulation and modulation of visual pathways. Clinical studies in epilepsy often implant electrodes in the pulvinar to treat seizure networks involving the occipital and temporal neocortex. However, the placement of electrodes within the pulvinar is uncertain as anatomical studies have revealed that pulvinar subnuclei are connected to distinct occipital, temporal, and parietal visual areas. Single-pulse electrical stimulation of such pulvinar subnuclei during multi-lead stereotactic EEG (sEEG) provides an opportunity to causally test whether electrical stimulation of pulvinar subnuclei selectively engages distinct anatomical pathways. We delivered biphasic single-pulse electrical stimulation within the medial and lateral pulvinar in nine neurosurgical patients undergoing clinical sEEG for drug-resistant epilepsy evaluation. We analyzed pulse-evoked potentials (PEPs) using a parameterization algorithm that computed the explained variance as a metric of response reliability for each trial. Stimulation in the medial and ventromedial pulvinar elicited reliable PEPs in the temporal neocortex that diminish as stimulation proximity increases toward the lateral pulvinar. Conversely, stimulation of the lateral pulvinar produced reliable striate/prestriate (V1-V2) and extrastriate (V3a/b, hV4, TO1-2, LO1-2, IPS0) PEPs, which diminishes as stimulation proximity increases towards the medial pulvinar. We also found that dorsomedial pulvinar stimulation evokes parietal (IPS1-5) PEPs with relatively limited striate/extrastriate and temporal responses. Altogether, our results highlight that stimulation of specific pulvinar subfields evoke reliable responses in the striate/extrastriate cortex, parietal cortex, and temporal neocortex (Figure 1). Identifying these pulvinar-cortical projection fields is crucial for understanding how the pulvinar modulates neural activity in visual pathways. Moreover, it is a significant step towards the clinical advancement of seizure network-specific deep brain stimulation in epilepsy.

Funding was provided by the National Institute of Mental Health Award Number R01 MH122258. The content is solely the responsibility of the authors and does not represent the official views of the NIH.

Talk 5, 6:15 pm, 35.15

A COMMON PULVINAR-CORTICAL ARCHITECTURE ACROSS VISUAL TASKS

Xingyu Liu[1] (liuxingyu987@gmail.com), Michael J. Arcaro[1];
[1]University of Pennsylvania

Through its widespread connections with cortex, the pulvinar nucleus within the primate thalamus plays a crucial role in visual and higher cognitive processing. However, our understanding of how the pulvinar's internal organization and connectivity with distinct cortical networks support vision remains limited. Using fMRI, we characterized the pulvinar's functional connectivity "fingerprints" (fFC) — distinct patterns of correlated activity between the pulvinar and specific cortical networks. During task-free, resting state, we identified a topographic organization within the pulvinar characterized by primary axes spanning the entire visual cortical hierarchy as well as non-visual association cortices. To test whether this organization reflects a common architecture for visual processing, we analyzed pulvinar-cortical fFC across diverse visual tasks including naturalistic movie viewing and controlled visual attention paradigms. Remarkably, cortical activation patterns during these tasks could be reconstructed from pulvinar activity using the resting-state fFC patterns, suggesting a stable underlying architecture. This stability was further supported by analyses showing that the topographic organization of pulvinar-cortical fFC was preserved across states. In contrast to the stable pattern of fFC across tasks, the strength of specific pulvinar-cortical connections was flexibly modulated by task demands — naturalistic vision enhanced fFC with visual cortical areas, while attentional tasks strengthened coupling with dorsal attention and cognitive control networks. Moreover, the disruption of this architecture during anesthesia and restoration upon recovery suggests that this stable architecture may serve as a hallmark of consciousness. These findings reveal a fundamental principle of thalamic organization: the pulvinar maintains a stable architectural framework while dynamically adjusting the strength of specific cortical connections based on visual processing demands. This work provides new insights into how the primate visual system achieves both stability and flexibility in visual processing through thalamo-cortical interactions.

This research was supported by National Institute of Mental Health (P50 MH132642)

Talk 6, 6:30 pm, 35.16

INDEPENDENT ENCODING OF SALIENCE, VALUE, AND ATTENTION IN PRIMATE SUPERIOR COLLICULUS

Matthew Murawski[1], James Herman[1]; [1]University of Pittsburgh

Neuronal activity in the primate superior colliculus (SC) is modulated by physical salience, reward value, and attention, but how these signals are integrated remains unknown. One hypothesis is that SC activity reflects a unified "priority map" of the visual field. However, it is unclear how a unified priority map might support distinct roles for SC activity in different contexts: If the same SC neurons are activated by saccade cues and attention cues, how can SC activity evoke a saccade in one context but covert orienting of attention in another context? We hypothesized that distinct sources of modulation cause dissociable rather than unified patterns of SC activation, which would facilitate a context-specific relationship between SC activity and behavior. To test this hypothesis, we recorded 220 SC neurons in a macaque performing two tasks: a spatially cued covert change detection task manipulating goal-directed attention, and a saccade task independently varying reward value and salience. All three factors influenced behavior: salience affected saccadic endpoint error, reward modulated reaction time, and attention cueing affected detection rate. Single-neuron ROC areas for each factor suggested that salience, reward, and attention exert independent influences on SC. At the population level, linear classifiers trained to decode conditions for one factor (e.g., reward) failed to generalize to other factors (e.g., salience), indicating independent population-level encoding. These findings challenge the unified priority map model and suggest SC could support flexible visually guided behaviors by selective routing. This mechanism could explain how SC contributes to overt orienting and covert attention depending on context, advancing our understanding of how the brain flexibly processes visual information to guide behavior.

This work was supported by the Hillman Foundation and the Eye and Ear Foundation of Pittsburgh.

Talk 7, 6:45 pm, 35.17

GOAL-DIRECTED VISUAL INFORMATION PROCESSING WITH GLUTAMATERGIC EXCITATION AND GABAERGIC INHIBITION IN POSTERIOR PARIETAL CORTEX

Sebastian Frank[1], Sinah Wiborg[1], Antonia Wittmann[1], Nina Beck[1], Markus Becker[1], Zhiyan Wang[1]; [1]University of Regensburg

Goal-directed visual information processing involves selecting relevant among irrelevant visual signals. This selection is facilitated if there are separate and stable representations of goal-relevant and goal-irrelevant visual information. Posterior parietal cortex is crucially involved in maintaining such separate representations but it is debated how they are implemented on a neuronal level. The sharpest separation would be achieved by representing goal-relevant information through increased excitatory activity and goal-irrelevant information simultaneously through increased inhibitory activity in different subpopulation of neurons within posterior parietal cortex. If this is the case, increased demands on maintaining separate representations of goal-relevant and goal-irrelevant visual information should be accompanied by a concomitant increase of both excitatory and inhibitory activity. Testing this prediction is difficult with functional magnetic resonance imaging because the contributions of excitatory and inhibitory activity to the hemodynamic response cannot be separated. Here, we measured the concentrations of glutamate, a chief excitatory neurotransmitter, and gamma-aminobutyric acid (GABA), a chief inhibitory neurotransmitter, in the posterior parietal lobe using time-resolved functional magnetic resonance spectroscopy (fMRS). During fMRS participants (n=30) performed a multiple object tracking task with low and high demands on maintaining separate representations of goal-relevant moving targets among goal-irrelevant moving distractors. Tracking trials were 12s long and followed by a jittered inter-trial interval. Changes in glutamate and GABA concentrations with different tracking conditions were measured as a time-series of consecutive 2s-long PRESS and 3s-long MEGA-PRESS fMRS-scans. The results showed greater concentrations of glutamate and GABA in the high than low demand tracking condition in the posterior parietal lobe. No such simultaneous increase of glutamate and GABA concentrations between tracking conditions was found in the occipital lobe. Our results suggest that a simultaneous increase in excitatory and inhibitory activity in posterior parietal cortex is involved in maintaining sharply separated representations of goal-relevant and goal-irrelevant visual information.

Deutsche Forschungsgemeinschaft (DFG): Emmy Noether Grant (Project Number 491290285); Julitta und Richard Müller Stiftung

Talk 8, 7:00 pm, 35.18

ATTENTIONAL EYE SELECTION AFFECTS NEURAL DYNAMICS OF BINOCULAR RIVALRY ACROSS VISUAL HIERARCHY

Chuan Hou[1] (chuanhou@ski.org), Junxian Rao[2]; [1]Smith-Kettlewell Eye Research Institute

Attention plays a role in binocular rivalry. However, previous studies have mostly focused on feature-based attention (Mitchell et al., 2004) or directing attention away from the rival stimuli (Zhang et al., 2011). Few studies (Ooi and He, 1999, 2005) have examined whether attentional eye selection affects the dynamics of rivalry, with the underlying neural basis remaining unclear. Here, we investigated how attentional eye selection affects neural dynamics of binocular rivalry across the visual hierarchy by presenting a pair of orthogonal gratings tagged with different temporal frequencies in each eye (horizontal at F1 = 7.5 Hz for the non-dominant eye; vertical at F2 = 6 Hz for the dominant eye) to observers with normal vision. Neural activity during binocular rivalry was recorded using fMRI source-localized high-density EEG to identify regions of interest (ROIs). We correlated neural responses with behavioral reports to examine activity during perceptual dominance and suppression phases and then compared the responses between two conditions: passive viewing vs. paying attention to the horizontal gratings presented to the non-dominant eye, in various ROIs. Our findings showed that neural activity biased toward the stimuli of the attended eye across ROIs. The primary visual cortex (V1) and extrastriate visual areas, including hV4, middle temporal (MT) and lateral occipital cortex (LOC), exhibited a similar attention effect (approximately a 50% increase in response to the attended stimuli compared to the passive viewing). In contrast, V3A, intraparietal sulcus (IPS), temporal pole, and frontal pole showed approximately a 80% attention effect. Our results provided evidence that attentional eye selection can modulate neural dynamics during binocular rivalry across the visual hierarchy, with more pronounced effects observed in the high-level cortices. These findings are consistent with behavioral studies (Ooi and He, 1999, 2005) and also align with the attention literature, demonstrating that neural activity is enhanced by attention.

This research was supported by NIH grant R01EY035346 to C.H.

Talk Session: Sunday, May 18, 2025, 5:15 – 7:15 pm, Talk Room 2

Color, Light and Materials: Cones to cognition

Moderator: Karl Gegenfurtner, Giessen University

Talk 1, 5:15 pm, 35.21

UNIFORM DISTRIBUTION OF SPECTRAL TUNING IN THE MOUSE EARLY VISUAL SYSTEM

Juan Santiago Moreno[1,2,3,4], Daniel Denman[1,2,3,4]; [1]University of Colorado Anschutz Medical Campus, [2]Medical Scientist Training Program, [3]Neuroscience Graduate Program, [4]Department of Physiology and Biophysics

Like other dichromats, mice expresses two wavelength-sensitive cone opsin proteins: a medium (M) opsin centered in the green region of the light spectrum and a short (S) opsin shifted in sensitivity toward the ultraviolet (UV) region. However, these opsins are not uniquely expressed by dedicated cone cells, nor are they expressed uniformly across the retina, yet they still support color discrimination. A variety of potential mechanisms for color opponency have been shown in the mouse retina, but few studies have captured distribution of spectral tuning downstream in lateral geniculate nucleus (LGN) or the primary visual cortex (V1) populations, nor the relationship between the two. Here, we used high-density electrophysiology to record from LGN (n = 888 units) and all layers of V1 (n = 2729 units) during the presentation of chromatic, achromatic, spatially structured, and uniform stimuli (n = 8 mice, 8 recordings). We defined spectral tuning based on the response-weighted average of M and S-opsin contrasts presented, providing a distribution in M-S contrast space that represents both color opponency (M-S) and luminance preference (M+S). Additionally, we developed a selectivity index that labels neurons as either color preferring (negative) or luminance preferring (positive). As opposed to a model where color and luminance are processed by distinct circuits, single neuron spectral tuning in LGN and V1 formed uniform distributions in contrast space, with their selectivity indices skewing towards luminance. We find an expansion in spectral tuning in V1 compared to LGN (LGN: range = [-0.184, 0.259], skew = 0.687; V1: range = [-0.328,0.621], skew = 1.488), suggesting a de novo origin of color information in cortical processing. Our findings provide a framework for color vision, where spectral tuning in single cells multiplexes color and luminance, and populations collectively encode the full breadth of the mouse visual spectrum without specialized parallel circuits.

NEI R00EY028612, NEI R00EY028612-S1, 1F30EY034775

Talk 2, 5:30 pm, 35.22

VARIATION OF SMALL SPOT COLOR APPEARANCE WITH LOCAL L/M RATIO

Maxwell Greene[1], Vimal Pandiyan[2], Ramkumar Sabesan[2], William Tuten[1]; [1]University of California, Berkeley, [2]University of Washington

Suprathreshold color appearance is surprisingly robust to considerable variations in L/M cone ratio between trichromatic observers. Previous studies suggest chromatic sensitivity may be reduced in individuals with biased cone proportions (Gunther & Dobkins, 2002; Hood et al, 2006). Due to technical limitations, these studies relied on indirect L/M ratio measurements. Additionally, variation of color perception with L/M ratio was assessed between subjects, making it hard to disentangle the effect of cone demographics from other individual differences. Building on this earlier work, we examined how local variations in cone spectral topography influenced responses to small, brief (67 ms) increments of red (680 nm) or green (543 nm) light in two male trichromats whose cone mosaics had been spectrally classified by optoretinography. Flashes subtending 2.25 arcmin (covering ~5 cones) were presented through an adaptive optics scanning laser ophthalmoscope to targeted retinal loci ~2° from fixation. On each trial, the subject indicated whether the stimulus appeared red, green, or achromatic, or whether it went unseen. Stimulus intensities were concentrated near threshold, with occasional high intensities to gauge suprathreshold color appearance. Overall, achromatic percepts predominated (73.9% of seen trials). We fit a generalized linear mixed-effects model to examine how the the probability of categorizing the 543 nm stimulus as “green” or the 680 nm stimulus as “red” depended on i) stimulus intensity (normalized by sensitivity) and ii) the spectral heterogeneity (i.e., proximity to a 1:1 cone ratio) of the stimulated retinal locus. As expected, the likelihood of responding “red” or “green” increased with intensity. Interestingly, we found that the ability to categorize stimuli under the expected hue name improved significantly as the L/M ratio approached unity. Hence, at a fine spatiotemporal scale, the arrangement of cones imposes a limit on the postreceptoral mechanisms subserving color appearance.

Air Force Office of Scientific Research (FA9550-20-1-0195, FA9550-21-1-0230), the National Institutes of Health (R01EY02359, T32EY007043, U01EY032055, P30EY001730), Research to Prevent Blindness Unrestricted grant, the Hellman Fellows Fund, and the Alcon Research Institute.

Talk 3, 5:45 pm, 35.23

NEURAL MECHANISMS OF COLOR SATURATION

Robert Shapley[1], Valerie Nunez[2], James Gordon[3]; [1]Center for Neural Science, New York University, [2]Center for Neural Science, New York University (now at Albert Einstein College of Medicine), [3]Hunter College, City University of New York

Color saturation is a quantitative estimate of how colorful something looks. In an attempt to understand the neural mechanisms of color appearance, we measured perceived color saturation in human observers with hue and saturation scaling, as used by Gordon et al. (1994). The color targets were equiluminant color-gray checkerboards (spatial frequency 2-3 c/deg) presented on a calibrated OLED monitor. The colors lay along the two cardinal axes of DKL color space: “Red/Green” (L-M/M-L) and “Blue/Yellow” (+S/-S). Five or six cone contrasts were used for each stimulus ranging from 0-10% for the L-M/M-L stimuli and from 0-50% for the +S/-S stimuli. The observers estimated saturation as the percentage of the entire sensation, chromatic and achromatic, that was chromatic. To compare color perception with activity in early visual cortex, we also measured the chromatic visual evoked potential (cVEP) over the same range of cone contrast and for stimulus patterns similar to those used in the behavioral experiments. The main results are: 1) saturation varies with the magnitude of cone contrast, and therefore equiluminant complementary colors appear equally saturated; 2) the slope of the saturation vs cone contrast line is 6-8 X shallower for the +S/-S stimuli than for L-M/M-L; 3) cVEP amplitude's dependence on cone contrast resembled the scaling data. It was interesting that, for many observers, cVEPs to +S and -S were approximately equal in amplitude when stimuli were equated for cone contrast magnitude. This result is somewhat surprising because the neurons that carry -S (“Yellow”) signals in the LGN might be expected to respond weakly to the checkerboard stimuli. It raises the possibility that there is a cortical contribution to -S signals. The results also suggest that saturation is likely to be a result of integration of color-evoked responses over the entire population of neurons in early visual cortex.

Our research was supported by a grant from the US National Science Foundation

Talk 4, 6:00 pm, 35.24

NEURAL GAIN COMPENSATION FOR COLOR DEFICIENCIES: EVIDENCE FROM COLOR CONTRAST ADAPTATION

Fatemeh Charkhtab Basim[1], Arsiak Ishaq[2], Erin Goddard[2], Michael A. Webster[1]; [1]Graduate Program in Integrative Neuroscience, University of Nevada, Reno, NV, USA, [2]School of Psychology, University of New South Wales, Sydney, Australia

Anomalous trichromacy (AT) is characterized by a reduced separation between the spectral peaks of the normal medium (M) and long (L) wavelength cone photopigments, resulting in a reduced LvsM comparison signal. Several studies suggest that the color vision of ATs is compensated for this sensitivity loss, so that their color perception is more similar to color-normal (CN) observers. However, it is often difficult to distinguish whether the compensation observed is sensory (e.g., neural gain) or post-perceptual (e.g., learning). To directly probe sensory gains, we compared contrast adaptation in 11 color-normal (CN) and AT observers (10 deutan and 5 protan), by measuring LvsM contrast thresholds before or after adapting to 1 Hz modulations of LvsM contrast for 120 sec. Adaptation increases the thresholds, and the losses increase with the adapting contrast. In ATs, LvsM sensitivity is lower, but adaptation should also be weaker, because the adapt contrast is a lower multiple of threshold. This predicts that the change in thresholds (post/pre adapt) should be weaker for ATs. The predicted

relation was quantified by measuring threshold changes in CNs after rescaling stimulus contrasts by varying amounts to simulate different levels of sensitivity loss. All deutan observers exhibited stronger adaptation than their threshold sensitivity losses predicted, consistent with partial (but incomplete) compensation for their reduced LvsM separation. The results for 4 protan observers were instead consistent with no compensation (with a 5th outlier in the direction of over-compensation). Chromatic contrast adaptation is thought to have an early cortical locus. The results thus provide direct psychophysical evidence for compensatory neural gain adjustments at or before early visual cortex that may partially offset the altered cone signals in some anomalous trichromats, and reinforce neuroimaging (fMRI) measures which have also implicated a cortical site of compensation for color deficiencies (Tregillus et al 2021).

Funding: EY010834

Talk 5, 6:15 pm, 35.25

PROLONGED FIXATION DURATIONS IN COLOR-DEFICIENT OBSERVERS REFLECT A GENERALIZED VISUO-MOTOR ADAPTATION

Doris I. Braun[1] (doris.braun@psychol.uni-giessen.de), Karl R. Gegenfurtner[1]; [1]Giessen University

About 300 million men worldwide have inherited color vision deficiencies (CVD). While their difficulties in distinguishing shades of red and green are well documented, much less is known about the broader aspects of their visuo-motor behavior. We examined eye movements of 27 color-normal and 24 color-deficient observers of different types and severities. They viewed high-quality digital reproductions of still-life paintings, presented both in their original colors and in grayscale. We measured saccade counts and amplitudes, fixation positions and durations, explored image areas (spread), scan paths, fixation heatmaps, and the chromatic properties of fixated regions. Color-deficient participants exhibited an average increase of about 15 ms in fixation duration compared to color-normal observers. Notably, this increase also appeared when viewing grayscale images, suggesting it is not solely related to color discrimination. The effect was especially pronounced for the very first fixation. Despite this increase in fixation duration, we observed remarkably little difference in the overall spatial distribution of fixations or the color distributions of fixated regions. This may reflect the distinct boundaries of objects within these artworks, allowing both groups to identify salient features similarly. Preliminary data from natural image viewing also show prolonged fixation durations, especially for initial fixations, while no such differences emerge during reading tasks. These findings suggest that color-deficient observers may have adapted a slightly different fixation strategy to cope with their challenges in processing visual information. Unlike previous accounts that focused on specific perceptual deficits, our results indicate a more general alteration in visuo-motor behavior.

ERC AdG Color 3.0 and DFG SFB/TRR 135 Cardinal mechanisms of perception.

Talk 6, 6:30 pm, 35.26

SPARSE CODING OF CHROMATIC NATURAL IMAGES RECOVERS UNIVERSALS IN COLOR NAMING AND UNIQUE HUES

Alexander Belsten[1,2] (belsten@berkeley.edu), Paxon Frady[3], Bruno Olshausen[1,2,4]; [1]Redwood Center for Theoretical Neuroscience, University of California, Berkeley, [2]Herbert Wertheim School of Optometry and Vision Science, University of California, Berkeley, [3]Neuromorphic Computing Lab, Intel, [4]Helen Wills Neuroscience Institute, University of California, Berkeley

Understanding the transformation from cone activations to color appearance is a central question in vision research. Here, we study the color statistics of natural images and propose a model of this transformation based on principles of efficient coding. We show that this model replicates universals in color naming from the World Color Survey (WCS) (Kay and Cook, 2023) and aligns with unique hues (Hering, 1878). Whether or not the unique hues have a privileged status in perception has been subject to debate (Conway et al., 2023), and our model provides additional data and computational analyses that help address this question. Using simulated long-, medium-, and short-wavelength (LMS) cone activations in response to natural scenes, we compute a decorrelating transform that spheres the activations (i.e., achieves unit variance in all directions). Utilizing overcomplete, non-negative sparse coding models, we derive optimal bases for representing the data distribution and analyze their tuning properties. The decorrelating transform aligns with the DKL color space (Derrington, Krauskopf, Lennie, 1984). The data distribution in this space exhibits asymmetrical, heavy-tailed structure within the chromatic plane, but not along the luminance axis. Adapting sparse coding models to this structure recovers bases that align with universal color categories identified by the WCS. Notably, the six basis-vector model aligns with the unique hues and replicates features of psychophysical characterizations of these hues, such as mutual exclusivity between opponent colors (e.g., red and green). This work demonstrates that a model based on principles of efficient coding can provide an account for the physiology of color processing and the psychology of color appearance. Specifically, sparse coding models adapted to the sparse structure of the LMS distribution recover bases aligned with universal color categories and the unique hues, supporting previous reasoning that the unique hues are a well-suited basis for describing color appearance.

AFOSR FA9550-21-1-0230 and NSF IIS231-3149

Talk 7, 6:45 pm, 35.27

EDGE-BASED IMAGE RECONSTRUCTION PROVIDES A UNIFIED ACCOUNT OF (MANY) LIGHTNESS ILLUSIONS

Srijani Saha[1] (srijanisaha@g.harvard.edu), George Alvarez[1,2], Talia Konkle[1,2]; [1]Harvard University, [2]Kempner Institute for the Study of Natural and Artificial Intelligence at Harvard University

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 computational goals 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 also require a recovered scene structure. Our work examines whether an edge-based reconstruction goal produces a range of lightness illusions. First, we trained a reconstructive U-Net to output a filled-in image from edge-only inputs of images, an objective analogous to filling-in surfaces from edge-selective neurons in the biological visual system. This model not only reconstructed images with minimal error, but also made systematic errors consistent with lightness illusions measured in people for both the Craik-O'Brien-Cornsweet illusion and the Anderson-Winawer illusion. This effect was robust across training parameter choices (32 combined variations between training datasets and model seeds) and illusion probe choices (contrast signal of edges). When the model was applied to a suite of additional lightness illusions (e.g., Adelson Haze Illusion, Snake Illusion, Koffka Illusions, and Kanizsa Square Illusion), we found that the model consistently recapitulated illusions when there are connected edges with consistent polarity bounding the illusory surface. When the same U-Net model architecture was trained with a different reconstructive goal – denoising different levels of Gaussian noise – the models did not recapitulate any illusions, indicating that edge-based reconstruction is critical and provides a plausible mechanism underlying many perceptual lightness illusions.

This work was supported by NSF PAC COMP-COG 1946308 to GAA.

Talk 8, 7:00 pm, 35.28

GENERALIZATION OF COLOR-CONCEPT ASSOCIATIONS MIMICS COLOR CATEGORY STRUCTURE

Melissa A. Schoenlein[1], Karen B. Schloss[2]; [1]High Point University, [2]University of Wisconsin-Madison

People learn associations between colors and concepts from experiences. Evidence suggests that color-concept associations for colors seen in the input generalize to unseen colors, and this generalization pattern depends on the distance between seen and unseen colors and color category membership—typicality and boundary location (Schoenlein & Schloss, 2022). The current study aimed to build a framework characterizing how association distributions are formed by evaluating how these factors combine. First, we used an associative learning paradigm to expose participants to co-occurrences between colors (yellows/pinks) and concepts (Filk/Slub alien species, respectively). Participants saw aliens in either prototypical (saturated) or non-prototypical (desaturated) yellows/pinks (between-subjects). Then, we assessed color-concept associations for seen colors and generalization to unseen colors. Participants rated how much they associated each species name (Filk/Slub) with colors from sequences spanning pink-to-orange and yellow-to-green, including colors that varied in typicality of color category membership. We constructed a model to predict these associations from: (1) color distance from the seen colors (CIE ΔE), (2) color typicality judgments, and (3) category boundary judgements (all judgments made by different groups of participants). Color-concept associations were correlated with each factor in isolation, but in the model, only typicality ($p<.001$) and color distance ($p<.001$) were significant predictors (no effect of category membership; $p=.898$). Effects of category boundaries were superseded by typicality, implying that associations extrapolated to other colors inside the category boundary shared by the seen colors, but more so for colors that were more typical of the category. Further analysis indicated that these results could not be understood merely in terms of perceptual similarity. This work demonstrates that color-concept association generalization mimics color category structure. These findings increase our knowledge of how color-concept association distributions are populated from sparse input, furthering our understanding of human judgments in visual cognition that rely on associations.

This work was funded by NSF grant BCS-1945303 to KBS

Talk Session: Monday, May 19, 2025, 8:15 – 9:45 am, Talk Room 1

Action

Moderator: Yala Mohsenzadeh, Western University

Talk 1, 8:15 am, 41.11

DECODING TOOL ACTIONS REGARDLESS OF THE OBSERVED ACTING BODY PART

Kyungji Moon[1] (kyungji.moon@georgetown.edu), Florencia Martinez-Addiego[1], Yuqi Liu[1,2], Maximilian Riesenhuber[1,3], Jody C. Culham[4,5], Ella Striem-Amit[1]; [1]Georgetown University Medical Center, [2]Institute of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Sciences and Intelligence Technology, Chinese Academy of Sciences, [3]Center for Neuroengineering, Georgetown University, [4]Department of Psychology, University of Western Ontario, [5]Brain and Mind at Western, Western Interdisciplinary Research Building, University of Western Ontario

What are the mechanisms for understanding tool-use actions? To what extent do we generalize across various parameters, such as acting body part, in the observed actions? Here, we tested if there is a shared neural substrate for observed tool-use actions regardless of if they are performed with the hand or with the foot. We leveraged functional neuroimaging in typically-developed controls and individuals born without hands to understand whether there are shared representations for tool-use, regardless of the observed executing body part, and regardless of one's motor experience. fMRI data from control subjects (n=18) and people born without hands (n=7) were collected while participants passively viewed complex and simple tool-use actions performed with either the hand or foot. We found shared neural responses across the observed body part for both typically-developed individuals and individuals born without hands, which suggested that observed tool-use actions are represented independently of both observed body part and personal sensorimotor experience. Specifically, univariate analyses revealed a consistent preference for action-type (simple or complex tool-use) regardless of observed executing body part in the left superior parietal lobe (SPL) for both typically-developed individuals and people born without hands. Further, multi-voxel pattern analysis successfully discriminated between observed simple and complex tool-use actions consistently across the body part in the bilateral SPL, inferior parietal lobe (IPL), and lateral occipitotemporal cortex (LOTC). Together the results suggest there are shared neural substrates for action understanding regardless of observed body part that are consistently differentiated in typically-developed individuals and people born without hands. This supports generalization across body parts in action perception and implies that motor experience is not necessary for core action understanding.

This work was supported by the Edwin H. Richard and Elisabeth Richard von Matsch Distinguished Professorship in Neurological Diseases (to E.S.A.).

TALK 2, 8:30 AM, 41.12

WORKING MEMORY IMPAIRMENTS CONTRIBUTE TOWARD POOR VISUALLY GUIDED REACHING IN SCHIZOPHRENIA

Jose Reynoso[1,2] (jreynoso@ur.rochester.edu), Maya Glasman[1], Duje Tadin[1,2,3], Brian Keane[1,2,3]; [1]Department of Brain and Cognitive Sciences, University of Rochester, [2]Center for Visual Science, University of Rochester, [3]Unversity of Rochester Medical Center

Common actions such as reaching for a bottle require a robust sense of proprioception and/or working memory (WM). Vision aids these movements, and improperly integrating visual signals leads to inaccuracies in reaching. In schizophrenia, it is unclear to what extent vision interacts with proprioception and WM when executing a reach, although impairments in visual perception, motor function, and visual WM have been studied. We hypothesize that patients will show significantly lower multisensory enhancement, and reduced reaching accuracy due to a memory delay. We used a virtual reality paradigm to assess reaching accuracy guided primarily by proprioception, WM, or vision. A total of 10 patients and 12 age- and sex-matched healthy controls were tasked with reaching out and tapping a virtual target with the index finger. To assess the role of proprioception, in half of the trials (randomly selected), the hand would become invisible at the start of a trial. To assess the role of WM, the target would become invisible shortly after appearing at the start of a trial, followed by a 1 second delay before initiating a reach. Groups did not differ in overall accuracy in proprioception-guided reaching or memory-guided reaching (both ps > .2). The multisensory enhancement gained by using both vision and proprioception to guide a reach was marginally worse in patients (one-tailed $p = 0.095$, $d = 0.58$). However, adding a memory delay worsened patients' accuracy more than controls' (one-tailed $p = 0.032$, $d = 0.84$). While data collection is still ongoing, these data suggest that impaired WM and perhaps multisensory integration contribute to poor dexterity in patients. We further showcase impairment of visual WM in schizophrenia within the context of action. Further work may also elucidate interactions between vision and proprioception (or lack thereof) and help further characterize deficits in these domains within psychotic disorders.

TALK 3, 8:45 AM, 41.13

IMPLICIT ADAPTATION'S EFFECT ON MOTOR AWARENESS AND CONFIDENCE

Marissa H. Evans[1], Jordan A. Taylor[2], Michael S. Landy[1]; [1]New York University, [2]Princeton University

Sensorimotor adaptation is essential for maintaining movement accuracy; it seeks to minimize the effects of external perturbations. It can occur explicitly, by adjusting the intended motor plan to overcome task errors, or implicitly, by automatically and incrementally calibrating the sensorimotor mapping (while the motor plan remains stable). While explicit adaptation has been shown to reduce sensorimotor confidence (confidence in the success of a motor action with a sensory-directed goal), it remains unknown if the operation of implicit adaptation also affects confidence in one's motor awareness. Participants made a slicing reach through a visual target with an unseen hand. We provided "error-clamped feedback": visual feedback of radial hand position moved forward with the hand but went in a fixed direction independent of hand position, which participants were instructed to ignore. Error-clamp direction varied over trials (square wave, amplitude +/-6 deg, 12 cycles/session, plus zero-mean noise per trial, range +/- 6 deg). They indicated perceived hand direction and reported confidence by adjusting the length of an arc centered on the indicated endpoint direction, with larger arcs reflecting lower confidence. Points were awarded if the arc encompassed the true reach direction; fewer points for larger arcs. This incentivized attentive confidence report and minimizing direction-report error. A leaderboard was presented every 50 trials. No other feedback was provided. A significant 12 cycle/block Fourier component in reach direction provides strong evidence for implicit adaptation from the error clamp. The same frequency component was also obtained for indication of reach direction. Thus, while adaptation is unconscious, a mismatch is created between the motor plan and proprioceptive signal, resulting in a judgment of an unsuccessful reach. However, confidence judgments of motor awareness were not affected by adaptation, indicating that sensorimotor confidence and confidence in one's own proprioceptive measurements are calculated differently.

Funding: NIH EY08266

TALK 4, 9:00 AM, 41.14

LEVERAGING PUPIL DIAMETER TO TRACK EXPLICIT CONTROL PROCESSES IN VISUOMOTOR ADAPTATION

Sean R O'Bryan[1], Joo-Hyun Song[1]; [1]Brown University

Visuomotor adaptation (VMA) enables us to recalibrate our sensorimotor mappings to overcome unexpected perturbations. While VMA has been traditionally characterized as an implicit process driven by sensory prediction error, recent work suggests that explicit cognitive control (e.g., working memory) is involved in many VMA tasks, where learning is achieved through a dynamic interplay of implicit adaptation and effortful, explicit strategies to minimize error. However, current approaches to measure explicit VMA (e.g., reported aiming direction) are cumbersome. As an alternative, we predicted that task-evoked pupil diameter (PD) could provide an index of explicit control during VMA. Across three experiments, participants reached to visual targets

while the direction of a cursor was unexpectedly rotated 45° relative to the hand or target. To dissociate explicit and implicit learning, we provided continuous (mixed; N = 30), delayed (explicit; N = 28), or error-clamped feedback (implicit; N = 30). For both mixed and explicit tasks, we found that PD rapidly increased in response to perturbations, consistent with the expectation that PD may track effortful, explicit processes recruited to reduce high initial error. Moreover, PD was significantly associated with individual differences in adaptation, where high performers exhibited larger task-evoked responses. In contrast, for the implicit task—which yields large sensory prediction errors that cannot be controlled—PD was insensitive to both the onset of the perturbation and to individual differences in adaptation. Collectively, our results point to PD as a promising tool to study the interplay of explicit and implicit learning mechanisms in VMA.

This work was supported by NSF BCS-2043328

Talk 5, 9:15 am, 41.15

THE EFFECTS OF SENSORIMOTOR UNCERTAINTY DURING NATURAL LOCOMOTION IN EARLY CHILDHOOD

Sara E Schroer[1], Mary M Hayhoe[1]; [1]University of Texas at Austin

While walking, vision is used to efficiently gather information to navigate, avoid obstacles, maintain balance, and find footholds. Adults flexibly adapt gaze allocation and gait when walking on challenging terrains, suggesting a high-level control of walking that accounts for visual information, energetic costs, and stability. Although locomotion develops through early childhood, how children use vision to control their walking has been minimally studied. Children 2- to 6-years-old and adults wore a head-mounted eye tracker while they walked on two natural terrains – a sidewalk and loose stones. Motor uncertainty is higher on the stones because each potential foothold is wobbly, increasing motor noise. In response to the increased uncertainty, walkers modified their behavior in three ways: more fixations were directed towards the ground, gaze was closer to their body, and speed decreased. On the sidewalk, an average of 61.2% of fixations were towards the ground – and this increased to 90.5% on the stones. Participants looked further ahead on the ground when walking on the sidewalk (ps<0.001; distance measured as participant's own leg length). Adults looked 3 leg lengths ahead when walking on the sidewalk (approximately 3 steps ahead) and only 2 leg lengths ahead on the stones. By 6-years-old, children's attention on the sidewalk is similar to adults (3 leg lengths ahead), but younger children looked further ahead on the ground (3-4 leg lengths, ps<0.014). Younger children's lookahead distance was also more broadly distributed, whereas older children's and adults' gaze peaked at 3 leg lengths ahead. On the stones, children of all ages look closer to their body (and closer than adults), just 1-2 leg lengths ahead. Lastly, children walked faster on the sidewalk than stones (p=0.027). Our work suggests that locomotion in young children, like adults, is controlled by complex decision mechanisms that take sensorimotor uncertainty into account.

NIH R01 EY05729 and T32 EY021462

Talk 6, 9:30 am, 41.16

AGE-RELATED DIFFERENCES IN GAZE DISTRIBUTION DURING LOCOMOTION: BALANCING SAFETY AND EXPLORATION IN REAL AND VIRTUAL ENVIRONMENTS

Sophie Meissner[1] (sophie.meissner@psychol.uni-giessen.de), Jochen Miksch[2], Sascha Feder[3], Sabine Grimm[2], Wolfgang Einhäuser[2], Jutta Billino[1]; [1]Justus Liebig University Giessen, Experimental Psychology, [2]Chemnitz University of Technology, Physics of Cognition Group, [3]Chemnitz University of Technology, Cognitive Systems Lab

During locomotion, attention guidance must maintain a balance between gathering relevant information from the environment and ensuring stable gait. Gaze allocation is key for this balance of exploration and safety. In particular for older adults, avoiding falls plays a critical role during locomotion. Prioritization of gaze towards the ground has been proposed as a compensatory strategy in older adults to maintain gait stability. In this study, we investigated age-related differences in gaze allocation during locomotion with and without an additional task. We considered locomotion in a real environment and in virtual reality (VR) to evaluate comparability of behavior in both settings. We studied locomotion in younger (N=24, M=26.1 years) and older (N=24, M=68.8 years) adults. Participants traversed a real hallway and a highly realistic virtual version of the same hallway, with or without the additional search task to locate and manipulate small target objects on the wall. Gaze behavior was assessed using mobile eye-tracking glasses and a VR headset, respectively. Our findings show a strong age-related bias in gaze allocation towards the floor that holds in the real world as well as in VR. Older adults focus on gait-related information, putatively stabilizing their postural control, at the cost of exploring their environment. However, when an explicit second task is introduced that requires allocation of attention away from the floor, older adults adapt their gaze behavior so that their gaze allocation patterns appear similar to those of younger adults, potentially deprioritizing safety in favor of exploration. We conclude that task demands during locomotion might critically put gait stability of older adults at risk. Gaze allocation seems similarly modulated in the real and the virtual world, supporting that virtual approaches provide an appropriate proxy to investigate real-world locomotion behavior across the lifespan.

This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), Collaborative Research Centre SFB/TRR 135: Cardinal Mechanisms of Perception, project number 222641018

Talk Session: Monday, May 19, 2025, 8:15 – 9:45 am, Talk Room 2

Theory: Artificial neural networks

Moderator: Michael F. Bonner, Johns Hopkins

Talk 1, 8:15 am, 41.21

MODEL MANIFOLD ANALYSIS SUGGESTS HUMAN VENTRAL VISION IS LESS LIKE AN OPTIMAL CLASSIFIER AND MORE LIKE A FEATURE BANK

Colin Conwell[1] (colinconwell@gmail.com), Michael F. Bonner[1]; [1]Johns Hopkins University

What do deep neural network models of biological vision tell us about the computational structure of the brain? A now common finding in visual computational neuroscience is that many different kinds of deep neural network models -- each with different architectures, tasks, and training diets -- are all comparably good predictors of image-evoked brain activity in the ventral visual cortex. This relative parity of seemingly diverse models may at first seem to undermine the common intuition that we can use these models to infer the key computational principles that govern the visual brain. In this work, we show to the contrary that the relative parity of models along certain dimensions does not preclude the differentiation of those same models on other dimensions -- and in particular, that metrics of manifold geometry may in certain cases reliably predict whether one model will yield better brain-alignment than another. To do this, we assess 21 manifold geometry metrics computed across a diverse set of over 1200 deep neural network models, curated to include multiple tasks, architectures, and input diets. We then use these metrics to predict how well each model aligns with occipitotemporal cortex (OTC) activity from the human fMRI Natural Scenes Dataset. We find that manifold signal-to-noise ratio (a metric previously associated with few-shot learning) is a robust predictor of downstream brain-alignment and supersedes the predictive power both of other manifold geometry metrics (i.e. manifold capacity, effective dimensionality) and of downstream task-performance (e.g. top-k recognition accuracy) across multiple different image sets (e.g. ImageNet21K versus Places365) and model comparison probes (e.g. category-supervised versus self-supervised models). These results add to a growing body of evidence that the ventral visual stream may serve more as a basis set (or feature vocabulary) for object recognition rather than as the actual locus of recognition per se.

TALK 2, 8:30 AM, 41.22

HUMAN VISUAL ROBUSTNESS EMERGES FROM MANIFOLD DISENTANGLEMENT IN THE VENTRAL VISUAL STREAM

Zhenan Shao[1,2,3], Yiqing Zhou[4], Diane M. Beck[1,3]; [1]Department of Psychology, University of Illinois, Urbana-Champaign, [2]Department of Computer Science, University of Illinois, Urbana-Champaign, [3]Beckman Institute, University of Illinois, Urbana-Champaign, [4]Department of Physics, Cornell University

Humans effortlessly navigate the dynamic visual world, yet deep neural networks (DNNs), despite excelling in visual tasks, are surprisingly vulnerable to image perturbations that are innocuous to humans. Aligning DNN representations with human neural representations, particularly those from higher-order regions of the ventral visual stream (VVS), has been shown to improve their robustness (Shao et al., 2024). Such observation suggests that the representational space in the VVS has desirable properties that support human robustness but are absent in DNNs. One particular framework posits that the VVS achieves robust inference by progressively disentangling neural category manifolds (Dicarlo & Cox, 2007). Specifically, neural population responses to identity-preserving changes of objects form continuous manifolds in the neural state space. These manifolds are initially tangled, i.e., linearly inseparable, but become progressively disentangled across stages of the VVS, naturally resulting in robust inference. Despite its theoretical appeal, empirical evidence for this framework remains limited. Here, using a computational characterization of neural manifold statistics (Chung et al., 2018) and a 7T fMRI dataset (Allen et al., 2022), we first demonstrate that category manifolds at different stages of the human VVS show increasingly desirable geometric properties: smaller radius and compressed dimensionality, together leading to improved overall linear separability. Importantly, we show that these properties are inheritable by DNNs through neural representation alignment and indeed predict subsequent robustness gains observed in previous work. Finally, to more directly test this framework, we propose “manifold guidance”, a method that aligns DNNs to human VVS on the granularity of category manifolds, without imposing strict individual representation matching commonly adopted in previous neural alignment studies. We show that manifold guidance is capable of leading to robustness improvements in DNNs. Our findings, thus, provide compelling evidence that human visual robustness arises from the disentanglement of category manifolds in the VVS.

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

FRAMED RSA: HONORING REPRESENTATIONAL GEOMETRY AND REGIONAL-MEAN RESPONSE PREFERENCES

JohnMark Taylor[1] (johnmarkedwardtaylor@gmail.com), Nikolaus Kriegeskorte[1]; [1]Columbia University

Representational similarity analysis (RSA) characterizes the geometry of neural activity patterns elicited by different stimuli while discarding information about neural response preferences, regional-mean activity and the absolute location or orientation of the patterns in the multivariate response space. Regional-mean activation analysis serves the complementary purpose of comparing the average population response to different stimuli. Invariance to certain aspects of the neural code is useful for comparing systems that might use superficially different encodings to implement the same computations. However, quantities such as mean activation are physiologically or mechanistically important, and so we may want for our models to predict them correctly. Here we introduce a novel analysis technique, framed RSA, which honors both the geometry and the regional-mean preferences in evaluating model-predicted representations. To achieve this, we augment the set of patterns that define the geometry by two reference patterns: the zero-point (origin) and a uniform constant pattern in the multivariate response space, enabling RSA to incorporate information about the global location, orientation, and mean activation of neural population codes. First we present the mathematical and methodological underpinnings of framed RSA, including how it interacts with different RSA analysis choices, such as the use of cross-validated dissimilarity estimates or whitened RDM comparators. Second, we show empirically that framed RSA improves accuracy for both brain region identification (using fMRI data from the Natural Scenes Dataset) and deep neural network layer identification relative to existing RSA approaches. Framed RSA thus offers the theoretical virtue of combining the strengths of two complementary and traditionally separate analysis approaches and the practical value of improved power for model-comparative inference.

This work was supported by the National Eye Institute of the NIH under award number 1F32EY033654.

Talk 4, 9:00 am, 41.24

BEYOND ONE-WAY MAPPING: LINKING MODEL-BRAIN ASYMMETRY TO BEHAVIORAL PREDICTIONS IN VISUAL OBJECT RECOGNITION

Sabine Muzellec[1,2,3], Kohitij Kar[3]; [1]Brown University, [2]University of Toulouse, [3]York University

Advancements in artificial neural networks (ANNs) have yielded object recognition models that closely mimic the primate ventral visual pathway. Traditional evaluation metrics focus mainly on how well ANN units predict neural activity, often overlooking the bidirectional nature of this relationship. In this study, we investigate the symmetry of predictive relationships between ANN components and inferior temporal (IT) neurons and explore its implications for aligning computational models with primate behavior. We conducted large-scale neural recordings from 288 sites across the IT cortex in two macaques engaged in 45 binary object discrimination tasks using 1,320 naturalistic images. Human behavioral data were collected from 80 participants, achieving an image-level reliability of 0.89. Our analysis revealed significant asymmetries in the bidirectional predictive relationships between ANN units and neural responses. By employing linear regression and centered kernel alignment (CKA), we tagged two classes of ANN units: "best" units (top 10th percentile explained variance, EV) demonstrated significantly higher CKA values compared to all units ($p<0.0001$) while "worst" units (bottom 10th percentile EV) showed significantly lower CKA values. This asymmetry was consistent across multiple architectures, including Vision Transformer (ViT), ResNet50v2, Inception-v3, and AlexNet. Crucially, we found that the "best" ANN units more accurately predicted both human and macaque object discrimination performance compared to the "worst" units ($p<0.05$, permutation test). This relationship also remained consistent across different object categories. Interestingly, monkey IT neurons identified as "best" units as predicted by other monkey ITs also demonstrated a similar enhanced prediction of human behavior, suggesting potential shared neural mechanisms across species. Taken together, our findings underscore that developing human-like object recognition in ANNs requires optimizing neural prediction accuracy and jointly ensuring representational symmetry with biological systems.

KK was supported by funds from the Canada Research Chair Program, the Canada First Research Excellence Funds (VISTA Program), Google Research Award, and a NSERC-DG. SM was supported by CFREF (VISTA Program).

Talk 5, 9:15 am, 41.25

TOPONETS: ARTIFICIAL NEURAL NETWORK MODELS WITH BRAIN-LIKE TOPOGRAPHY

Mayukh Deb[1,2] (mayukh@gatech.edu), Mainak Deb, Haider Al-Tahan[1,2], N. Apurva Ratan Murty[1,2]; [1]Cognition and Brain Science, School of Psychology, Georgia Tech, Atlanta, GA 30332, [2]Center of Excellence in Computational Cognition, Georgia Tech, Atlanta, GA 30332

The human visual system exhibits topographic organization across multiple scales: micro-scale orientation selective pinwheels, macro-scale faces, bodies, and scene-selective regions, and large-scale biases for real-world size and animacy. Artificial neural networks (ANNs) on the other hand, lack this inherent organization. To address this problem, we introduce TopoLoss, a novel loss function for ANNs inspired by synaptic connectivity pruning mechanisms that sculpt topographic representations in the brain. The resulting models, TopoNets, exhibit brain-like topography while maintaining task performance (on Imagenet). We validated TopoLoss on convolutional and transformer-based ANN architectures (ResNet-18s, ResNet-50s, ViT-b-32), collectively TopoNets. We show that TopoNets outperform all previous topographic models on ImageNet performance while delivering similar (or higher) levels of induced topography. But what are the representational consequences of brain-like topography? We find that inducing topography drives visual representations to be lower dimensional and in-turn more brain-like. This representational change manifests in two key ways (1) improved ability to predict neural data (on BrainScore for example), and (2) in recapitulating key topographic signatures observed in the visual system (category-selective maps for example). Next we used TopoNets to examine why the brain might adopt such a characteristic topographic design. Our findings revealed that TopoNets are exceptionally parameter efficient and exhibit robustness to lesioning (L1 pruning) and downsampling. These features indicate a functional and evolutionary advantage of cortical topography in brains. Given the ubiquity of topography in the cortex, we extended TopoLoss beyond vision to language (NanoGPT, GPT-Neo-125M) and audition models (multi-task models). Remarkably, TopoNets retained high task performance while replicating key signatures of brain-like topography observed in language and auditory cortices. Taken together, TopoLoss offers a simple and flexible way to instill brain-like topography into modern ANNs. TopoNets offer a unique combination of competitive task performance and brain-like topographic organization, heralding a new generation of brain-inspired AI.

This work was funded by the NIH Pathway to Independence Award from the NEI (R00EY032603) and a startup grant from Georgia Tech (to NARM)

Talk 6, 9:30 am, 41.26

FACTORIZED CONVOLUTION MODELS FOR PREDICTING AND INTERPRETING NEURONAL TUNING IN NATURAL IMAGES

Binxu Wang[1,2,3], Carlos Ponce[1,3]; [1]Harvard University, [2]Kempner Institute for the Study of Natural and Artificial Intelligence, [3]Harvard Medical School

Convolutional neural networks have been extensively used to model neurons in the visual systems of primates and rodents. However, this regression is ill-posed because the number of image-response pairs are often far fewer than the feature regressors. To address this, previous approaches used unsupervised feature reduction (e.g. PCA) and penalized regression (e.g. LASSO). However, these solutions discard the spatial structure of feature, usually leading to non-smooth or non-local weight, harming interpretability. We present a supervised feature reduction method that applies tensor factorization to the covariance between image features and neuronal activations. This factorization generates paired spatial masks and feature vectors, offering clearer insights into which features matter and where they are localized. The method matches the accuracy of penalized regression approaches while providing more precise localization of neuronal receptive fields. To validate our method, we performed closed-loop experiments on neurons recorded from V1, V4, and inferotemporal (IT) cortex in two primates. Using the Evolution paradigm, neurons guided image synthesis through generative networks, and the resulting image-response pairs trained a factorized convolution model. The factor structure in these models allowed further manipulation and ablation study. We shuffled different components in these models as control models and synthesized their maximal activating images. In the same recording session, we presented back these synthesized "optimal" images and measured neuronal responses. Synthesized "optimal" images from the factor model were generally more activating than those from controls; by comparison, we identified the necessary components of the factor model for each neuron. With adversarially trained backbones, we found low-rank, localized read-out weights can synthesize the preferred images as effectively as dense weights, while simplifying the images. In this way, we transformed the dense "black-box" model of a neuron into a part-based model, that was easier to describe and manipulate, helping us understand their natural image tuning.

NIH (1DP2EY035176), NSF (CAREER 2143077), Quan Pre-Doctoral Fellowship to B.W., Kempner Post-Doctoral Fellowship to B.W.

Talk Session: Monday, May 19, 2025, 10:45 am – 12:15 pm, Talk Room 1

Eye Movements: Gaze strategies

Moderator: Emily Cooper, UC Berkeley

Talk 1, 10:45 am, 42.11

THE INTERPRETATION ATTRIBUTED TO OBSERVED GAZE SHIFTS AFFECTS THEIR ATTENTION CUEING EFFECT

Shlomit Yuval-Greenberg[1,2] (shlomitgr@tau.ac.il), Amit Zehngut[1]; [1]School of Psychological Sciences, Tel Aviv University, [2]Sagol School of Neuroscience, Tel Aviv University

In environments rich with stimuli, individuals must rely on attentional cues to identify the most relevant targets. As social beings, a key strategy involves observing where others focus their attention and following their lead, under the assumption that areas attended to by others are more likely to be of interest. Since gaze shifts represent a visible indicator of attention, mirroring the gaze of others serves as an effective social-attentional strategy. Indeed, research has shown that observing others' gaze redirection triggers a reflexive shift in attention, improving perceptual performance for objects located at the gazed-at positions. This phenomenon, known as the Gaze Cueing Effect (GCE), is well-documented and often regarded as reflexive. However, in social interactions, gaze shifts do not always signify attentional orientation. For instance, people often avert their gaze during effortful cognitive processing. Correctly interpreting gaze shifts is essential for effective gaze-based attention shifts. Here, we challenge the reflexive nature of the GCE by examining its dependency on gaze interpretation. Across two preregistered experiments (total N = 110), participants watched videos of gaze shifts while performing a perceptual task. The experimental context was manipulated: one group was primed to interpret the gaze shifts as reflecting cognitive processing rather than overt attentional shifts, and the other was not. Results revealed that the GCE was suppressed in the group primed to view gaze shifts as cognitive processing. This finding suggests that the GCE is influenced by the social interpretation of observed gaze shifts. We conclude that the Gaze Cueing Effect is modulated by social context and is not purely reflexive. These findings highlight how social interpretation can significantly shape fundamental attentional mechanisms.

The study was funded by ISF grant 1960/19 to S-Y.G and by BSF grant 2020308 to S-Y.G and Prof. Robert Knight

Talk 2, 11:00 am, 42.12

SCENE VIEWING FROM KINDERGARTEN TO RETIREMENT - LEARNING CANONICAL GAZE

Ben de Haas[1], Marcel Linka[1]; [1]Experimental Psychology, Justus-Liebig-Universität Giessen

Two adults viewing the same scene tend to fixate overlapping parts of it. This has led to decades of computational modeling, trying to predict average fixation densities. However, individual fixation patterns deviate from this in highly reliable ways. How do we get to this point? Do young children fixate scenes in stereotypical ways and acquire individual preferences over time? Or is children's gaze idiosyncratic before becoming more canonical? We present eye-tracking data from >6,500 participants in a museum. Participants ranged from 5-72 years of age, freely viewing 40 complex scenes. This large dataset allowed us to trace the development of individual differences by estimating average pairwise correlations of fixation patterns, separately for age bins of two years. We find that preschool children tend to fixate fewer elements of a scene and agree on those to an appreciable degree. For older children, image exploration and the number of fixated elements rapidly increase. In parallel to this, pairwise similarities steeply drop. The gaze of children becomes increasingly idiosyncratic until age 14. Then, the trend reverses and patterns of gaze become more and more similar despite a continued increase in image exploration. Pairwise similarity only plateaus from the early twenties, at its highest level. These results show that the degree to which adult gaze is canonical takes decades to develop. I will speculate on the reasons for this protracted development and on its potential relationship with scene understanding.

ERC StG INDIVISUAL

TALK 3, 11:15 AM, 42.13

QUANTIFYING SPATIOTEMPORAL GAZE DYNAMICS USING PERCEPTUAL SEGMENTATION

Concetta Brusco[1] (concetta.brusco@einsteinmed.edu), Sophie Molholm[1], Nathaniel Killian[1], Ruben Coen-Cagli[1]; [1]Albert Einstein College of Medicine

Eye movements reveal the spatiotemporal dynamics of visual attention. Image saliency is a well-known predictor of gaze location. However, it has been proposed that spatiotemporal gaze patterns are also organized by the spatial boundaries of visual objects. Here, we test this rigorously with natural images for the first time. We leverage our novel measurements of perceptual segmentation and concurrent eye-tracking to test the hypothesis that subjectively perceived segments predict spatial and temporal dynamics of scanpaths. In each of 35 sessions, participants (n=8) viewed a natural or texture image for a few seconds, then made a series of quick perceptual judgments that we used to algorithmically reconstruct their perceptual segmentation maps (PSMs, i.e. the most probable segment for each pixel). First, we tested if PSMs predict gaze location by calculating the mutual information (MI) between PSMs and gaze density during initial image viewing. We found significant MI for 12/22 natural image sessions ($p<0.05$, two-sided permutation test with 1000 shuffles of gaze data). Comparatively, all sessions had significant MI between gaze and the image's saliency map generated by DeepGazeIIE, the state-of-the-art scanpath prediction model. However, the PSMs had greater information gain than the saliency maps (162% vs 56% increase from shuffled gaze data, average of 12 sessions), suggesting that subjectively perceived segments influence the spatial density of the scanpath. We also investigated temporal dynamics: on average, segments were viewed for 910-msec periods containing 2 consecutive fixations and 2-3 saccades. These values did not depend on segment size but were significantly shorter for natural images than textures. This suggests that semantic features of segments, and not just the spatial area they occupy, influence temporal gaze dynamics. Overall, our findings indicate that perceptual segments are meaningful spatial divisions that impact spatiotemporal gaze and attentional dynamics, and could potentially improve individualized scanpath prediction.

R01EY031166 (RCC), Rose F. Kennedy Intellectual and Developmental Disabilities Research Center, Albert Einstein College of Medicine (RCC)

TALK 4, 11:30 AM, 42.14

EYE MOTION IMPROVES ACUITY UNDER EMULATED CONE LOSS

Hannah K. Doyle[1] (hannahdoyle@berkeley.edu), James Fong[1], Ren Ng[1], Austin Roorda[1]; [1]University of California, Berkeley

Retinal degenerative diseases degrade vision through cone loss. Prior work has probed the visual function of patients with these diseases on measures such as acuity (Ratnam 2013, Foote 2018), contrast sensitivity (Alexander 1992), and motion detection (Turano 1992). However, this area of study is limited by recruitment of patients with such diseases. We have developed a system for performing optical stimulation on a cone-by-cone basis, which we used to emulate cone loss in healthy subjects and study the impact of eye motion on acuity under retinal degeneration. We used an adaptive optics scanning light ophthalmoscope to simultaneously image the retina at 840 nm and deliver cone-by-cone stimulation at 543 nm. We conducted a 4AFC Landolt C task with "cone dropout," where a random percentage of cones were excluded from the stimulation, emulating cone loss that was fixed to the retina as it moved across the stimulus. For a separate "image dropout" condition, we removed pixels fixed to the stimulus itself, preventing the subject from gathering new information about the stimulus through eye motion. First, we measured acuity thresholds for varying percentages of cone and image dropout through an interleaved staircase procedure. Second, we varied stimulus duration and compared performance between the cone and image dropout conditions. For all 3 subjects tested, acuity declined logarithmically with the fraction of cones removed from the mosaic. Above 50% dropout, acuity was improved under the cone dropout condition as compared to image dropout. This benefit was also apparent in experiments varying stimulus duration, where acuity improved with duration under the cone dropout condition and showed an advantage over the image dropout condition, especially at longer durations. Our results show that the visual system makes use of information gathered through eye motion in order to improve its acuity when sampling with a degraded cone mosaic.

Funded by Air Force Office of Scientific Research grants FA9550-20-1-0195 and FA9550-21-1-0230, and National Institutes of Health grant R01EY023591.

TALK 5, 11:45 AM, 42.15

LANGUAGE GUIDED SEARCH WITH MULTI-MODAL FOVEATED DEEP NEURAL NETWORK FINDS OBJECTS AT UNEXPECTED LOCATIONS

Parsa Madinei[1], Miguel P Eckstein[1]; [1]University of California, Santa Barbara

Introduction: Human eye movements during visual search are guided by scene context and object co-occurrence. Search accuracy deteriorates, and response times increase when the target appears at an unexpected location (e.g., a toothbrush on a toilet seat). In everyday life, observers can overcome the detrimental effects of out-of-context target placement when linguistic instructions guide their search: "The toothbrush on the toilet seat". Here, we present a foveated language-guided search model (FLGSM) that combines a multi-modal transformer with foveated architecture (Freeman & Simoncelli, 2011) and a reinforcement learning agent to locate objects in real-world scenes using language instructions. We assess FLGSM's accuracy for images with targets at expected and unexpected locations accompanied by informative language. Methods: The FLGSM uses a transformer to perform visual grounding using referring expressions and foveated feature maps, while an Advantage Actor-Critic agent with a Long-Short Term Memory architecture to select fixation points and optimize detection performance (rather than mimicking human scanpaths; Yang et al., 2020). Results: FLGSM's detection accuracy (Area under the ROC, AUC) was significantly higher when the target appeared at expected (AUC=0.766) vs. unexpected location (AUC=0.597). When the FLGSM received language input about the location of the target, the performance gap between expected and unexpected locations substantially decreased

(AUC=0.722 vs. 0.694, difference=0.028). Conclusions: These results demonstrate that integrating language guidance with foveated visual processing enables more robust object detection, particularly for targets in unexpected locations. Our model's performance suggests that combining linguistic information with strategic eye movements can help overcome the limitations of context-based visual search, more closely matching human visual search capabilities.

Talk 6, 12:00 pm, 42.16

DISTINCT HUMAN EYE FIELDS UNRAVELED BY FMRI VISUAL FIELD AND OCULOMOTOR MAPPING TASKS

Uriel Iascombes[1], Sina Kling[1], Guillaume S Masson[1], Martin Szinte[1]; [1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Université, Marseille, France

Eye movements, primarily saccades and pursuit, are essential to active visual perception and are therefore intricated with visual neural representations. While research in animals has elucidated the roles of the frontal eye fields (FEF) and parietal eye fields (PEF) in controlling these eye movements, the topography of their human homologues remains poorly understood. To comprehensively map the human eye fields, we used advanced neuroimaging and modeling techniques. High-field functional MRI was used to first localize cortical frontal and parietal regions involved in generating saccadic and smooth pursuit eye movements in 20 participants. We then explored the visuospatial characteristics of these regions using population receptive field mapping, identifying visuomotor clusters within the parietal and frontal cortices that exhibits retinotopic organization. Our findings reveal a topographically organized architecture of the human eye fields underlying the sophisticated control of saccades and smooth pursuit. Interestingly, we found differences in cortical magnification patterns between the FEF and PEF clusters, suggesting specialized functional roles. This study challenges the conventional understanding observed in primates, as the human eye fields appear to be composed of distinct visuomotor areas.

This research was supported by an ANR JCJC and a Fyssen Foundation grant to MS.

Talk Session: Monday, May 19, 2025, 10:45 am – 12:15 pm, Talk Room 2

Multisensory Processing

Moderator: Peter Kok, University College London

Talk 1, 10:45 am, 42.21

AUDIOVISUAL INTERACTIONS ALTER THE GAIN OF CONTRAST-RESPONSE FUNCTIONS IN HUMAN VISUAL CORTEX

Minsun Park[1] (vd.mpark@gmail.com), Sam Ling[1]; [1]Boston University

Humans are inherently multisensory, seamlessly integrating information across senses to create a unified perception. Recent studies have challenged the traditional notion that early sensory areas are dedicated to a specific sensory modality such as vision by demonstrating its susceptibility to another sensory modality such as hearing. However, how cross-modal interactions influence early sensory processing remains unclear. In particular, it is unknown whether cross-modal interaction can modulate sensory gain in early sensory cortex. In this study, using functional magnetic resonance imaging (fMRI), we investigated whether audiovisual interactions (AVI) modulate neural activity in early visual areas (V1-V3) by measuring its effects on the population contrast-response function, a fundamental building block in vision. Participants viewed vertical gratings, which moved either leftwards or rightwards, with parametrically varied contrast intensities (9 contrast levels, 3-96%). While viewing these stimuli, they also heard synchronous binaural auditory motion. The direction of auditory motion was either congruent or incongruent with the direction of visual motion, alongside a stationary sound condition. We measured BOLD responses in V1-V3 to examine whether and how AVI alters the gain of the contrast response. Results showed higher contrast sensitivity when the audiovisual motion direction was congruent compared to incongruent or stationary conditions, the pattern of which was observed consistently across V1-V3. These findings demonstrate that AVI improves visual sensitivity by modulating sensory gain in visual cortex, particularly under the condition of the congruent direction of audiovisual motion. This suggests that interactions between vision and hearing can influence sensory computations in visual cortex, traditionally considered specific to visual processing.

This work was funded by the National Research Foundation of Korea (NRF) Grant RS-2024-00407838 to M. Park and by the National Institutes of Health (NIH) Grant R01EY028163 to S. Ling.

Talk 2, 11:00 am, 42.22

AUDITORY SPATIAL CUES INFLUENCE CROSS-MODAL RECRUITMENT OF VISUAL PREFRONTAL CORTEX

Abigail Noyce[1], Wusheng Liang[1], Madhumitha Manjunath[1], Christopher Brown[2], Barbara Shinn-Cunningham[1]; [1]Carnegie Mellon University, [2]University of Pittsburgh

Lateral prefrontal cortex (PFC) contains discrete regions that are preferentially recruited for visual versus auditory attention and working memory (WM), including regions in superior and inferior precentral sulcus and in mid inferior frontal sulcus. During auditory WM for spatial locations, these visual-biased regions are also significantly recruited. This recruitment during spatial auditory cognition may indicate that additional resources are required to buttress audition's poor precision for spatial location (consistent with the multiple-demand account of PFC), or it may reflect task-specific recruitment of vision's cortical machinery for representing space. In order to better understand the role of visual-biased PFC in auditory spatial cognition, we used fMRI to first label these structures in individual subjects (N = 20) using a direct contrast of visual vs. auditory 2-back WM blocks. Then, in an independent task, we estimated the recruitment of visual-biased PFC during auditory spatial WM under 3 different spatial cue conditions. Auditory locations were cued using either a weaker spatial cue (interaural time differences, ITDs), a moderate spatial cue (interaural level differences, ILDs), or a strong spatial cue (head-related transfer functions, HRTFs). The multiple-demand account predicts that visual PFC recruitment will be strongest under ITDs, because this is the case when the most effort is required to represent spatial location. Instead, we observe that visual PFC recruitment is lowest (but still positive) under ITDs, and is highest under HRTFs, suggesting that visual PFC is representing spatial location rather than allocating resources. These results support a task-specific, not multiple-demand, account of visual-biased PFC's cognitive role.

Supported in part by ONR MURI N00014-19-12332

Talk 3, 11:15 am, 42.23

EARLY VISUAL CORTEX ENCODES MULTISENSORY POSTDICTIVE PERCEPTION WITH RETINOTOPIC SPECIFICITY: A LAYER-SPECIFIC FMRI STUDY

Pieter Barkema[1,2] (skgtpwb@ucl.ac.uk), Joost Haarsma[1], Christoph Koenig[1], Peter Kok[1]; [1]Department of Imaging Neuroscience, UCL Queen Square Institute of Neurology, University College London, London, UK, [2]Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

Postdiction is a phenomenon where later incoming information influences how we perceive earlier sensory input. Little is known about the neural mechanisms of postdiction, despite its important role in shaping perception. Accumulating research suggests that neural representations in the early visual cortex (EVC) are not solely determined by bottom-up retinal inputs, but additionally reflect top-down modulations of subjective perception. Bottom-up and top-down signals in EVC are reflected in different cortical layers. Here, we extend this framework to postdictive perception. Inspired by Kok et al. (2016) we hypothesized that neural responses to postdictive illusions in deep cortical layers would reflect perception, and that the effect is retinotopically specific. We used the Audiovisual Rabbit paradigm (Stiles et al., 2018) to induce a postdictive illusory flash using sound. Twenty-four participants were selected for high susceptibility to the illusion and took part in the experiment as well as in retinotopic mapping during 7T functional Magnetic Resonance Imaging (7T fMRI). EVC neural response amplitude reflected retinal input but not perception, with no layer-specific differences. Multivoxel pattern analysis, however, revealed that the EVC activity pattern evoked by an illusory flash was similar to that evoked by a real flash and this effect was retinotopically specific. This effect was stronger in deep layers than middle layers, in line with a top-down effect. These findings extend the emerging role of EVC in perception by, for the first time to our knowledge, implicating it in multisensory postdiction. Moreover, these findings support the emerging view that the amplitude of neural responses in EVC is primarily driven by retinal input, whereas EVC activity patterns reflect subjective perception.

This work was supported by a Wellcome/Royal Society Sir Henry Dale Fellowship [218535/Z/19/Z] and a European Research Council (ERC) Starting Grant [948548] to P.K. The Wellcome Centre for Human Neuroimaging was supported by core funding from the Wellcome Trust [203147/Z/16/Z].

Talk 4, 11:30 am, 42.24

BRAIN REPRESENTATIONS OF NUMEROSITY ACROSS THE SENSES AND PRESENTATION FORMAT

Ying YANG[1] (yyangpsy@gmail.com), Michele Fornaciai[1], Irene Togoli[1], Iqra Shahzad[1], Alice Van Audenhaege[1], Filippo Cerpelloni[1,2], Olivier Collignon[1,3]; [1]Institute of Psychology (IPSY) and Institute of Neuroscience (IoNS), University of Louvain, Belgium, [2]Department of Brain and Cognition, Leuven Brain Institute, KU Leuven, [3]The Sense Innovation and Research Center, HES-SO Valais-Walis (Lausanne and Sion)

Whether it's three pens on a table, three knocks on a door, or three strikes of a hammer on a nail, we can automatically perceive "threeness". Such ability to seamlessly encode numerosity across the senses and presentation formats has led people to assume the existence of an abstract numerical code in our minds. Yet, researches on the existence of such an abstract representation in the human brain have yielded inconsistent results. The current study used multivariate pattern analysis and representational similarity analysis to comprehensively investigate how the brain represents numerosities (range 2-5) across different modalities (auditory, visual) and formats (sequential, simultaneous; symbolic, non-symbolic). We identify a set of dorsal brain regions, from early visual cortex to the intraparietal and frontal regions, that encode specific non-symbolic numerosity across formats and modalities. The numerical distance effect, a hallmark of magnitude encoding, was observed in parietal regions. We also observed aligned representation of numerosities across visual and auditory modalities in the intraparietal and frontal subregions, but only when they shared a sequential presentation format. Further exploration of the unique contributions of modality and format factors to numerosity representation revealed that the distributed numerical activity in lateral intraparietal sulcus (IPS) subregions is mostly influenced by the modality of presentation, the anterior IPS by the format of presentation, while the medial IPS exhibited equal contributions from both factors. Our study provides a detailed description of the geometry of numerosity representation across the senses and formats in the human brain and provides support for abstract numerical representation across the senses when the presentation format is equivalent.

Talk 5, 11:45 am, 42.25

MULTISENSORY CONTINUOUS PSYCHOPHYSICS: HEADING PERCEPTION IS FASTER BUT NOT MORE PRECISE WHEN BOTH SOUND AND VISUAL CUES ARE PRESENT

Bjoern Joerges[1], Jong-Jin Kim[1], Laurence Harris[1]; [1]York University, Toronto

Heading perception is an inherently multisensory phenomenon that can involve, among others, visual and auditory cues. Like in many other tasks, the presence of multisensory over unisensory cues is expected to lead to both higher precision in responses and lower reaction times – findings that are fairly well established in the trial-based tasks that are typical of this area of study. Here, we used a novel paradigm from vision science - continuous psychophysics - to investigate whether such enhancements of multisensory heading perception were found. We immersed 25 participants in a virtual environment in which they either experienced unisensory visual or auditory information consistent with self-motion that continuously changed direction, or consistent visual and auditory information at the same time. They were asked to continuously align a joystick with their direction of motion. Contrary to our expectations, we did not find any differences in precision between the three (auditory, visual, and visuo-auditory) conditions. However, we did find that participants reacted faster to changes in the stimulus in the visuo-auditory condition than in either of the unisensory conditions. While this discrepancy between our results and what has generally been reported in the literature (e.g., Ernst & Banks, 2002) might be the consequence of a speed-accuracy trade-off (Drugowitsch et al., 2014), it underlines the importance of testing long-established findings using novel paradigms in new, diversified contexts. References: Drugowitsch et al. (2014) eLife 3, e03005 Ernst and Banks (2002) Nature 415, 429-33

Talk 6, 12:00 pm, 42.26

EXAMINING THE TIME COURSE OF VISUOPROPRIOCEPTIVE INTEGRATION USING THE MIRROR BOX ILLUSION

Grant Fairchild[1] (grant.t.fairchild@gmail.com), Yuqi Liu[2], Riwa Safa[3], Jared Medina[1,3]; [1]Emory University, [2]Chinese Academy of Sciences, [3]University of Delaware

The mirror box illusion creates a visuoproprioceptive conflict between the actual position of the hand behind the mirror and the reflection of the participant's viewed hand. After synchronous movement or even passive viewing of the hand over time, participants will report feeling their hand where they see it. Although the illusion involves evidence accumulation over time, the relationship between evidence accumulation and perceptual qualia is poorly understood. One possible mechanism is an abrupt state transition in which the visual estimate of hand position overrides the proprioceptive estimate as soon as the accumulated evidence favoring vision surpasses a certain threshold. Alternatively, the illusion may reflect a gradual shift in weighting given to the visual vs. proprioceptive estimates as evidence favoring vision accumulates. To investigate the temporal progression of the mirror box illusion, we recorded observers' perceived hand position every five seconds in a series of experiments manipulating several factors that modulate the congruence between visual, motor, and proprioceptive signals. As expected, we find that the illusion is strengthened by factors increasing the congruence between the hidden and reflection hand, including synchronous movements, decreased scalar or angular distance, and reduced biomechanical constraints. Notably, we find that the illusion can proceed along both abrupt and gradual trajectories: Sometimes, participants perceive an abrupt transition in the perceived position of their hand, as though their hand suddenly snapped into the position of the reflection, while other times, participants perceive a gradual shift in hand position, as though their hand slowly drifted towards the position of the reflection. Furthermore, there appear to be consistent individual differences in observers' tendency to experience the illusion abruptly or gradually. We discuss the implications of these findings for the mechanisms of resolution of visual-proprioceptive conflict.

This project was funded by NSF grant 1632849.

Talk Session: Tuesday, May 20, 2025, 8:15 – 9:45 am, Talk Room 1

Spatial Vision: Neural mechanisms

Moderator: Elisha Merriam, NIMH/NIH

Talk 1, 8:15 am, 51.11

EVOKED RESPONSES MODULATE PERCEPTUAL SENSITIVITY ACROSS VISUAL SPACE OVER TIME CONSISTENT WITH A CORTICAL TRAVELING WAVE.

Zachary Davis[1], Ashley Royston[2], Dylan Jensen[1], Emma Rudolph[1], Akshay Parchure[1]; [1]University of Utah, [2]University of Denver

Sensitivity for faint visual targets is not fixed but varies from moment to moment. One factor that has been proposed to influence visual sensitivity is the state of intrinsic fluctuations in subthreshold cortical activity that can travel as waves across the cortical surface (Davis, Muller, et al., Nature 2020). Sensory and motor events such as visual stimuli and saccades generate traveling waves of cortical activity (Muller et al., Nat. Communi. 2014; Zanos, et al., Neuron 2015) but it is unknown whether these waves also modulate sensitivity like intrinsically generated traveling waves. To test this, we asked healthy adult human subjects to report their detection of faint visual targets while triggering evoked responses (a visual stimulus or an eye movement) that putatively form traveling waves along a foveal-peripheral axis in the visual cortex. The targets appeared randomly at various eccentricities and delays with respect to the timing of the wave-triggering event and were titrated in size and contrast to achieve similar levels of performance. We tracked eye movements as subjects reported their detection of faint visual targets with a mouse button press. Detection performance followed a rhythmic time course with respect to trigger onsets consistent with previous reports (Fiebelkorn et al., Current Biology 2013; Hogendoorn, J. Cog. Neurosci. 2016). When we separated target performance based on target eccentricity, we found that the modulation time course shifted with eccentricity, following a progression consistent with the rhythmic fluctuation traveling as a wave. These results were recapitulated in a model that generated similar rhythmic performance modulations by activity propagation in the human visual cortex across the retinotopic map. This data is consistent with the view that spatiotemporal fluctuations in cortical activity modulate sensory processing and perceptual sensitivity and suggest that evoked activity fluctuations can impact sensitivity similar to intrinsically generated activity.

Talk 2, 8:30 am, 51.12

VARIABILITY IN HEMODYNAMIC RESPONSE FUNCTIONS CAN MASQUERADE AS DIFFERENCES IN RETINOTOPIC SELECTIVITY

D. Samuel Schwarzkopf[1,2] (s.schwarzkopf@auckland.ac.nz), Ecem Altan[1], Catherine Morgan[3,4], Steven Dakin[1,5]; [1]School of Optometry & Vision Science, University of Auckland, New Zealand, [2]Experimental Psychology, University College London, United Kingdom, [3]School of Psychology & Centre for Brain Research, University of Auckland, New Zealand, [4]Centre for Advanced MRI, UniServices Ltd., Auckland, New Zealand, [5]Institute of Ophthalmology, University College London, United Kingdom

Population receptive field (pRF) analysis has become the most popular method for retinotopic mapping with functional magnetic resonance imaging. It requires assumptions about the hemodynamic response function (HRF) to model a voxel's response. Most pRF studies use canonical HRFs based on normative data or individual HRFs estimated via independent, event-related stimulation paradigms. However, the choice of HRF can influence pRF size estimates substantially (Lerma-Usabiaga et al., 2020, PLOS Comp. Biol. 16(6): e1007924). To understand how this might affect results in practice, here we concurrently fit a double-gamma HRF with five free parameters as part of the pRF model. Using simulated data, we first demonstrate that this algorithm accurately recovers different ground truth HRFs used for generating the data. Next, we reanalyzed several empirical pRF datasets collected using different stimulus paradigms (bar sweeps, combined wedge+ring), magnetic field strengths (1.5T, 3T, 7T), and pulse sequences. Concurrent HRF fitting affected pRF size estimates considerably. It typically improved the presumed validity of model fits by reducing the proportion of artifactually small pRFs. Importantly, the best-fitting HRF varied between datasets. It also differed substantially from all canonical HRFs and from HRFs measured independently, indicative of response non-linearities. Finally, HRF shape also differed substantially between early and higher visual cortex. All these differences could theoretically affect the interpretation of reported findings, such as claims about inhibitory modulation or contextual processing. To test this, we conducted further simulations. Even when the ground truth pRF had a circular Gaussian profile, analyzing these data with an inhibitory surround model (difference of Gaussians pRF) produced spurious estimates of surround inhibition. Such errors were greatly exacerbated when assuming a different HRF than the one used for data generation. Our findings therefore show pRF parameters estimated when assuming a fixed HRF should be treated with caution.

Supported by a Research Development Fund allocation from the Faculty of Medical & Health Sciences of the University of Auckland to DSS

Talk 3, 8:45 am, 51.13

DECODING THE ORIENTATION SERIAL DEPENDENCE EFFECTS FROM V1 NEURONAL RESPONSES USING A TRANSFORMER MODEL

Xin Wang[1] (2301110701@stu.pku.edu.cn), Shi-Ming Tang[1], Cong Yu[1]; [1]Peking University

When two lines are presented sequentially, the second line may be perceived as either slightly tilted towards the first one (attraction, influenced by Bayesian priors) or away from it (repulsion, driving by efficient coding), contingent on the experimental conditions. To understand the computational principles underlying these serial dependence effects, we employed two-photon calcium imaging to simultaneously record the responses of >1000 V1 superficial-layer neurons to sequentially presented Gabors embedded in white noise in two awake, fixating macaques. The parafoveal Gabor varied in 12 orientations and 4-5 contrasts (0.03-0.50), while the background noise varied in 5-6 RMS contrasts (0-0.29). The stimulus intervals were 1000-ms long, separated by 1500-ms inter-stimulus intervals. The average neuronal responses displayed maximum repulsion effects around 15 deg orientation differences. Consequently, downstream brain areas must readout V1 responses to generate attraction effects. We developed a transformer model that incorporated two self-attention mechanisms, each designed to identify the most relevant neurons and their effective connections for decoding the orientation of one stimulus line. Additionally, a cross-attention mechanism was included to capture interactions between the outputs of two self-attention mechanisms. The model was trained to reconstruct the actual repulsion effects and hypothesized attraction effects (repulsion effects flipped). A PCA analysis of the cross-attention maps corresponding to the outputs of the repulsion and attraction models revealed distinct distributions of neuronal responses within the top three PC dimensions. Moreover, within each distribution, cross-attention values of trial pairs with the largest effect sizes were clustered in the PCA space. These results suggest that the repulsive and attractive serial dependence effects in orientation perception are likely mediated by distinct and independent neural computations based on the same neuronal responses. Moreover, downstream brain areas can efficiently readout relevant V1 information to produce attraction effects, as indicated by the cross-attention values in the three-dimensional PCA space.

STI2030-Major Projects grant (2022ZD0204600)

Talk 4, 9:00 am, 51.14

CHARACTERIZING POPULATION RECEPTIVE FIELDS IN HUMAN VISUAL CORTEX UNDER WIDE-VIEW STIMULATION

Pei-Yin Chen[1,2] (pychen1116@gmail.com), Atsushi Wada[1,2]; [1]Center for Information and Neural Networks, NICT, Osaka, Japan, [2]Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan

Population receptive field (pRF) analysis is a widely used method for investigating the retinotopic organization of human visual cortex. Yet, conventional fMRI visual presentation methods are typically limited to measuring pRF properties within 20 degrees of visual angle. Here, we utilized wide-view stimulation to explore the pRF properties of human visual cortex in both central and far-peripheral vision. The retinotopic stimuli consisted of rotating wedges and contracting concentric rings with checkerboard textures moving across the visual field. Observers viewed the stimuli through a wide-view binocular visual stimulation system (Wada et al., 2015, OHBM) in the MRI scanner, which allowed the visual presentation to cover around 90 degrees of the visual field. We applied a 2D Gaussian pRF model to the BOLD activations of voxels in areas V1 to V3, hV4, and V3A/B. Our results show a topographical representation of contralateral visual hemifield in all regions of interest, spanning both central and far-peripheral vision. When examining the relationship between the estimated pRF size and the eccentricity of the pRF center for both central and far-peripheral regions, we found that the relationship is better described by a logarithmic regression function rather than a linear regression function employed in previous studies which assessed only central visual regions. In addition, in areas V1 to V3, both the pRF center eccentricity and pRF size increased with the cortical distance from the posterior pole up to a certain level and then began to decrease. The range where the values increase can be captured by an exponential function, reflecting the cortical magnification in each area. On the other hand, the range where the values decrease may reflect potential limitations of the pRF estimation when the true pRF is located near the boundary of the visual presentation which can only be partially stimulated.

JSPS KAKENHI (Japan) 24K16880, 19K12745

Talk 5, 9:15 am, 51.15

VISUAL RESPONSE ADAPTATION DYNAMICS DEPEND ON LUMINANCE POLARITY AND SPATIAL FREQUENCY IN PRIMARY VISUAL CORTEX BUT NOT SUPERIOR COLLICULUS NEURONS

Yue Yu[1], João B. Bittar[1], Ziad M. Hafed[1]; [1]University of Tuebingen

Stimulus-driven neuronal responses are often characterized by an initial spike burst followed by a gradual reduction towards lower steady-state activity. Such temporal adaptation is mathematically equivalent to high pass filtering, suggesting that different adaptation dynamics translate into different abilities to faithfully track time varying stimuli. Here, motivated by the idea that luminance polarity (dark versus bright contrasts) and spatial frequency may not be equally likely at all possible temporal frequencies in natural dynamic scenes, we asked whether primary visual cortex (V1) and superior colliculus (SC) neurons exhibit different adaptation time constants along these feature dimensions. For luminance polarity, we recorded from 408 V1 (two monkeys) and 238 SC (three monkeys) neurons. In each trial, a disc (0.51 deg radius and variable dark or bright Weber contrast) appeared within the neurons' response fields. For spatial frequency, we recorded from 268 V1 and 127 SC neurons (two monkeys) while presenting a static Gabor grating (five spatial frequencies). SC neurons exhibited faster adaptation dynamics than V1 neurons. However, there was neither dependence on luminance polarity nor spatial frequency. In contrast, V1 neurons exhibited clear feature-dependence in their adaptation. For luminance polarity, adaptation was considerably slower for low contrast (10%) bright than dark stimuli. On the other hand, at 100% contrast, adaptation was considerably faster for brights than for darks. As for spatial frequency, V1 neurons always showed the fastest adaptation at intermediate (4 cycles/deg) spatial frequencies. We hypothesize that the faster SC adaptation allows visual-motor SC neurons to quickly recruit saccade-related bursts after the sensory responses. Instead, V1 neurons need to track scene dynamics. For example, clouds are low contrast bright stimuli that have slow temporal dynamics, not necessitating fast neural adaptation. Conversely, fixational eye movements enhance the temporal retinal image modulations of intermediate spatial frequencies, requiring faster V1 dynamics to represent them.

Talk 6, 9:30 am, 51.16

ORIENTATION SELECTIVITY IN MOUSE SUPERIOR COLLICULUS MODELED WITH CENTER-SURROUND RECEPTIVE FIELDS

Austin Kuo[1,2,3] (austinchkuo@gmail.com), Justin Gardner[1], Elisha Merriam[2,3]; [1]Stanford University, [2]National Institutes of Health, [3]National Institute of Mental Health

Can a neural population be selective for properties of a stimulus none of its constituent neurons are selective for? Foundational single-unit physiology experiments show orientation-selective neurons emerging in primate V1 but not in subcortical structures such as LGN or superior colliculus (SC). Accordingly, a single-unit perspective suggests linear readout of stimulus orientation should be possible from V1, but not LGN or SC. We tested whether Ca2+ imaging results of orientation-selective populations from mouse SC necessarily implies SC neurons with V1-like, elongated receptive field (RF) structure, by simulating neural responses from a study (Liang et al., 2023, Nat Commun 14:4756) that measured mouse SC responses to visual stimuli varying in orientation (0, 15, …, 165°), spatial frequency (0.01-0.32 cycles/°), size (radii: 19, 29, 40°), and shape (circle, square, diamond). We simulated V1-like and center-surround RFs at single-unit and population scales, with RF sizes (~50-250 deg2) and spatial frequency selectivity (0.01-0.32 cycles/°) based on prior measurements. Reproducing empirical findings, neural populations simulated from either V1-like or center-surround RFs showed similar orientation preferences dependent on aperture location, shifting from radial to anti-radial orientation preferences as stimulus spatial frequency increased. At the single-unit scale, simulated center-surround, rather than V1-like RFs, matched empirical orientation preferences along stimulus apertures. Our simulations provide a unified explanation of edge-related orientation selectivity, suggesting measured orientation selectivity from SC populations need not imply V1-like RFs, but spatial frequency selectivity in single units can instead confer orientation selectivity to the population. This result, however, does not preclude existence of V1-like RFs in prior empirical data. Rather, our simulations provide a tool for experimenters to determine RF properties sufficient for producing measured activity. Broadly, our results demonstrate that neural populations can exhibit emergent selectivity absent in their constituent neurons, which in principle could be read out by downstream perceptual and motor systems.

ZIAMH002966, Wu Tsai Neurosciences Institute

Talk Session: Tuesday, May 20, 2025, 8:15 – 9:45 am, Talk Room 2

Visual Memory: Imagery, memorability, long-term

Moderator: Wilma Bainbridge, University of Chicago

Talk 1, 8:15 am, 51.21

WHAT MAKES MENTAL IMAGES VIVID? SHARPNESS AS THE KEY VISUAL DIMENSION

Xueyi Huang[1], Angela Shen[2], Emil Olsson[2], Kiarra Michelle I. Garcia[2], Nadine Dijkstra[3], Megan A. K. Peters[2], Jorge Morales[1]; [1]Northeastern University, [2]University of California, Irvine, [3]University College London

Visual mental images vary in their subjective vividness. Conventionally, this vividness has been measured on a unidimensional numerical scale. But how different visual qualities influence our experience of vividness is poorly understood. Here, we present results from a novel method for reconstructing visual properties of mental images as they vary across different subjective vividness ratings. On each trial, subjects saw a line drawing of an object and were instructed to imagine it in as much detail as possible. Then, they rated the vividness of their mental image on a 1-to-5 scale. Finally, a Voronoi tessellation pattern—a tiling of differently-colored shapes—appeared on the screen along with three sliders that controlled its sharpness, opacity, and saturation. Subjects abstracted the qualities of their mental image and applied them to the Voronoi pattern by adjusting the sliders. With a mixed-effects linear model that displayed high goodness of fit, we found that all three dimensions significantly predicted subjects' vividness ratings. Importantly, at the group level, sharpness was the strongest predictor (followed by opacity and then saturation), and it also explained most of the random effects variance. At the subject level, sharpness was the top predictor of trial-by-trial vividness ratings as well. While our results showed individual differences in what visual features influence vividness ratings, the clear primacy of sharpness as the strongest and most common predictor can help guide further research into the visual properties and neural basis of mental imagery. Additionally, our modeling results were not correlated with the overall vividness of subjects' mental imagery, suggesting that the pattern in which people relied on the three visual dimensions was independent of their overall mental imagery capacity. In conclusion, despite the subjective nature of mental imagery, these results show that our method can robustly reconstruct vividness ratings by quantifying the contributions of different visual features.

This work was supported by the Templeton World Charity Foundation, Inc. (funder DOI 501100011730) under the grant https://doi.org/10.54224/22032 to (JM, MAKP & ND) and in part by the Canadian Institute for Advanced Research (CIFAR, Fellowship in Brain, Mind, & Consciousness; to MAKP).

Talk 2, 8:30 am, 51.22

THE ART OF MEMORY—ART CONTEST REVEALS THAT ARTISTS CAN CREATE MEMORABLE ARTWORKS, PREDICTED BY AI

Trent M Davis[1] (trentdavis@uchicago.edu), Yifei Chen[1], Wilma A Bainbridge[1]; [1]University of Chicago

Why do we remember some artworks but not others? Previous research has shown that paintings have an intrinsic memorability, where the same pieces are consistently remembered and forgotten across viewers (Davis & Bainbridge, 2023). Importantly, the memorability of famous paintings was significantly higher than non-famous paintings as predicted by ResMem, a neural network designed to predict the memorability of images without any outside information or context (Needell & Bainbridge, 2022). This was especially surprising given previous work showing that people are inconsistent in their ability to predict an image's memorability (Revsine & Bainbridge, 2022). Inspired by these findings, we organized a national art contest with the following question in mind: Can artists intentionally create memorable and forgettable artwork? 87 artists from across the United States submitted to the contest, with about two thirds attempting to make memorable artwork and the remaining third, forgettable. Pieces were judged via a continuous recognition memory experiment (N=131), where we found entries to the memorable category were significantly more memorable than the forgettable category, suggesting that artists are able to manipulate the memorability of their artwork. We then created a hybrid art exhibit and in-person experiment where the 20 most memorable and forgettable artworks were displayed for a month at Connect Gallery in Chicago, IL. Viewers of the exhibit were invited to test their memory of the artworks (N=61), and we also tested people's memories at a satellite event at VSS 2024 (N=140). Visitors in both samples remembered memorable pieces significantly better than forgettable ones ($p<0.001$), demonstrating artists were able to manipulate memory. Further, ResMem was significantly able to predict people's memory performance both online and in-person ($p<0.001$). These results suggest that memory for artwork is manipulable and predictable, which has major implications for the design of both artwork and museums.

Talk 3, 8:45 am, 51.23

SINGLE NEURONS IN HUMAN MTL TRACK THE DEPTH-OF-PROCESSING ELICITED BY VISUAL REPRESENTATIONS OF IMAGES

Aalap Shah[1] (aalap.shah@yale.edu), Yuchang Tian[1], Qi Lin[2], Runnan Cao[3], Shuo Wang[3], Ilker Yildirim[1]; [1]Yale University, [2]RIKEN, [3]Washington University in St. Louis

Decades of research have pointed to the human Medial Temporal Lobe (MTL) as the locus interfacing spontaneous visual processing and memory formation. The hippocampus, for instance, receives input from the ventral visual stream and encodes episodic memories. Yet, the algorithmic basis of how stimulus-driven visual processing modulates activity across MTL structures remains unclear. Here, we address this gap by leveraging computational techniques to model a unique dataset of single-cell recordings from the hippocampus (n=362 neurons) and amygdala (n=446 neurons) of humans (n=15) passively viewing real-world images. We hypothesize that a core psychological theory, the 'depth-of-processing' of visual inputs (Craik and Lockhart, 1972), explains neural processing across MTL structures. A recent study (Lin et al., 2024) introduced an image-computable signature of depth-of-processing: compression-based reconstruction error of visual representations obtained by training a linear sparse autoencoding model on the activations of a pre-trained convolutional neural network. However, the stimuli used in our neural dataset included several basic-level categories (e.g., animals, airplanes, buildings etc.) that proved challenging for the linear model to compress effectively. Therefore, we developed a higher-capacity, non-linear sparse autoencoding model trained on a large-scale, multi-category image dataset, ImageNet. We found that the same computational signature of Lin et al. (2024)—compression-based reconstruction error—positively correlated with the firing rates of hippocampus singe-cells under this more powerful model. No alternative model demonstrated this relationship with the hippocampus, including distinctiveness of visual representations and reconstruction error under a non-sparse variant of our model. Moreover, given the predominantly inhibitory connections from the hippocampus to amygdala; we predicted—and subsequently confirmed—a statistically significant, but notably negative correlation between reconstruction error and the firing rates of amygdala single-cells. Only our model captured this dissociative relationship across the two MTL structures: hippocampus and amygdala. These results suggest depth-of-processing as an algorithm-level account of stimulus-driven visual processing in the MTL.

TALK 4, 9:00 AM, 51.24

MEMORABILITY BEYOND SEMANTIC FEATURES: PROBING MEMORABILITY OF SEMANTICALLY SIMILAR IMAGES THROUGH GENERATIVE DIFFUSION MODEL

Hyewon Willow Han[1,2], Yalda Mohsenzadeh[1,2]; [1]Western Center for the Brain and Mind, Western University, London, Ontario, Canada, [2]Vector Institute for AI, Toronto, Ontario, Canada

What makes an image more memorable? Recent findings suggest that semantic features exert a stronger influence than perceptual features on the memorability of object images (Kramer et al., (2023)). However, semantically similar images can still exhibit a diverse range of memorability. What features of an image make it more memorable among semantically identical images? To address this question, we propose a novel framework that generates cloned stimuli from original images, replicating both the image features and the memorability of the originals. We utilize the THINGS dataset (Hebart et al., (2023)), a naturalistic image collection comprising diverse kinds of object concepts and their associated memorability scores. First, we develop a memorability predictor model tailored to the THINGS dataset using CLIP (Radford et al., (2021)) and leverage recent advancements in large-scale generative diffusion models (Rombach et al., (2022)) to generate clones of these images. We demonstrate that the cloned images successfully preserved image attributes and features from the originals and exhibited comparable memorability scores. Subsequently, we further investigate what makes an image more memorable with the most forgettable and most memorable cloned images by analyzing groups of artificial neurons responsible for specific image features through the decomposition of image representations (Gandelsman et al., (2024)). Our findings reveal that cloned images not only successfully replicate the memorability of their originals but also exhibit variability in their memorability. By analyzing these variations, we gain deeper insights into perceptual features that drive image memory in addition to semantic features. This study highlights the potential of generative diffusion models to explore the cognitive and computational attributes underlying image memorability.

TALK 5, 9:15 AM, 51.25

LIMITED INFLUENCE OF SLEEP ON DRAWINGS MADE FROM MEMORY

Samuel R. Rosenthal[1], Emma Megla[1], Wilma A. Bainbridge[1]; [1]Department of Psychology, University of Chicago

It is well known that memory worsens the longer it has been since encoding (Ebbinghaus 1885), and that memory is better after a period of sleep than the same period awake (Ellenbogen et al., 2006). While this is widely known to be true for recognition-based visual memory (Wagner et al., 2007), it is less clear for visual recall. To assess the effect of sleep on visual recall, we conducted two experiments. First, Prolific participants (N=188) encoded 4 scene images before undergoing a 10-hour delay, during which they were either awake or asleep. After the delay, the participants drew the 4 images from memory. Participants also encoded and immediately drew 4 scene images, to replicate prior effects without a delay (Megla et al., 2024). Following this, separate Prolific participants (N=393) rated which objects were present in the drawings. As expected, participants recalled significantly more images, and more objects from those images, immediately after encoding than after a delay. However, we surprisingly found no significant difference in memory quality between the sleep and wake conditions. In the second experiment, we increased interference between images by having participants memorize and draw scenes within the same category, which has been shown to result in worse memory quality (Hall et al., 2021). Participants (N=175) encoded 8 scene images, in which 4 were from the same scene category (“within”) while the remaining 4 were of different categories (“between”). Participants recalled significantly more images from the “within” scene category, but significantly more objects from the “between” scene category. However, despite successfully increasing interference between images, we found no significant difference between sleep and wake for either metric. These results challenge the commonly held belief that sleep benefits all types of memory and suggests that visual recall is not impacted by sleep.

TALK 6, 9:30 AM, 51.26

CONTRIBUTIONS FROM LONG-TERM MEMORY EXPLAIN SUPERIOR WORKING MEMORY PERFORMANCE FOR MEANINGFUL OBJECTS

Hyung-Bum Park[1] (hbpark@uchicago.edu), Edward Awh[1]; [1]University of Chicago

Visual working memory (VWM) capacity has been claimed to be higher for meaningful objects compared to simple visual features, possibly due to their rich and distinctive representations. However, prior demonstrations have made this observation by comparing working memory performance for trial-unique objects and repeated sets of simple stimuli (e.g., a limited set of color categories). Unfortunately, this design includes a confound between meaningfulness and the strength of proactive interference, which is virtually absent for trial-unique object images. Thus, improved behavioral performance with meaningful stimuli could reflect contributions from episodic long-term memory (LTM) that are not accessible with repeated stimuli typically used in standard VWM capacity studies. To test this hypothesis, Experiment 1 measured VWM performance for trial-repeated colors, trial-repeated objects, and trial-unique objects. The results replicated the advantage for trial-unique objects over simple colors, but this advantage was eliminated with trial-repeated objects. Equivalent performance with colors and trial-repeated objects appears to contradict the claim that enhanced distinctiveness enables more meaningful objects to be stored in VWM. Instead, these findings indicate that LTM contributions in the trial-unique condition are eliminated by PI in the trial-repeated condition. To further test this interpretation, Experiment 2 measured contralateral delay activity (CDA), an electrophysiological marker of active VWM storage, for trial-repeated colors and trial-unique meaningful objects. Here again, we replicated the behavioral advantage for trial-unique objects, but CDA amplitudes plateaued at equivalent set sizes for objects and colors. Thus, our findings suggest that an equivalent number of trial-unique objects and colors can be stored in VWM, although testing with trial-unique objects invites contributions from familiarity-based representations in LTM (e.g., Endress & Potter, 2014; JEP:G).

Talk Session: Tuesday, May 20, 2025, 10:45 am – 12:15 pm, Talk Room 1

Attention: Temporal, spatial

Moderator: Sarah Shomstein, George Washington University

Talk 1, 10:45 am, 52.11

VOLUNTARY TEMPORAL ATTENTION ENHANCES INFORMATIONAL CONNECTIVITY ACROSS CORTICAL NETWORKS

Jiating Zhu[1], Karen Tian[1,2], Marisa Carrasco[2], Rachel Denison[1,2]; [1]Boston University, [2]New York University

Motivation: Selective attention enhances communication between brain areas. Whereas spatial attention alters communication through rhythmic synchronization, the impact of temporal attention on communication remains unexplored. Here we investigated whether and how voluntary temporal attention—the goal-directed selection of visual information at specific points in time—dynamically routes stimulus information across cortical networks. Methods: We recorded MEG data in human observers performing orientation discrimination in a two-target temporal cueing task. Observers were instructed with a precue to attend to one of two sequential grating stimuli (T1 and T2) separated by 300 ms, and with a response cue to judge the tilt of the target. To track the flow of stimulus information across the cortex, we constructed a dynamic informational connectivity network using source-reconstructed data. We measured correlations in stimulus orientation decoding accuracy (edges) across all pairs of atlas-based regions (nodes) within sliding time windows. We estimated a region's closeness centrality as its average inverse correlation distance to all other regions. Higher closeness centrality indicates greater informational connectivity with the rest of the network. Results: We found both early (~100 ms) and late (400-725 ms) modulations of closeness centrality in the informational connectivity network due to temporal attention. These enhancements emerged for both T1 and T2, primarily in the occipital and temporal lobes. We also observed an early occipital motif that recurred periodically in the entorhinal-parahippocampal cortex, but only when a target was temporally attended. The timing of this motif overlapped with the periods of closeness centrality enhancements. Conclusion: The results reveal how temporal attention dynamically routes stimulus information through cortical networks, indicating both early and late selection mechanisms at a network level.

This research was supported by National Institutes of Health National Eye Institute R01 EY019693 to M.C., T32 EY007136 to NYU, and F32 EY025533 to R.D., National Defense Science and Engineering Graduate Fellowship to K.J.T., and startup funding from Boston University to R.D.

Talk 2, 11:00 am, 52.12

TRACKING THE NEURAL SIGNATURES OF INTERNAL VISUAL AND MOTOR PRIORITIZATION ACROSS SPACE AND TIME

Irene Echeverria-Altuna[1,2], Sage Boettcher[2], Freek van Ede[3], Kate Watkins[2], Kia Nobre[1,2]; [1]Yale University, [2]University of Oxford, [3]Vrije Universiteit Amsterdam

Visual and motor contents within working memory can be flexibly and dynamically prioritized to guide adaptive behavior. Such prioritization is accompanied by modulations in frequency-specific electroencephalography (EEG) activity. Namely, lateralized posterior alpha-band (8-13 Hz) modulation tracks changes in internal attention to item locations, whereas lateralized central mu/beta-band (8-30 Hz) modulation tracks changes in response-plan prioritization. The limited spatial resolution of EEG, however, precludes resolving the cortico-subcortical networks underlying alpha and mu/beta modulations during internal attention. To investigate and compare the networks controlling internal visual and motor selection associated with these EEG markers, we completed a combined simultaneous fMRI-EEG study. Participants held two items (tilted bars) in working memory. In turn, items appeared on the left or right side and required a left- or right-hand response, with location and hand manipulated orthogonally. In half the trials, an informative retro-cue prompted participants to prioritize one encoded visual item and its associated action plan and, with the passage of time, to shift their focus to the other spatial location and response plan. In the other half of the trials, retro-cues were uninformative, so neither visual nor motor prioritization was possible. The EEG analyses replicated patterns of contralateral alpha modulation for spatial item selection and contralateral mu/beta modulation for motor selection. Analysis of fMRI data revealed the engagement of frontal and parietal areas associated with internal attention control during internal spatial attention and of the corresponding sensorimotor hand representation during internal action selection. Ongoing analyses will probe for cortico-subcortical systems tracking shifts in internal spatial and motor attention. The results promise interesting clues about the sources underlying the modulations of frequency-specific activity that accompany flexible sensory and motor prioritization in working memory.

Talk 3, 11:15 am, 52.13

ATTENTIONAL MODULATION OF STIMULUS-SYNCHRONIZED BOLD OSCILLATIONS IN THE HUMAN VISUAL CORTEX

Reebal Rafeh[1], Geoffrey Ngo[2], Lyle E. Muller[3], Ali R. Khan[4], Ravi S. Menon[4,5], Taylor W. Schmitz[2], Marieke Mur[6,7]; [1]Neuroscience Graduate Program, Western University, [2]Department of Physiology and Pharmacology, Schulich School of Medicine & Dentistry, Western University, [3]Department of Mathematics, Faculty of Science, Western University, [4]Department of Medical Biophysics, Schulich School of Medicine & Dentistry, Western University, [5]Centre for Functional and Metabolic Mapping, Robarts Research Institute, Western University, [6]Department of Physiology and Pharmacology, Schulich School of Medicine & Dentistry, Western University,

[7]Department of Psychology, Faculty of Social Science, Western University, [8]Department of Computer Science, Faculty of Science, Western University

Visual cortical neurons synchronize their firing rates to periodic visual stimuli. EEG is commonly used to study directed attention by frequency-tagging brain responses to multiple stimuli oscillating at different frequencies but is limited by its coarse spatial resolution. Here we leverage frequency-tagged fMRI (ft-fMRI) to study the influence of directed attention on the fine-grained spatiotemporal dynamics of competing stimulus-driven visual cortical oscillations. In this 7T fMRI experiment, participants (n=7) distributed their attention to simultaneously presented visual checkerboard stimuli oscillating at 0.125 Hz and 0.2 Hz. Our analysis revealed that distinct populations of visual cortical neurons exhibited either in-phase or anti-phase synchronization with the oscillating stimuli. The spatial topographies of these populations were highly replicable across scan sessions within participants, indicative of a fine-grained map of competitive feature-tuned responses (Dice's coefficient > 0.59; K-S test: p < 10-39). In accordance with this observation, we found that directed attention homogeneously increased the amplitude of anti-phase BOLD oscillations across the visual hierarchy, consistent with a distributed attention-driven suppressive field (Reynolds & Heeger, 2009). In contrast, attentional modulation of in-phase BOLD oscillations increased hierarchically from V1 to hV4, consistent with the effects of target enhancement reported in prior monkey electrophysiology (Moran & Desimone, 1985) and event-related fMRI work (Kastner et al., 1998). Finally, the strength of anti-phase (Wilcoxon test: p = 0.016), but not in-phase (p = 0.16), modulation predicted psychophysical correlates of attentional performance, further highlighting the mechanistic dissociation of attention-driven target enhancement and surround suppression. Together, our findings support the biased competition model of attention, which posits that attention modulates competing neural populations through concurrent mechanisms of enhancement and suppression. ft-fMRI extends the boundaries for research on the neural basis of biased competition in humans by providing a non-invasive method for distinguishing between concurrent stimulus-synchronized in-phase (enhancing) and anti-phase (suppressing) BOLD oscillations.

Talk 4, 11:30 am, 52.14

EXPLORING THE ROLE OF EYE MOVEMENTS IN THE ATTENTIONAL BLINK PHENOMENON: INSIGHTS FROM NATURALISTIC VISUAL EXPLORATION

Shahar Messika[1], Shlomit Yuval-Greenberg[1,2]; [1]School of Psychological Sciences, Tel Aviv University, [2]Sagol School of Neuroscience, Tel Aviv University

Humans can fixate for extended durations in experimental settings, but this behavior is highly unnatural. The visual system is inherently explorative, and suppressing this drive through fixation tasks may have profound perceptual consequences. We hypothesize that the visual system, due to its explorative nature, is not optimized to perceive sequential foveal targets presented in close temporal proximity. To investigate this, we revisit the well-known attentional blink (AB) phenomenon—a reduced ability to perceive a second target (T2) appearing 200-500ms after an initial target (T1). While various interpretations of the AB exist, its origins remain debated. We propose that the AB arises from the visual system's intrinsic drive to explore. Specifically, detecting T1 activates an explorative mechanism, shifting attention from the fovea to the periphery in search of the next saccade target, leading to frequent misses of the consecutive foveal target (T2). In two experiments (Total N=47) we tested this hypothesis using a novel AB-like design. Stimuli (digits/letters) appeared sequentially at 10Hz in changing, but predictable, spatial locations along a horizontal path with variability in the vertical-axis. T1 and T2 (letters/digits) were embedded within this sequence, separated by 1-5 lags and participants were asked to report both targets. Participants were not instructed to fixate and naturally tracked the moving stimuli with their gaze. Intermittent trials included a classic AB task with centrally presented stimuli. Results revealed a pronounced AB effect in the foveal condition but a significantly diminished effect in the novel task. These findings suggest that allowing free eye movements and aligning attentional shifts with peripheral exploration reduce the AB effect. We conclude that AB may stem from the unnatural constraints of fixation tasks. This research underscores the importance of dynamic visual exploration in cognition and suggests that classical cognitive phenomena studied under fixation should be re-evaluated in free-viewing conditions.

The study was funded by ISF grant 1960/19 to S-Y.G

Talk 5, 11:45 am, 52.15

CAN YOU ATTEND BROADLY IN SPACE WHILE ATTENDING NARROWLY IN TIME?: ON THE GENERALITY OF ATTENTIONAL BREADTH

Merve Erdogan[1], Anna C. Nobre[1], Brian Scholl[1]; [1]Yale University

A salient aspect of spatial attention is its variable *breadth*: sometimes we select narrowly (e.g. when hunting for lost keys in a cluttered drawer), while other times we select broadly (e.g. when viewing the overall configuration of players on a soccer field). And an analogous dynamic applies across time: sometimes we attend to relatively high-frequency changes (e.g. when listening to fast jazz) while other times we focus on events changing at a slower pace (e.g. waves crashing ashore). How are these different forms of attentional breadth related? For example, can you attend broadly in space while simultaneously attending narrowly in time? One might have no effect on the other. Or attending broadly in one domain might facilitate attending broadly in the other. Or broad vs. narrow attention might draw in part on different resources, such that attending broadly in one domain is easier when attending narrowly in the other. Participants were presented with four types of stimuli at once: (a) visual probes flashed in a relatively narrow ring around fixation (spatially narrow), (b) visual probes flashed in a wider ring relatively far from fixation (spatially broad), (c) repeated-tone probes in a high-frequency auditory stream (temporally narrow), or (d) repeated-tone probes in a lower-frequency stream (temporally broad). Across trials, participants were instructed to attend broadly vs. narrowly in space, and (independently) broadly vs. narrowly in time. As expected, attending to one visual ring impaired performance in the other -- and ditto for the two auditory streams. Critically, there were also cross-dimension interactions: for example, participants were better at focusing spatial attention *broadly* (detecting probes in the wider ring) when their temporal attention was focused *narrowly* (detecting probes in the higher-frequency stream). The interplay between spatial and temporal attention may thus depend on its relative breadth in each domain.

TALK 6, 12:00 PM, 52.16

UNRAVELING LEARNED DISTRACTOR SUPPRESSION: INSIGHTS FROM PSYCHOPHYSICS AND COMPUTATIONAL MODELING

Jan Theeuwes[1], Dock Duncan[1], Dirk van Moorselaar[1]; [1]Vrije Universiteit Amsterdam

The ability to ignore salient yet irrelevant stimuli is essential to accomplish everyday life activities. Previous research has shown that individuals improve their ability to suppress distracting items through experience; however, the mechanisms underlying this learned suppression remain unclear. The current study employed a psychophysical approach combined with computational modelling investigating how learned spatial suppression affects perception. The results show that items presented at suppressed locations are perceived as less bright than those in non-suppressed areas, suggesting that learned suppression directly affects the perceived saliency of items. To determine how saliency changes impact visual search, computational modelling was used to compare different models of attentional selection. The analysis favored a model in which learned suppression reduces the saliency of objects in suppressed locations during the initial salience computation. As a result, these items are less likely to compete for attentional processing and are therefore less likely to capture attention. To build on these findings, an additional set of experiments focused on high-probability distractor features instead of spatial distributions. The data were analyzed using the same modeling framework to assess whether feature learning operates in a similarly proactive manner. The results indicate that feature learning differs from spatial learning: Instead of a reduction in attentional capture, feature learning results in faster disengagement of attention from high probability distractor features such as frequently encountered colors and shapes. Collectively, these experiments suggest that spatial suppression operates proactively reducing attentional capture at suppressed locations, while feature suppression operates reactively facilitates faster disengagement from frequently encountered features. Together, these findings highlight the distinct mechanisms that drive learned distractor suppression.

ERC advanced grant [833029 - LEARNATTEND] and NWO SSH Open Competition Behaviour and Education grant [406.21.GO.034].

TALK SESSION: TUESDAY, MAY 20, 2025, 10:45 AM – 12:15 PM, TALK ROOM 2

Scene Perception

Moderator: Caroline Robertson, Dartmouth

TALK 1, 10:45 AM, 52.21

THE OCCIPITAL PLACE AREA (OPA) REPRESENTS LEFT-RIGHT INFORMATION IN 8-YEAR-OLDS, BUT NOT 5-YEAR-OLDS

Rebecca J. Rennert[1], Daniel D. Dilks[1]; [1]Emory University

Recent fMRI studies suggest that the occipital place area (OPA) – a scene-selective cortical region involved in “visually-guided navigation” – is surprisingly late developing, not involved in visually-guided navigation at all until around 8 years of age and thus develops abruptly (discontinuously). However, it could nevertheless be the case that OPA still supports visually-guided navigation (in some “weaker” form) continuously, for example, with earlier emerging representations of one of the most primitive kinds of navigationally-relevant information: “sense” (left-right) information. Indeed, studies have shown that young toddlers and even non-human species, including fish and insects, can perceive and utilize sense information for navigation. Thus, how does OPA develop, continuously or discontinously? Here we directly address this question by investigating the development of sense representation in OPA. More specifically, using functional magnetic resonance imaging (fMRI) adaptation in children at 5 and 8 years of age, we measured the response in OPA to repeating images of identical scenes (“same” condition), mirror-image reversals of scenes (“mirror” condition), or different scenes (“different” condition). Importantly, OPA in both groups adapted to the “same” condition (demonstrating that the adaptation paradigm is working). Crucially, however, we found that OPA in 8-year-olds, like adults, was sensitive to the “mirror” condition (thus representing sense information), while OPA in 5-year-olds was not (thus not representing sense information). Taken together, these findings i) reveal further evidence that the visually-guided navigation system undergoes protracted development, not even supporting sense information until around 8 years of age, and ii) raise the intriguing possibility that OPA develops discontinuously, developing abruptly around 8 years of age.

TALK 2, 11:00 AM, 52.22

BEHAVIORAL AND NEURAL CORRELATES OF SURROUND SUPPRESSION IN DYNAMIC NATURAL SCENES

Merve Kınıklıoğlu[1] (mervekiniklioglu@gmail.com), Daniel Kaiser[1,2]; [1]Neural Computation Group, Department of Mathematics and Computer Science, Physics, Geography, Justus Liebig University Gießen, Germany, [2]Center for Mind, Brain and Behavior (CMBB), Philipps-University Marburg, Justus Liebig University Gießen and Technical University Darmstadt, Germany

Sensitivity to a central grating often decreases when it is presented with an annular grating. Such surround suppression is observed in neural activity within regions like MT and also manifests in motion perception. Notably, suppression strength decreases when the central and surrounding stimuli move in opposite directions. While previous studies have primarily investigated surround suppression using low-level stimuli like drifting gratings, evidence from ecologically relevant contexts remains limited. In this study, we investigated the behavioral and neural correlates of surround suppression in humans using dynamic natural scenes. Additionally, we examined whether the suppression effect is modulated by the content similarity between the central and surrounding scenes. Participants viewed panoramic videos created by moving static images behind a circular occluder, with central images shown through a 1.9° aperture and surrounding images through a 2.5°-10.4° annular aperture. The resulting videos depicted natural scenes with varied relationships between the center and surround, including identical exemplars, different exemplars of the same basic-level category, and videos from different basic-level or superordinate categories. In a behavioral experiment, participants reported the superordinate categories of the central and surrounding videos. Results showed that surround suppression was strongest when the central and surrounding videos were from different superordinate categories and weakened as content similarity increased. Furthermore, suppression decreased when the videos moved in opposite directions. In a complementary fMRI experiment, participants viewed the same stimuli while detecting rare target events. Findings revealed that surround suppression was evident in hMT+, with stronger suppression during same-direction videos. In contrast, regions such as V1 and scene-selective areas (OPA and PPA) exhibited surround facilitation. Critically, no content-specific effects were observed in any region of interest. These findings suggest that hMT+ activity alone cannot fully explain the perceptual suppression effect, suggesting complex interactions among multiple brain regions in surround suppression with dynamic natural scenes.

D.K. is supported by the DFG Grants SFB/TRR135, Project Number 222641018; KA4683/5-1, Project Number 518483074, “The Adaptive Mind,” funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art, and an ERC Starting Grant PEP, ERC-2022-STG 101076057.

Talk 3, 11:15 am, 52.23

THREE CATEGORIZATION TASKS YIELD COMPARABLE CATEGORY SPACES: A COMPARISON USING REAL-WORLD SCENES

Pei-Ling Yang[1] (plyang2@illinois.edu), Diane M. Beck[1]; [1]University of Illinois

Categorization is a fundamental cognitive function that depends on the similarity among items. Many tasks have been developed to measure item similarities to better understand the underlying similarity space of categories. Most such studies have focused on object categories. The current study compares the measured similarity outcomes of three categorization tasks for real-world scenes: an arrangement, odd-one-out, and same-different judgment task. Our study asks whether a stable representation for scene categories (i.e., beach, city, mountain) exists across tasks. To assess the reliability of the tasks, each task was conducted twice for each participant (N = 98 for each task). The arrangement task asked participants to place each scene image relative to the three text anchors that they set at the beginning of the task. The odd-one-out task required participants to choose the scene that is the most dissimilar to the other two in a triplet of scenes. The same-different judgment task asked participants to respond whether a pair of scenes was from the same category or not. The similarity matrices were derived from distances in pixel space for the arrangement task, the probability of choosing a scene as dissimilar for the odd-one-out task, and the probability of reporting two scenes as ‘same’ for the same/different task. The rank correlations were calculated between two repeats of the task to examine the reliability of the similarities. All three tasks showed comparable rank correlation reliability: arrangement (0.61), odd-one-out (0.60), same-different (0.69). Ordinal multidimensional scaling on the similarity matrices of each task was used to construct 3-D category spaces, where the distances reflect the similarity among stimuli. These rank-derived spaces were moderately correlated across tasks: arrangement & same-different (0.67), arrangement & odd-one-out (0.58), odd-one-out & same-different (0.57). These results imply some stable representation of scene category similarity space that is worth further investigation.

Talk 4, 11:30 am, 52.24

UNEXPECTED SCENE VIEWS ARE PRIORITIZED IN PERCEPTUAL AWARENESS IN REAL-WORLD ENVIRONMENTS

Anna Mynick[1], Michael A. Cohen[2], Adithi Jayaraman[1], Kala Goyal[1], Caroline E. Robertson[1]; [1]Dartmouth College, [2]Amherst College

Although we only perceive a small portion of the visual world at any given moment, we operate within the world with remarkable efficiency. How does memory for the world around us help facilitate perception in naturalistic environments? Here, we examined how memory for immersive, 360º environments shapes perceptual awareness of the world around us across head turns. Participants (N=64) first studied a set of immersive real-world scenes from the local college campus in head-mounted virtual reality (VR). After studying these scenes, each trial began by showing participants a limited view from a studied scene (prime). Then, the prime disappeared and participants turned their head left or right 90º towards a target view. Targets were initially masked with continuous flash suppression, and participants’ task was to indicate when the target entered perceptual awareness. 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). Targets depicted another view from the same scene, either displayed in their learned spatial position (‘expected view’) or displayed 180º opposite their learned position (‘unexpected view’). We found that unexpected scene views entered perceptual awareness roughly 200ms faster than expected scene views (p=.006). Intriguingly, scene memory alone did not explain these results: the same pattern emerged in a new group of participants (N=31) who completed the same task on a set of unfamiliar, never-studied scenes (p=.04). Taken together, these results suggest that extrapolation beyond the current field of view is sufficient to draw expected and unexpected scene views into perceptual awareness at different rates. More broadly, our findings indicate that expectations of upcoming scene views shape how scene information is prioritized in perceptual awareness across head turns, supporting efficient interaction in immersive environments despite our limited field of view.

Talk 5, 11:45 am, 52.25

MODELING HUMAN SCENE UNDERSTANDING FIXATION-BY-FIXATION USING GENERATIVE MODELS

Ritik Raina[1] (ritik.raina@stonybrook.edu), Alexandros Graikos[1], Abe Leite[1], Seoyoung Ahn[2], Gregory Zelinsky[1]; [1]Stony Brook University, [2]UC Berkeley

The gist-level of understanding that humans extract from brief exposures to a scene becomes quickly elaborated with a more detailed scene understanding that emerges with each new fixation made during scene viewing. Each fixation enables the extraction of rich visual and contextual features that encode new information about the scene's objects and their spatial interrelationships. Successive fixations lead to an incrementally evolving representation of the scene's meaning and layout that parallels the viewer's evolving understanding of the scene, but this dynamic evolution of scene understanding has yet to be modeled. To this end, we present SparseDiff, a novel approach that leverages generative (latent diffusion) modeling to incrementally generate images that reflect progressively detailed levels of scene understanding that can be used as hypotheses for behavioral evaluation. Our model uses a self-supervised image encoder (DINOv2) to extract visual representations from the fixated regions of an image that capture both local object and broader scene context features. These fixation-grounded features then condition a pre-trained text-to-image diffusion model to generate full coherent scene representations. As SparseDiff accumulates information from successive fixations, it iteratively generates increasingly refined hypotheses based on the scene's objects that were fixated while filling in unattended regions with contextually plausible content. We evaluated SparseDiff using the COCO-FreeView dataset and found that increasing the number of fixations provided to the model led to enhanced visual and semantic similarity between generated and original images as measured by image similarity metrics (e.g. CLIP, DreamSim) that capture high-level semantic alignment as well as mid-level feature similarity. Future work will use same-different tasks to evaluate whether generated hypotheses are scene "metamers" for what was perceived during viewing, and use SparseDiff as a tool to study individual differences in visual perception by comparing the different scene understandings generated from different viewing scanpaths.

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

TALK 6, 12:00 PM, 52.26

CONTOURS, NOT TEXTURES DETERMINE ORIENTATION TUNING IN HUMANS

Seohee Han[1] (seohee.han@mail.utoronto.ca), Dirk B. Walther[1]; [1]University of Toronto

Ever since the seminal discoveries of Hubel and Wiesel, we know that cortical representations of visual input start with oriented edges and lines in the primary visual cortex. Determining exactly how orientation information is processed in the human brain is critical for understanding the computational and neural mechanisms of vision. Filter and contour-based methods, including simple bars and Gabor filters, have historically been used interchangeably for detecting orientation activity. While these approaches effectively capture orientation, filter-based methods typically aggregate data across spatial frequencies, blending texture and contour information. This overlap raises important questions about what specific orientation features are most relevant for perception and neural representation. In this study, we address this important question with two complementary approaches. First, we investigated human orientation judgments using image patches with maximal and minimal differences between average orientations computed by steerable pyramid filters and a contour-based method. Behavioural results revealed that human judgments align with contour-based orientation but not with filter-based orientation when the two were in conflict. Observers clearly prioritized contours over textures when summarizing orientation in complex scenes. Second, we evaluated the impact of orientation computation methods on neural maps of orientation selectivity, using Roth and colleagues' (2022) image-computable model as a benchmark. By comparing filter-based methods applied to photographs and line drawings with a contour-based method, we assessed how the choice of computation influences model fit and voxel-level orientation preference in the visual cortex. Again, we found a clear advantage of contours over textures in explaining orientation tuning in the visual cortex. These findings underscore the importance of oriented contours rather than textures as the elemental building blocks of vision. By highlighting the importance of contours for human orientation judgments and neural selectivity in the visual cortex, our work emphasizes the need for methodological alignment in visual neuroscience research.

TALK SESSION: TUESDAY, MAY 20, 2025, 2:45 – 4:45 PM, TALK ROOM 1

Plasticity and Learning

Moderator: Aaron Seitz, Northeastern University

TALK 1, 2:45 PM, 54.11

HOMOTOPIC REORGANIZATION OF THE VISUAL WORD FORM AREA FOLLOWING SURGICAL RESECTION OF CONNECTED CORTEX

Beth Rispoli[1] (bmr99@georgetown.edu), Tina Liu[1], Kyungji Moon[1], Radhika Chatterjee[2], Kareem Zaghloul[2], Sara Inati[2]; [1]Georgetown University Medical Center, [2]National Institute of Neurological Disorders and Stroke

The visual word form area (VWFA), typically located in the left ventral temporal cortex (VTC), is crucial for word reading due to its privileged connectivity to higher-order visual, semantic, and language areas. While acute lesions involving the VWFA in adulthood can result in reading impairments (e.g., alexia), the impact of lesions to connected cortex on VWFA function remains unclear. In this study, we ask whether a resection of anterior temporal lobe (ATL) impacts reading selectivity in the VTC and how outcomes may vary by hemisphere of resection. We analyzed a clinical dataset of adult patients (n=9) who underwent surgical resection of ATL as a treatment for drug-resistant epilepsy. Patients completed an fMRI reading task both pre-operatively and one year post-operatively. During the reading task, patients read on-screen stories and viewed a matched control condition. VTC anatomical ROIs were drawn in the native space of each patient's brain for each hemisphere, and mean responses for the reading and control conditions were computed. Reading selectivity was calculated in each ROI as the difference in beta weights between the reading and control conditions for each session. We found that the hemisphere of resection was predictive of a change in reading selectivity after surgery in the intact, homotopic VTC. Following left ATL resection (n=5), all patients showed a consistent increase in reading selectivity in the intact right VTC. Conversely, patients with right ATL resection (n=4) did not show this pattern, likely because language and reading functions remained lateralized to the left hemisphere. Together, these findings highlight the plasticity of reading-related circuitry in adulthood, showing that disruption of left hemisphere VWFA connectivity via ATL resection can promote compensatory reliance on the intact right VWFA. Importantly, this homotopic reorganization of the VWFA appears to be asymmetric, as no comparable reorganization occurs following right hemisphere ATL resection.

TALK 2, 3:00 PM, 54.12

SHORT-TERM MONOCULAR DEPRIVATION IN ADULT HUMANS: A META-ANALYSIS AND NEW PERSPECTIVES

Cecilia Steinwurzel[1] (cecilia.steinwurzel@protonmail.com), Giacomo Pennella[2], Claudia Lunghi[3], Paola Binda[1]; [1]University of Pisa, [2]University of Florence, [3]Ecole Normale Supérieure & CNRS, Paris

Few hours of monocular deprivation in adult humans produces a transient shift of ocular dominance in favor of the deprived eye. This phenomenon has been investigated with several methodologies in 75 studies since 2011. We compiled a meta-analysis of these studies, structured following the PRISMA checklist and selectively including studies on healthy humans. Each study includes multiple experiments (for a total of about 180); for each, we computed a standardized effect size and, where possible, we quantified the decay rate of the effect. In the majority of studies, deprivation was achieved by covering one eye with an opaque or a translucent patch, with comparable outcomes (two-sample $t(108) = -0.02$, $p = 0.98$). Deprivation effects were mainly measured with three types of tasks: binocular cooperation, binocular competition, and monocular thresholds. Effects are larger when measured with binocular cooperation compared to binocular competition (two-sample $t(100) = 4.66$, $p = 0.001$), possibly reflecting the different time windows used to estimate effects. Longer periods of deprivation induce larger ($r(107) = 0.32$, $p = 0.001$) and longer lasting effects ($r(38) = 0.38$, $p = 0.020$). Patching either eye (dominant or non-dominant) produces a similar outcome (two-sample $t(98) = -0.74$, $p = 0.46$). Depriving one eye produces opposite effects in the two eyes; however, pooling across studies, we find that opaque patching primarily suppresses the non-deprived eye, while translucent patching primarily boosts the deprived eye (two-sample $t(31) = 2.99$, $p = 0.005$). A growing number of studies show that monocular deprivation effects can be mimicked by manipulations that do not impact the strength of the monocular signal, including degradation of image quality (pink noise, kaleidoscope), image inversion in space or time, or just monocular delay. We conclude that the observed shift in eye dominance results from a complex interplay between bottom-up visual stimulation and top-down signals

Research funded by: - grant PREDACTIVE (n. 101170249) - grant (No 948366-HOPLA) - grant (THE, CUP I53C22000780001), - grant PRIN 2022 (Project RIGHTSTRESS, Grant no. 2022CCPJ3J, CUP I53D23003960006) - grant SMILY (Italian Ministry of University and Research - FARE-2).

TALK 3, 3:15 PM, 54.13

HOLISTIC BAYESIAN MODEL OF PERCEPTUAL ADAPTATION THAT COMBINES EFFICIENT CODING AND NOVELTY DETECTION

Jiang Mao[1], Alan Stocker[1]; [1]University of Pennsylvania

Perceptual systems continually adapt to statistical changes of their sensory environment. This leads to changes in perception (i.e. aftereffects) that have been well characterized. The exact functional purpose of adaptation, however, remains unclear. The efficient coding hypothesis is one prominent account of adaptation. It proposes that neural representations adapt to changes in the input statistics in order to maximize encoding efficiency given overall resource limitations. Another account for adaptation is that it enhances the detection of novel stimuli. Neural repetition suppression as well as increased discriminability at the adaptor indicate that new stimuli that differ from the adaptor are more salient. Here we present a model of adaptation that incorporates both of these rationales within the framework of holistic Bayesian inference. The model assumes efficient sensory encoding that adapts to the distribution of future stimuli predicted by previous input statistics. Specifically, we constrain encoding accuracy with previously characterized, efficient encoding changes caused by prolonged adaptation to a single adaptor stimulus (Mao et al. 2024). The model also incorporates novelty detection by assuming that the observer performs categorical inference to determine whether the next stimulus originates from a distribution different from the distribution of the adaptor (hence is novel) or not. We show that this model can account for many of the previously reported perceptual aftereffects measured with estimation tasks (method of adjustment). It accurately explains the characteristic repulsive biases and changes in estimation variance after adaptation to a single adaptor. Furthermore, the model predicts aftereffects for arbitrary adaptors, in particular adaptors that consist of a distribution of feature values. In conclusion, we present a holistic inference model for adaptation that unites two major functional objectives of sensory adaptation, efficient coding and novelty detection.

TALK 4, 3:30 PM, 54.14

TRIAL-LEVEL PUPILLARY RESPONSE REVEALS THAT NEURAL GAIN IS ROBUSTLY LINKED TO PERCEPTUAL PLASTICITY AND EXPLAINS THE ROLE OF FEEDBACK IN LEARNING

Aaron Cochrane[1] (aaron_cochrane@brown.edu), Soph Bililies[1], Shreya Gulati[1], Kiley Haberkorn[1], Yuka Sasaki[1], Takeo Watanabe[1]; [1]Brown University

Visual perceptual learning arises from neuroplasticity leading to improved low-level inferences about environmental signals. This experience-dependent neural change can be influenced by various factors, including explicit and informative feedback, which is thought to boost neuroplastic changes. If feedback works through such a signal boosting mechanism to improve learning, global measures of “neural gain” (e.g., pupillary dilation in response to a stimulus) should mediate the effects of feedback on perceptual learning. Here, we examined the detailed time course of perceptual improvements during five days of training on discrimination of low-coherence dot-motion under three

conditions involving explicit feedback: no feedback, intermittent feedback, and first-day-only feedback. Performance improved significantly more for the first-day-only feedback group compared to the no-feedback group, even days after feedback was removed. This finding suggests a lasting effect in which initial feedback enabled participants to learn how to generate endogenous feedback signals; such signals may be revealed by measures of neural gain. To identify how the dynamics of perceptual change are linked to neural gain, we measured task-evoked pupillary response during each trial. Using local derivative analyses, we linked trial-based pupillary responses to changes in perceptual threshold (i.e., first derivative). Two striking effects emerged: First, differential pupillary responses to trials of varying difficulty (referred to as "adaptive neural gain") were reliably associated with the trial-to-trial changes in perceptual threshold. Specifically, trials with greater adaptive neural gain occurred during periods of larger decreases in threshold, providing evidence of a single-trial mechanism for perceptual plasticity. Second, the presence of feedback enhanced this association, demonstrating that feedback strengthens the learner's ability to dynamically regulate neural gain which leads to plasticity. These results provide evidence for the role of within-trial perceptual signal boosting and reveal a mechanism of neural gain through which feedback improves perceptual learning.

NIH R01EY019466, R01EY027841, R01EY031705, NSF-BSF BCS2241417

TALK 5, 3:45 PM, 54.15

BRIEF MEMORY REACTIVATIONS ENABLE GENERALIZATION OF PERCEPTUAL LEARNING

Taly Kondat[1,2], Yuka Sasaki[3], Takeo Watanabe[3], Nitzan Censor[1,2]; [1]Sagol School of Neuroscience, Tel Aviv University, [2]School of Psychological Sciences, Tel Aviv University, [3]Brown University, Providence, USA

Perceptual learning can significantly improve human visual sensitivity. However, it requires extensive stimuli exposure, and the ability to generalize learning to untrained conditions is often limited. Traditionally, perceptual learning is attributed to practice-dependent plasticity mechanisms. Nevertheless, recent studies suggest that brief memory reactivations can efficiently improve visual perception, exhibiting comparable gains through distinct mechanisms which recruit higher-level parietal brain regions. Here, we provide evidence that such memory reactivation mechanisms promote generalization of learning. Human participants first encoded a standard texture discrimination task with the target stimulus at retinotopic location A. Brief memory reactivations of only five trials each were performed on separate days at location A. Then, generalization was tested at retinotopic location B. Results indicate remarkable enhancement of location B performance following memory reactivations, pointing to efficient offline generalization mechanisms. A control experiment with no reactivations showed minimal generalization. These findings suggest that beyond dramatically reducing stimuli exposure, reactivation-induced learning further enhances learning efficiency by promoting generalization, and may extend to additional learning domains, with potential future clinical implications.

The study was supported by the European Research Council (ERC-2019-COG 866093) and the NSF-BSF grant (National Science Foundation 2241417, Binational Science Foundation 2022666). Taly Kondat is a fellow of the Ariane de Rothschild Women Doctoral Program.

TALK 6, 4:00 PM, 54.16

THE ROSTROLATERAL PREFRONTAL CORTEX IS ACTIVATED DURING A RULE-BASED VISUAL CATEGORIZATION TASKS

Jacob Vaughn[1] (jacovaug@ttu.edu), Shinyoung Jung[1], Sean R. O'Bryan[1], Tyler Davis[1], Miranda Scolari[1]; [1]Texas Tech University

Categorization is a fundamental cognitive process that allows us to make sense of the external world by extrapolating from previous visual experiences to better understand novel ones. One under-investigated aspect of categorization is how this extrapolation is neurologically supported; however, prior research indicates that humans use inferential processes, such as hypothesis testing and logical rules, when confronted with novel visual stimuli. One area that is hypothesized to instantiate these inferential processes is the rostrolateral prefrontal cortex (rlPFC). To investigate this, we conducted a visual rule-based categorization tasks while participants were in a functional magnetic resonance imaging (fMRI) machine. During the first phase of the experiment, participants (N = 50) learned the task rule via trial-by-trial feedback. During the second phase, participants were tasked with categorizing stimuli without feedback that contained novel visual features or feature combinations that were not presented during learning. This generalization phase consisted of six types of trials: three of which were ambiguous with respect to the learned rules and three were unambiguous. Our results showed that the rlPFC was significantly activated during the learning phase, and only during the ambiguous trials in the generalization phase. In contrast, primary visual cortex and MTL were activated during all trials of the generalization phase. Our results suggest that the rlPFC may be critical for early learning (times of uncertainty) and then, once a rule is learned, only activates again in situations where learned rules are inapplicable (ambiguous trials). Critically, the lack of rlPFC activation during the unambiguous trials suggests that visual feature novelty by itself is not enough to recruit the rlPFC.

This material is based upon work supported by the National Science Foundation under Grant No. 1923267 awarded to M.S.

TALK 7, 4:15 PM, 54.17

PERCEPTUAL LEARNING IMPROVES PATTERN CLASSIFICATION IN THE PRIMARY VISUAL CORTEX REPRESENTATION OF THE TRAINED LOCATION

Pinar Demirayak[1], Paul Stewart[1], Elam Cutts[1], Madeline Ragland[1], Elliot Maxwell[1], Marcello Maniglia[2], Samyukta Jayakumar[2], Jaap Munneke[3], Nicholas Turk-Browne[4], Aaron Seitz[3], Kristina Visscher[1]; [1]Department of Neurobiology, University of Alabama at Birmingham, Birmingham, AL, USA, [2]Department of Psychology, University of California at Riverside, Riverside, CA, USA, [3]Department of

Psychology, Northeastern University, Boston, MA, USA, [4]Department of Psychology, Yale University, New Haven, CT, USA

Deterioration of the photoreceptors in the center of the retina can lead to loss of central vision while leaving peripheral vision relatively spared. Individuals with central vision loss often spontaneously develop a "preferred retinal locus" outside the region of vision loss which they use for daily tasks needing finely detailed vision. To better understand the functional brain plasticity associated with this compensation following central vision loss, we used a simulated scotoma paradigm in healthy individuals. Twenty-two participants were trained on basic visual tasks, including contrast sensitivity and contour integration, at a Trained peripheral Retinal Location (TRL) either to the left or right of an artificial scotoma. A control region on the other side was defined as an Unpreferred Retinal Locus (URL). We collected a total of six runs of task-based fMRI data in each pre- and post-training MRI session. Using Multi-Voxel Pattern Analysis, we calculated classification accuracy for differentiating the cortical responses to orientation discrimination stimuli (Gabors) vs contour integration stimuli (contours made up of small Gabors) at the V1 cortical area contralateral to the trained location (cTRL) or contralateral to the URL (cURL). We observed that the cTRL improved classification accuracy more with training than the cURL did (significant ROI x session interaction). Our results indicate that experience leads to more distinct neural activity in a trained location.

NEI 1R01EY031589-01

Talk 8, 4:30 pm, 54.18

FILTERING MECHANISMS IN EARLY VISUAL CORTEX MODULATE THE OCCURRENCE OF VISUAL PLASTICITY AND LEARNING

Markus Becker[1] (markus.becker@ur.de), Sebastian M. Frank[1]; [1]University of Regensburg

Visual perceptual learning (VPL), a type of skill learning, is an example of visual plasticity in the adult brain and can occur by mere repeated exposure to task-irrelevant visual features. However, task-irrelevant VPL poses the risk of learning features that are not necessarily adaptive and could jeopardize the stability of previously acquired VPL for task-relevant features. Therefore, one might predict that there are mechanisms that minimize the occurrence of task-irrelevant VPL. An efficient solution would be to filter out task-irrelevant features already at the earliest stage of cortical visual processing (early visual cortex) before they are further processed in higher-order visual areas. Here, we examined this possibility by exposing 30 participants to a task-irrelevant feature (coherent motion in one direction) that was either salient (easily detectable) or weak (difficult to detect) in the visual periphery while they performed a rapid-serial-visual-presentation (RSVP) task at central fixation for six sessions on separate days. Before the first and after the final exposure session functional magnetic resonance spectroscopy (fMRS) using consecutive MEGA-PRESS scans was conducted while participants were exposed to the task-irrelevant feature and performed the RSVP task. The behavioral results showed that task-irrelevant VPL occurred for the weak but not salient task-irrelevant feature. The imaging results showed that glutamate-levels in early visual cortex were significantly lower during exposure to the salient task-irrelevant feature than the weak task-irrelevant feature in each fMRS session. This difference in glutamate-levels in early visual cortex was not found in a control experiment, in which the peripherally exposed salient and weak features were rendered task-relevant. Our results suggest that filtering mechanisms exist at the earliest stage of cortical visual processing, which modulate the occurrence of visual plasticity and learning. This filtering might occur through downregulating excitatory, glutamatergic activity in early visual cortex for task-irrelevant features that are sufficiently salient.

Deutsche Forschungsgemeinschaft (DFG): Emmy Noether Grant (Project Number 491290285); Julitta und Richard Müller Stiftung

Talk Session: Tuesday, May 20, 2025, 2:45 – 4:45 pm, Talk Room 2

Motion: Models, Neural mechanisms

Moderator: Ralf M. Haefner, University of Rochester

Talk 1, 2:45 pm, 54.21

DURATION OF OCCLUSION EFFECT IS BETTER EXPLAINED BY FEEDFORWARD THAN RECURRENT COMPUTATIONS

Giuliana Martinatti Giorjiani[1,2], 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

The ability to predict future locations of a moving object from its past trajectory underpins a wide range of complex behaviors. Already in the retina, motion extrapolation is found in ganglion cells that fire in anticipation of a stimulus entering their receptive fields, presumably compensating for neuronal transmission delays on short time scales (milliseconds) in a feedforward manner. In contrast, goal-oriented behaviors like target interception or tracking under occlusion rely on accumulated information from an object's past trajectory over longer time scales (seconds), requiring a storage component in addition to feedback. Recurrent processing could serve as such a storage mechanism for goal-oriented motion extrapolation. We tested this by comparing the performance of feedforward and recurrent neural networks to human participants in a motion extrapolation task involving visual occlusion. Human participants covertly and continuously tracked a target moving along a circular trajectory at constant speed. An invisible occluder masked half of the trajectory. A "shot" (red dot) could briefly appear either ahead, on top of, or behind the (visible or occluded) target location at different time points. Participants judged the shot's position relative to the target via a two-alternative forced-choice. Results show a decay in precision with increasing occlusion time, accompanied by an acceleration bias in target location estimates. The networks were trained on a synthetic dataset with 640.830 periodic trajectories. They were then tested by predicting the "occluded" segment of the trajectory from the input ("visible" segment). Both networks exhibited a decline in precision over time, with the feedforward network more closely resembling the decay found in human data. However, neither network replicated the human acceleration bias. These findings suggest that decay in precision may stem from feedforward computations, while the acceleration bias likely reflects learned priors, potentially carried through feedback processing rather than from feedforward or recurrent processing.

Talk 2, 3:00 pm, 54.22

CONFLICTING HEADING BIASES EXPLAINED BY DIFFERENT REFERENCE FRAMES

Renate Reisenegger[1,2], Ambika Bansal[3], Laurence R Harris[3], Frank Bremmer[1,2]; [1]Applied Physics and Neurophysics, Philipps-Universität Marburg, Karl-v-Frisch Str 8a, 35043 Marburg, Germany, [2]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt, Hans-Meerwein-Str. 6, 35032 Marburg, Germany, [3]Center for Vision Research, York University, Toronto, Ontario, Canada

Walking or moving around the world generates a characteristic visual stimulus called optic flow. Optic flow provides essential information about our self-motion, including direction (heading). Heading perception has been studied extensively with some studies showing a bias towards straight-ahead (centripetal) while others report the opposite bias – away from straight-ahead (centrifugal). It was recently suggested that response ranges cause the opposing biases, with smaller ranges generating centripetal and larger ranges producing centrifugal biases (Sun et al., bioRxiv, 2024). However, we noticed that centripetal biases are observed when participants are asked to "point" to their perceived heading along a horizontal line in an egocentric reference frame. In contrast, centrifugal biases arise when participants imagine themselves from above (a "bird's eye view") and report their heading as an angle on a circle or arc in an allocentric reference frame. This suggests egocentric vs allocentric reference frames may induce different heading biases. We assessed heading perception using the "Edgeless Graphics Geometry display" (EGG), a very large-range (~224° field of view) edgeless display. Participants viewed an optic flow stimulus (600ms) simulating self-motion across a ground plane in various directions within ±55° of straight ahead. They then reported their perceived heading by positioning a target on either (i) a horizontal line spanning ±110° (egocentric condition) or (ii) an arc spanning ±110° as seen from above (the "bird's eye view" or allocentric condition). Results showed a centripetal bias in the egocentric condition and a centrifugal bias in the allocentric condition. These findings challenge the recent proposal that heading biases are solely driven by response range. Both our conditions had the same large (±110°) response range, which would predict centrifugal biases, yet we observed a centripetal bias in the egocentric condition. Instead, we show that biases in heading perception are primarily driven by the reference frame used.

Supported by the Deutsche Forschungsgemeinschaft (DFG: IRTG-1901 and CRC/TRR-135), the Hessian Ministry of Higher Education, Science, Research and the Arts (HMWK: Clusterproject The Adaptive Mind), and the EU (PLACES). The EGG display was provided by the Canadian Foundation for Innovation.

Talk 3, 3:15 pm, 54.23

COMPARING YOUNG WITH OLDER ADULTS IN TERMS OF CAUSAL INFERENCE DURING COMPLEX MOTION PERCEPTION

Maeve Silverman[1], Xinyi Yuan[1], Sabyasachi Shivkumar[2], Ralf M. Haefner[1]; [1]University of Rochester, [2]Columbia University

Surround suppression has been documented neurally and psychophysically in multiple sensory modalities and brain areas (Tadin et al. 2003). Interestingly, prior work has found differences in the strength of psychophysical surround suppression correlated with IQ (Melnick et al. 2013), mental disorders, and age (Betts et al. 2005, 2009). Surround suppression is also conjectured to be of critical importance for scene segmentation. The strongest visual cue for segmentation is common motion ("common fate"). Recently, motion perception for center-surround stimuli was formalized as Bayesian causal inference over reference frames, and shown to accurately capture the range of percepts reported by human observers for different center-surround configurations (Shivkumar et al. 2024). Here, we leverage the previously developed subjective motion direction estimation task in order to study potential differences in motion perception and causal inference between young and older adults. Fitting a causal inference model to the data allowed us to characterize differences not just in terms of the raw data (e.g. strength of bias) but also in terms of the underlying model parameters. We collected data from 10 college age observers (age 18-25), and 10 older adults (above 65). Each observer contributed 330 perceived direction reports of coherently moving dots surrounded by either moving or stationary dots. Varying the relative direction of center and surround, we could constrain the underlying causal inference process that characterizes the transition from integration to segmentation. Interestingly, we found no significant differences in the response biases between young and old. However, we found significant differences between two model parameters: sensory noise and computational noise were almost an order of magnitude higher in older adults than in young adults ($p<0.01$ and $p<0.05$, respectively) implying a high sensitivity of our task to age-related changes, but no noticeable change to the beliefs underlying causal inference over the lifespan.

We acknowledge funding support from NIH/U19 NS118246, and NSF/CAREER IIS-2143440.

Talk 4, 3:30 pm, 54.24

V1-INDEPENDENT DEVELOPMENT OF DIRECTION TUNING IN HIGHER ORDER VISUAL CORTEX

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

Early postnatal visual experience drives immature cortical circuits to refine their tuning for features like direction of motion. The network-level mechanisms that underly these changes remain unclear. Many believe visual cortical development occurs sequentially and beginning in primary visual cortex (V1). However, early projections from first order visual thalamus to more rostral visual areas suggest parallel streams of input could support V1-independent development of higher order visual cortex. In the present study, we used two processing stages in the ferret visual motion pathway, V1 and higher motion area PMLS, as a platform to test the necessity of canonically 'first' order visual inputs in the functional development of higher order visual cortex. Visually naïve kits (postnatal day (P) 28-32; ferrets open their eyes ~P30) were anesthetized and exposed to an acute experience paradigm previously observed to induce rapid functional development of direction tuning in ferret visual cortex. Using simultaneous electrophysiological recordings of spiking activity in both areas, the initial direction tuning was assessed before bidirectional drifting gratings were shown to the kits for 8 hours. We inactivated V1 during stimulus presentations with localized hypothermia to assess the contributions of its inputs to the stimulus-driven changes in PMLS response properties. After the 8 hours of visual experience, direction tuning was assessed again in both areas. Analysis of single- and multi-unit responses using conventional tuning metrics showed increased direction tuning in PMLS, but not in V1, suggesting that functional development of direction tuning in PMLS can occur independently of V1 inputs. Preliminary analysis using representational distances of population responses hints that, despite no significant change at the single unit level, orientation tuning in V1 may still increase at the population level. This would suggest encoding of stimulus orientation in V1 responses can improve independently of visually driven activity in V1.

This work was supported by the NIH (1R01EY035807) and the Kavli NDI

Talk 5, 3:45 pm, 54.25

THE ROLE OF ACTIVE VISION IN THE PRIMARY VISUAL CORTEX OF FREELY-MOVING MARMOSETS

Jingwen Li[1], Vikram Singh[1], Jude Mitchell[2], Alexander Huk[3], Cory Miller[1]; [1]UC San Diego, [2]University of Rochester, [3]UC Los Angeles

Historically, studies of visual cortex have been performed while nonhuman primates are head-fixed viewing visual stimuli on a screen. In the real world, however, visual processing must accommodate how we actively explore the environment. Despite its significance, little is known about how the primate visual system supports natural, active vision in freely moving animals. To address this problem, we leveraged an innovative, head-mounted eye-tracking system developed for marmosets in our lab while simultaneously recording the activity of single neurons in V1 to examine the effect of eye, head, and body movement on visual representations. In these experiments, monkeys are first head-fixed to quantify traditional visual stimuli for receptive field and tuning properties, and then brought to freely explore a large arena where high-contrast visual stimuli are shown on the wall. We first successfully recapitulated the receptive fields and turning properties of the V1 neurons in the head-restrained scenario. In the freely-moving scenario, we found that the primate V1 population consistently response to gaze shift and fixation with a suppression followed by an enhancement in a sequential latency. A model of gaze response built upon this mechanism well recapitulates the observed data. The effect of locomotion, however, modulates the baseline of neural activity and is speed correlated. To further dissect the neural responses driven by visual input and movement, we performed the experiment in the dark condition. Surprisingly, most responses to gaze movement are disrupted and the effect of locomotion is no longer significant in the dark. Preliminary analyses of visual scenes show a different statistic of visual input induced by locomotion, influencing V1 neurons differently depending on their turning properties. These data are the first to examine the neural basis of active vision in a freely-moving primate and have a significant influence on our conceptions of natural vision.

The research is funded by NIH UF1 NS116377, NIH R01 NS118457, AFOSR 19RT0316, Kavli Institute for Brain and Mind Postdoctoral Award.

Talk 6, 4:00 pm, 54.26

MOTION-INDUCED OBJECT POSITION ADAPTATION IN MACAQUE IT CORTEX: TEMPORAL PROCESSING LIMITATIONS OF NEURAL NETWORKS

Elizaveta Yakubovskaya[1], Hamidreza Ramezanpour[1], Kohitij Kar[1]; [1]York University

Recent studies have demonstrated that the macaque inferior temporal (IT) cortex, a key area in the ventral visual pathway, supports not only object identification but also object position estimation—a function previously attributed to dorsal-stream mechanisms. In parallel, artificial neural networks (ANNs) optimized for object recognition replicate this positional decoding capability. Such findings invite an intriguing question: If these ventral-stream-aligned ANNs can recapitulate positional coding, can they also exhibit systematic positional biases induced by adaptation to motion, analogous to the well-known human aftereffects? We tested this by simulating adaptation in ANNs through the exponential decay of model features (Vinken et al., 2020). Using "brain-mapped" ANN architectures—feedforward convolutional networks (AlexNet, VGG), networks with skip connections (ResNet), and transformer-based models (ViT)—pre-trained on ImageNet, we analyzed responses from their most "IT cortex-like" layers to naturalistic test images. While these ANNs robustly decoded object positions under static conditions, none exhibited motion-adaptation-induced positional shifts. In contrast, when we presented motion-adapting rightward or leftward moving grating stimuli (3000 ms) to two passively fixating macaques and recorded large-scale IT responses to succeeding test images (40 images containing 1 of 8 objects, with varying latent parameters embedded in naturalistic backgrounds), the resulting position decodes from IT population activity showed directionally specific biases matching human perceptual aftereffects. These results suggest that the neural mechanisms in IT supporting adaptive shifts in perceived object position are not fully captured by current ANN models. Therefore, we hypothesized that additional history-dependent, nonlinear transformations might explain these dynamic adaptation effects. Testing a state-of-the-art dynamic video recognition model (SlowFast with ResNet-50 backbone) showed that even this more temporally sophisticated model failed to reproduce adaptation-induced positional aftereffects. Our results reveal a key gap: while current ANNs can decode position, they lack the temporal processing needed to replicate the adaptive positional biases in the macaque ventral stream.

KK was supported by Canada Research Chair Program, Simons Foundation Autism Research Initiative (SFARI, 967073), Brain Canada Foundation (2023-0259), CFREF (VISTA Program), and a

NSERC DG. EY was supported by NSERC CGS-M and CFREF (Connected Minds).

TALK 7, 4:15 PM, 54.27

THE MODULATORY EFFECT OF INTENTIONALITY ON NEURAL TUNING TO PERCEPTUAL CUES OF INTERACTIVITY

Sajjad Torabian[1], John A Pyles[2], Hongjing Lu[3], Emily D Grossman[1]; [1]University of California, Irvine, [2]University of Washington, [3]University of California, Los Angeles

Introduction. Humans share strong intuitions about interactive behaviors that imply intentions and goals. Perceived intentionality exists on a spectrum ranging from minimal intentionality to clear attributions of goal-directedness. Here, we use a computational model that extracts two key attributes of interactions as conveyed in decontextualized animations: the degree of intentionality and the extent to which the movements violate the laws of physics (Shu et al., 2021). We then evaluate how each of these high-level attributes, and those of perceptual features, best explain the neural representational similarity structure in the multivariate patterns within social cognitive brain systems. Methods. In a rapid event-related fMRI design, participants rated short animations (3s each) of two moving shapes as either interacting agents or physical objects. Each animation was characterized by the degree of interactivity conveyed in perceptual features, such as distance between the shapes, as well as low-level properties (e.g. speed), and high-level attributes indexing the degree of intentionality and violations of physics. The similarity structure derived from feature scores were then compared against the neural similarity elicited by the animations. Results. Each parcel in the superior temporal sulcus (STS) reflected a unique weighting of features defining interactivity, perceptual properties, or human ratings that accounted for variance in similarity structure. The attribute capturing the extent of violations of physics best explained neural activity in the anterior intraparietal sulcus (AIP). Considering only the similarity of the agentic interactions revealed the similarity structure in the same parcels in the STS (but not the AIP) to be further weighted by the degree of intentionality. Conclusion. Together these results demonstrate neural tuning for features that convey interactivity are modulated by the extent to which those features convey intentionality. These results are consistent with a model in which the salience of social behaviors is enhanced by the implied goal-directedness.

BCS-1658560 to EG & BCS-1658078 to JP & NSF BCS 2142269 to HL

TALK 8, 4:30 PM, 54.28

FOR MSTD AUTOENCODING IS ALL YOU NEED

Oliver Layton[1], Scott Steinmetz[2]; [1]Colby College, [2]Sandia National Labs

Neurons in brain area MSTd demonstrate tuning to optic flow patterns that resemble those encountered during navigation. Computational neural models have been developed to elucidate the link between MSTd and self-motion perception, most of which possess a hierarchical design wherein MSTd-like optic flow tuning emerges after successive stages. For example, models tend to contain a V1-like stage in which local motion is extracted, a MT-like stage in which direction and speed is estimated over extended regions of space, and a MSTd-like stage that contains tuning to optic flow patterns. Deep neural networks (DNNs) likewise adopt this hierarchical design and when trained to accurately classify natural images they have been successful at modeling brain areas along the primate ventral stream. When we trained DNNs to perform the analogous task of accurately estimating self-motion from optic flow (Layton & Steinmetz, 2024), interestingly, we found poor correspondence with MSTd optic flow tuning properties compared to the simpler non-negative matrix factorization (NNMF) model of Beyeler et al. (2016). Rather than attempting to accurately estimate self-motion, NNMF reconstructs MT motion inputs from a low dimensional representation. To determine whether this difference in computational objective accounts for the discrepancy, here we investigate whether MSTd-like optic flow tuning emerges in autoencoders, artificial neural networks that share a common computational objective with NNMF. While we find that autoencoders produce more MSTd-like tuning than accuracy-optimized DNNs, the correspondence with MSTd is weaker than with NNMF. Training autoencoders on a MT-like representation rather than a motion vector representation of optic flow substantially improves the alignment with MSTd properties. Making the same adjustment to accuracy-optimized DNNs does not improve the correspondence. Our results suggest that the computational objective of autoencoders aligns more closely with that of MSTd and the motion representation in MT may critically shape optic flow tuning in MSTd.

TALK SESSION: TUESDAY, MAY 20, 2025, 5:30 – 7:15 PM, TALK ROOM 1

Eye Movements: Perceptual advantages and disadvantages

Moderator: David Melcher, NYU Abu Dhabi

TALK 1, 5:30 PM, 55.11

EXPLORING THE LIMITS TO THE "ILLUSION OF RELATIVE STABILITY" OF IMAGES SLIPPING CONSISTENT TO FIXATIONAL EYE MOTION

Josephine C. D'Angelo[1] (josephine_dangelo@berkeley.edu), Pavan Tiruveedhula[1], Raymond J. Weber[2], David W. Arathorn[2], Jorge Otero-Millan[1], Austin Roorda[1]; [1]University of California, Berkeley, [2]Montana State University

The human visual system is exquisitely sensitive to detecting relative motion; this ability requires the presence of world-fixed images in the scene which serve as frames of reference. However, images moving in the direction consistent with retinal slip appear to have little to no motion relative to world-fixed retinal image background content, suggesting that under this special condition perceptual stabilization overrides motion sensitivity (1). This phenomenon was called the "Illusion of relative stability." We aimed to explore the limits to this illusion by measuring the perceived motion of images slipping opposite to and in the same direction as eye motion with magnitudes closer to the transition between directions. For example, would a stimulus moving in a direction consistent with retinal slip but only slipping with 10% of the retinal motion of a world-fixed stimulus also appear relatively stable? We used an adaptive optics scanning light ophthalmoscope to present stimuli that moved contingent to fixational eye motion. To quantify the perceived motion of these stimuli, we presented a second stimulus that moved in a random trajectory and the subjects adjusted its magnitude of motion until it appeared perceptually equivalent to the retina-contingent stimulus. We found that images slipping opposite to eye motion appear less in motion while images moving in the same direction appear to have a higher magnitude of motion which on average is similar to the image's magnitude of world motion. Even stimuli slipping with only 10% of the retinal motion of a world-fixed stimulus appear relatively stable. These results suggest a discontinuity in motion perception which is dependent on the direction that images move contingent to fixational eye motion. (1) D'Angelo, Josephine C et al. "A paradoxical misperception of relative motion." Proc. Natl. Acad. Sci. U.S.A. vol. 121,48 (2024): e2410755121. doi:10.1073/pnas.2410755121

NIH R01EY023591; NIH T32EY007043; NIH R00EY027846; Berkeley Center for Innovation in Vision and Optics

Talk 2, 5:45 pm, 55.12

TEMPORAL PREDICTIONS UNDERLIE THE EXTRA-FOVEAL PREVIEW EFFECT ACROSS SACCADES

David Melcher[1], Michele Deodato[1]; [1]New York University Abu Dhabi

In natural viewing, we look around objects in the scene. Thus, visual processing of objects typically involves a time course in which we make a saccadic eye movement to a target object that was first viewed, and selected, using extra-foveal vision. Previous studies have shown that this extra-foveal preview can strongly influence perceptual judgments about that stimulus after it is fixated. However, the mechanisms underlying trans-saccadic perception remains a matter of debate. Here, participants performed a gaze-contingent task in which an extrafoveal saccade target stimulus (the preview) was replaced with a tilted stimulus (the target) during a saccade directed to it. On separate trials, the preview stimulus was either identical to the target (valid preview) or different (invalid preview). We ran multiple experiments, with invalid previews consisting of faces with different identities, inverted versus upright faces, or faces versus houses. Critically, on some trials we added a brief, blank delay at the beginning of the post-saccadic fixation period, before the appearance of the target. The added temporal gap at the beginning of the new fixation dramatically modulated the preview effect, leading to an elimination or even a reversal of the usual preview validity benefit. Overall, these findings suggest that saccadic programming and processing of the extrafoveal saccade target creates a category-level and temporal prediction which underlies the parafoveal preview effects in trans-saccadic perception.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYU Abu Dhabi Research Institute grant CG012. Part of the work was conducted at the Brain Imaging lab within the Core Technology Platforms at NYU Abu Dhabi

Talk 3, 6:00 pm, 55.13

VOXEL-WISE PREDICTIVE ENCODING MODELS REVEAL EVIDENCE FOR PRE-SACCADIC REMAPPING IN THE HUMAN VISUAL CORTEX

Yong Min Choi[1] (choi.1696@osu.edu), Julie D. Golomb[1]; [1]The Ohio State University

Retinal visual inputs shift drastically across saccadic eye movements. To align pre- and post-saccadic visual information and aid perceptual stability, neurons' receptive fields (RFs) are predictively remapped in anticipation of upcoming saccades toward the future RF location (forward remapping) or the saccade target location (convergent remapping). Despite extensive evidence of pre-saccadic RF remapping in non-human primates, there remain open debates about the nature of remapping in different brain areas, and, more generally how remapping operates in the human visual system. In the current study, we developed a novel fMRI paradigm to investigate pre-saccadic RF remapping in the human brain using a hypothesis-driven, voxel-wise predictive encoding model approach. Adult participants completed two fMRI sessions: (1) a static population receptive field (pRF) mapping session to estimate voxel-wise static pRFs, and (2) a main experiment session presenting noise-like visual stimuli either during stable fixation or during saccade preparation. Using each voxel's estimated static pRF, we constructed different hypothetical models of pre-saccadic remapping (e.g., "no predictive remapping", "forward remapping", "convergent remapping") and evaluated how well each model predicted actual neural responses in the main experiment session. During stable fixation, as expected, the neural activity of most voxels in visual areas was best predicted by models assuming no predictive remapping. Critically, during saccade preparation, many voxels showed improved predictability for models incorporating forward and/or convergent RF remapping. These findings provide novel evidence for pre-saccadic remapping at a voxel-level across the human visual cortex, with this new approach offering exciting potential to broaden our understanding of how RF remapping operates across the brain and links to behavior.

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

Talk 4, 6:15 pm, 55.14

PRE-**SACCADIC ATTENTION IN PARKINSON'S DISEASE:** EFFECTS OF AGING AND DOPAMINE LEVEL

Oliver L. Steiner[1,2,3] (oliverr.e.steiner@hotmail.de), Nina M. Hanning[3,5], Sarah Melchert[1], Sven-Florian Jaeger[1], Fabian Klostermann[1,2], Martin Rolfs[2,3,4]; [1]Klinik für Neurologie mit Experimenteller Neurologie, Charité – Universitätsmedizin Berlin, Deutschland, [2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Deutschland, [3]Institut für Psychologie, Humboldt-Universität zu Berlin, Deutschland, [4]Bernstein Center for Computational Neuroscience, Berlin, Deutschland, [5]Department of Psychology and Center for Neural Science, New York University, New York, NY 10012, USA

Navigating dynamic and unpredictable environments requires allocating attentional resources for impending movement. While

mechanisms of pre-saccadic attention – the anticipatory enhancement of sensory processing at an imminent saccade target – have been characterized in cortical networks, the role of subcortical structures remains largely unexplored. Based on the important role of the basal ganglia for attention allocation and movement preparation, we hypothesized that dopamine depletion in this subcortical network disrupts the temporal dynamics of pre-saccadic attention. To test this, we assessed patients with Parkinson's Disease (PD) – a neurological movement disorder characterized by dopamine degeneration in the basal ganglia – both ON and OFF dopaminergic medication. We compared their performance to age-matched controls and young observers. Participants fixated on a central bull's-eye while two placeholders, each consisting of four black dots, were positioned 10° left and right of fixation. Upon stable fixation, a central line cue indicated the saccade target. Participants were instructed to "look as fast and precisely as possible" at the indicated placeholder. At varying times from approximately 200 ms before saccade execution up to saccade onset, we presented ±45 degree oriented 1/f noise for 83.3 ms, embedded in a dynamic stream of random 1/f noise. Once participants had executed the saccade, they verbally reported the stimulus orientation (clockwise or counterclockwise from vertical).
Both elderly controls and PD patients exhibited prolonged saccadic latencies compared to young observers, accompanied by a flatter slope of attentional enhancement approaching saccade onset. PD patients showed a reduced pre-saccadic attentional benefit immediately before saccade execution. No significant differences were observed between ON and OFF medication in PD patients. Collectively, our data delineate the temporal dynamics of pre-saccadic attention in PD. Our findings suggest that aging reduces the build-up of pre-saccadic attention. This process is exacerbated by PD, but without a detectable influence of dopamine.

Talk 5, 6:30 pm, 55.15

FIXATIONAL EYE MOVEMENTS AND VISUAL ACUITY IN PATIENTS WITH SCHIZOPHRENIA

Sanjana Kapisthalam[1,2] (skapisth@ur.rochester.edu), Howard Bi[1,2], Youjia Zhang[1], Ashley M. Clark[1,2], Judy L. Thompson[1], Martina Poletti[1,2], Brian P. Keane[1,2]; [1]University of Rochester, [2]Center for vision science

In non-clinical populations, visual acuity and oculomotor behavior at fixation are tightly linked. For the first time, we leveraged high-precision eye tracking to determine whether fixational eye movements are abnormal in schizophrenia and whether such abnormalities may contribute to reduced visual acuity. Visual acuity was first assessed with habitual correction in 14 patients and 13 age-matched healthy controls, none with ocular pathologies or worse than 20/20 Snellen acuity. To record eye movements with high precision, participants were asked to view stimuli without their habitual visual correction and were best corrected using a Badal Optometer. Subjects performed a 4AFC acuity task. Stimuli consisted of digits in Pelli's font, whose size changed based on an adaptive staircase. Each trial began with a 400-ms blank screen followed by a 500-ms stimulus presentation. Although both groups averaged near 20/20, patients exhibited poorer Snellen acuity (20/20 vs. 20/16; $p = 0.01$, $d = 1.06$). Visual acuity thresholds measured with the adaptive procedure were slightly elevated but not significantly different in patients compared to controls ($p = 0.3$, Cohen's $d = 0.52$). However, patients exhibited significantly greater between-subject variability in these thresholds ($p = 0.03$, Levene's test). Our findings also revealed that fixation stability, quantified using the bivariate contour ellipse area, was reduced in patients (0.12± 0.03 deg^2 (patients) vs. 0.08±0.04 deg^2 (controls); $p = 0.005$, $d = 0.98$). Further, whereas healthy controls effectively suppressed microsaccades during stimulus presentation, patients did not and were characterized by a higher microsaccade rate (0.3±0.25 ms/s (controls) vs. 0.88±0.46 ms/s (patients), $p = 0.003$, $d = 1.3$). These findings show that patients with schizophrenia exhibit worse Snellen acuity, less stable fixation, and a higher rate of microsaccades, highlighting the possibility that these ocular abnormalities may contribute to an overall worse acuity.

NIH R21EY035001; Schmitt Program in Integrative Neuroscience (through Del Monte Institute at URMC)

Talk 6, 6:45 pm, 55.16

SACCADIC SUPPRESSION IN AREA MT/MTC IS ABSENT DURING SIMULATED SACCADES IN THE VISUAL INPUT.

Amy Bucklaew[1], Shanna Coop[2], Jude Mitchell[2]; [1]University of Rochester, Neuroscience, [2]University of Rochester, Brain and Cognitive Sciences

To maintain a stable percept of the world, visual information during saccades is suppressed to avoid retinal blur, a phenomenon termed saccadic suppression. Extra-retinal motor signals, potentially originating from efference copy feedback from frontal eye fields (FEF) or superior colliculus (SC), are thought to contribute to this suppression (Sommer & Wurtz, 2008). More recent studies show that saccadic suppression can start as early as the retina from wide-field rapid visual flow that simulates the visual input during saccades (Indrees et al., 2020; Baumann & Hafed 2024). In the present study we sought to clarify the involvement of suppression from the retina versus efference copy feedback at the level of visual cortex. Marmoset monkeys freely viewed either natural images, while in the dark, or while saccades were simulated by intermittently moving viewed natural images. We recorded from motion-selective areas MT and MTC. Previously, we reported that a subpopulation of neurons (18%), predominantly found in area MTC more so than MT, respond with early transients to saccades that are tuned for saccade direction and are faster than conventional visual response latencies (25-35ms vs 40-50ms). These short-latency responses could reflect an extra-retinal signal involved in driving saccadic suppression (Bucklaew et al., 2023). Here we tested if these short-latency responses could instead reflect a response to rapid full-field visual flow by simulating saccades with moving natural images and comparing it against saccades in free-viewing of natural images. We find that short-latency responses are absent with simulated saccades, as well as signs of the typical saccadic suppression in the rest of the population. Short-latency responses in the dark were highly attenuated, suggesting a modulatory rather than driving role for potential extra-retinal inputs. These results suggest extra-retinal feedback likely plays the key role in driving saccadic suppression in visual cortex.

Funding: AB, SC, and JFM from NIH EY030998, AB from NIH T32EY007125 and NIH F31EY035866

TALK 7, 7:00 PM, 55.17

DARK CONTRASTS ARE IMMUNE TO SACCADIC SUPPRESSION IN THE PRIMARY VISUAL CORTEX

Wenbin Wu[1] (wenbin.wu@student.uni-tuebingen.de), Yue Yu[1], Tatiana Malevich[1], Matthias P. Baumann[1], Tong Zhang[1], Carlotta Trottenberg[1], Ziad M. Hafed[1]; [1]University Tübingen

Saccade generation is accompanied by a dramatic reduction in perceptual sensitivity for perimovement stimulus onsets. Neuronal correlates of this phenomenon have been observed in multiple brain areas, including the retina, superior colliculus (SC), and primary visual cortex (V1). However, how each area specifically contributes to the perceptual effect itself remains unknown. Here, we were motivated by previous observations that perisaccadic perceptual detectability is similarly impaired for dark and bright stimuli, and that this is also true in the SC (Wu & Hafed, 2024). We asked whether these observations are universal (and thus observable in other brain areas), or whether they reveal a particular role for the SC in mediating them. We recorded from 248 SC neurons (three monkeys) and 325 V1 neurons (two monkeys). In each trial, a disc (0.51 deg radius; bright or dark) appeared within the recorded neurons' response fields. We measured stimulus-evoked visual response strength as a function of stimulus onset time relative to microsaccades. In the SC, we replicated the earlier observations (Wu & Hafed, 2024) that suppression strength is similar for darks and brights. Surprisingly, this was not the case in V1: all of our dark contrasts were completely immune to saccadic suppression. Moreover, bright contrasts underwent weaker suppression than in the SC. These results suggest that perisaccadic perceptual detectability (similarly suppressed for darks and brights) is not mediated by V1. However, this does not mean a complete lack of V1 impact on perisaccadic vision. In five human subjects, we repeated the same experiments but now presented supra-threshold oriented bars (either dark or bright). The bars (at 20% contrast) were perisaccadically detectable in 100% of the trials. Nonetheless, orientation discrimination thresholds were still elevated, but only for the bright stimuli. Thus, there is a highly selective saccadic suppression of exclusively ON processing pathways in V1.

TALK SESSION: TUESDAY, MAY 20, 2025, 5:30 – 7:15 PM, TALK ROOM 2

Face and Body Perception: Development, clinical, individual differences, experience

Moderator: Adrian Nestor, University of Toronto

TALK 1, 5:30 PM, 55.21

AN AGE BIAS IN OTHER-RACE FACE PERCEPTION: NEURAL AND BEHAVIORAL EVIDENCE

Moaz Shoura[1], Yong Zhong Liang[1], Dirk B. Walther[1], Adrian Nestor[1]; [1]University of Toronto

The other-race effect (ORE) refers to a well-documented disadvantage in recognizing faces of other races compared to one's own. Yet, the perceptual biases underlying ORE remain less understood. This research combines behavioral similarity measures, neural decoding and style-based generative adversarial networks (StyleGAN2; Karras et al., 2020) to examine visual biases, with focus on age misrepresentation in other-race (OR) face perception. To this end, first, East Asian and White participants (n = 190) rated the pairwise visual similarity of OR and same-race (SR) synthetically generated faces. The similarity structure of behavioral data was mapped onto that of StyleGAN2 latent representations. Relying on this mapping, we employed a novel procedure for GAN-based image reconstruction to recover SR and OR face percepts across our participants. The procedure generated hyper-realistic visualizations of face percepts. More importantly, facial attribute analysis, using RetinaFace (Deng et al., 2019), revealed that OR image reconstructions appeared systematically younger than SR ones. Second, electroencephalography (EEG) data were collected from East Asian and White participants (n = 40) who viewed OR and SR faces. Neural representations of these faces, recovered through EEG-based face decoding and image reconstruction (Nemrodov et al., 2018), were subsequently rated by another group of East Asian and White validators (n = 46) for perceived age and typicality. OR face reconstructions were consistently rated as younger and more typical for their race than SR reconstructions across both validator groups. Together, these findings provide convergent evidence for age misrepresentation in OR face perception across different data types and different reconstruction methodologies. These results open new avenues for investigating representational differences in face perception and highlight new implications for cross-racial interactions.

TALK 2, 5:45 PM, 55.22

CATEGORY SELECTIVITY OBSERVED IN THE HUMAN FFA, PPA, AND EBA IS DISTINCT FROM CATEGORY SELECTIVITY OBSERVED IN ARTIFICIAL NEURAL NETWORK UNITS

Alish Dipani[1,2] (alish.dipani@gmail.com), N Apurva Ratan Murty[1,2]; [1]Cognition and Brain Science, School of Psychology, Georgia Institute of Technology, [2]Center of Excellence in Computational Cognition, Georgia Institute of Technology

Category-selective responses to faces, scenes, and bodies have been robustly observed in the human brain (e.g. FFA, PPA, and EBA) and more recently, even in artificial neural network (ANN) models. This raises a key question: How similar are the category-selective responses observed in the human brain compared to the category-selective responses seen in vision ANNs? To address this question, we first identified the category-selective voxels and ANN units following an identical localization procedure. Consistent with prior findings, we successfully isolated the category-selective voxels and ANN model units. Next, we compared the voxel-averaged response patterns in the FFA, PPA, and EBA (from the Natural Scenes Dataset) with the unit-averaged response patterns observed in category-selective ANN units. Response patterns observed in the brain were remarkably consistent across participants (median subject-subject R = 0.68). In contrast, the response patterns of category-selective ANN units were significantly distinct from those observed in category-selective voxels (median subject-ANN R = 0.38). This divergence between ANN units and brain regions was robust: it persisted for all ANN models evaluated (N=13) and across different localizers (N=3), generalized to multivariate tests between model units and brains, and held across all brain regions tested. Finally, we asked whether category-selective ANN units could be linearly combined (encoding model) to predict responses in human category-selective regions. Surprisingly, while the encoding models had high predictivity, lesioning the category-selective ANN units did not significantly impact the model's prediction accuracy. This finding is further evidence that category-selective ANN units may not directly correspond to human category-selective brain regions. Together, these results underscore the need for caution: not all forms of category selectivity can be considered equivalent. The category selectivity observed in FFA, PPA, and EBA is a specific kind, fundamentally distinct from the category selectivity exhibited in ANN units.

This work is supported by a CoCo fellowship by the School of Psychology, Georgia Tech to AD and by the NIH Pathway to Independence Award (R00EY032603), and a startup grant from Georgia Tech to NARM.

Talk 3, 6:00 pm, 55.23

FROM EYES TO IDENTITY: EARLY NEURAL SENSITIVITY TO THE EYE REGION IS ASSOCIATED WITH FACE RECOGNITION ABILITY

Anthony Proulx[1] (anthonyproulx21@gmail.com), Isabelle Charbonneau[1], Justin Duncan[2], Vicki Ledrou-Paquet[1], Chanelle Demeule[1], Caroline Blais[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais, [2]Ottawa University

Recognizing faces is a fundamental social skill that varies widely among individuals. While most people identify faces effortlessly, others experience significant difficulties, with extreme cases characterized as prosopagnosia. Research on individual differences and prosopagnosia has emphasized the importance of the eye region in face identification (Caldara et al., 2005; Tardif et al., 2019), with greater reliance on the eyes predicting better face recognition ability (Royer et al., 2018). However, this association may arise from different underlying mechanisms. It is possible that the superior use of the eyes results from more efficient perceptual processing, better facial memory, or a combination of both. To explore these hypotheses, 36 participants completed a series of face and object recognition tasks, with a principal component analysis (PCA) generating a global face recognition ability score. Participants then performed a face identification task while their brain activity was recorded using EEG. Faces were presented through small Gaussian apertures ("Bubbles"; Gosselin & Schyns, 2001), and classification images were generated from -200ms to 800ms post-stimulus, revealing pixels associated with stronger or weaker voltage at PO8. These images were aligned with each participant's N170 and N250 latencies and combined to capture feature processing within these time windows. Weighted averages based on PCA scores highlighted individual differences in classification images relative to face recognition ability. Pixel tests ($p<.05$; Stat4Ci Toolbox; Chauvin et al., 2005) on these weighted images revealed that better face recognition ability was significantly associated with increased sensitivity to both the contra- and ipsilateral eyes shortly before the N170, and with increased sensitivity to the contralateral eye before the N250. These results suggest that better face recognizers make superior use of the eye region early in perceptual processing, which also manifests as more effective utilization of the contralateral eye when activating memory representations.

Talk 4, 6:15 pm, 55.24

PREDICTING INDIVIDUALIZED CATEGORY-SELECTIVE FUNCTIONAL TOPOGRAPHIES IN DEVELOPMENTAL PROSOPAGNOSIA USING CONNECTIVITY HYPERALIGNMENT

Ian Abenes[1] (ian.abenes@gmail.com), Jiahui Guo; [1]The University of Texas at Dallas

Category-selective functional topographies are mostly similar across individuals, but considerable variability exists in the exact topographical location, size, and shape of these areas. It has been demonstrated that individualized category-selective topographies can be estimated with high fidelity using hyperalignment (Jiahui et al., 2020, Jiahui et al., 2023). However, previous work only included typical participants, and it is unclear whether this method can be extended to neuropsychological populations. To address this, we analyzed data from 12 individuals with developmental prosopagnosia (DP), who display profound face recognition deficits in the absence of brain lesions or broad neurodevelopmental problems, alongside 16 typical participants. We first estimated functional topographies of four categories: faces, bodies, objects, and scenes from a classic dynamic localizer. We then used connectivity hyperalignment (CHA) to derive transformation matrices for each individual using a separate functional scan from other participants within their group. This scan was taken while participants completed a one-back task to identify if the previously shown face was the same in identity, expression, or view. We applied the transformations back to the localizer data to project all other participants' localizer data to the given individual's space. The estimated topographies using CHA for the given individual were the mean across all other participants' contrast maps. Finally, we estimated the topographies using anatomical alignment (AA) by directly averaging contrast maps based on surface-aligned localizer runs within each group. We correlated each participant's own localizer map with the CHA or AA estimated maps to measure the performance of the predictions. On average, CHA showed higher correlations than AA for estimating individual face-selective topography, and similar results were found for other categories. Notably, both DP and typical participants showed similar high-fidelity estimations, suggesting successful estimation of category-selective topographies using connectivity hyperalignment, despite significant impairments in face recognition abilities in the DP individuals.

Talk 5, 6:30 pm, 55.25

CATEGORY SELECTIVITY AS AN EXPLANATION FOR MULTIDIMENSIONAL TUNING IN HUMAN OCCIPITOTEMPORAL CORTEX

Hans Op de Beeck[1], Elahe' Yargholi[1]; [1]KU Leuven, Leuven Brain Institute, Dpt. Brain & Cognition, Leuven, Belgium

Recent studies have revealed an exceptionally rich landscape of functional selectivity in human occipitotemporal cortex. While the first seminal brain imaging studies of object recognition emphasized the existence of category-selective regions, alternative perspectives have emerged. Explicit comparisons suggest that dimensions outperform categories in predicting brain responses (Contier et al., 2024, Nature Human Behavior), and that the tuning of category-selective neurons can be explained by domain-general tuning for nonface dimensions (Vinken et al., 2023, Science Advances). This begs the question whether the notion of category selectivity is outdated and should be abandoned in favor of a dimensional view (Ritchie et al., 2024, arXiv). Here we review existing findings and present new findings that illustrate the power of the notion of category selectivity to explain object representations and the observed dimensional tuning profiles. Previous studies suggest that the most important dimensions in the literature, animacy and stubby/spiky, could possibly be attributed to category selectivity. Furthermore, in a new fMRI study (N = 22) with participants viewing social scenes that all contain the same categories, we find a strong modulation of neural responses in face-, body-, and scene-selective regions by variation in how categories are depicted in natural images, consistent with realistic category-based models. Overall, the notion of category selectivity provides a parsimonious explanation for the strongest forms of selectivity which are observed in human occipitotemporal cortex.

METH/24/003

Talk 6, 6:45 pm, 55.26

A NEW COMPUTATIONAL MODEL OF HUMAN FACE RECOGNITION THAT LEARNS CONTINUOUSLY BY GENERATING IMAGES FROM MEMORY

Naphtali Abudarham[1], Galit Yovel[1]; [1]Tel Aviv University

In recent years, deep convolutional neural networks (DCNNs) have reached human-level performance in face recognition and have been shown to exhibit human-like face effects including the face inversion effect, the other race effect, and sensitivity to human-like critical features. Nevertheless, there are still significant discrepancies between the operations of DCNNs and the human face system. One fundamental discrepancy is the training regime – DCNNs are trained in batch mode, learning the whole training set at once, while humans learn to recognize different identities gradually over time. Another discrepancy is that humans can learn and recognize new faces instantly after a single exposure, whereas training DCNNs on new identities requires extensive exposure to each identity and may lead to catastrophic forgetting of previously learned identities. To bridge these gaps, we propose a novel computational cognitive model for human face recognition. Our model is composed of three main components: 1. An embedder, which is trained incrementally on a small number of faces in each step, using a Continual Learning paradigm and is used to create face representations. 2. A face-generator module, which is trained in parallel to the embedder and is used to generate images of old faces for training the embedder concurrently with new faces. 3. A memory module, which stores embedding statistics of familiar/trained identities, enabling the recognition of familiar faces using a nearest-neighbor search, and instant learning of new identities. We show how this model can learn and recognize faces nearly as good as a batch model, even when it is trained on a very small number of identities. Furthermore, our model shows a human-like familiarity benefit on a face sorting task. Our model proposes a comprehensive and realistic approach to human face recognition, which can be expanded beyond faces to study mechanisms of human memory.

Talk 7, 7:00 pm, 55.27

EXAMINING THE FUNCTIONAL ORGANIZATION OF MARMOSET INFEROTEMPORAL CORTEX USING CALCIUM IMAGING

David G. C. Hildebrand[1], Santiago Otero-Coronel[1], Alipasha Vaziri[1], Winrich Freiwald[1]; [1]Rockefeller University

A characteristic feature of sensory cortex in primates is its functional organization into continuous maps, where neurons are arranged according to their functional properties (e.g., retinotopy, orientation pinwheels). In contrast, high-level visual areas in inferotemporal (IT) cortex have been parcellated into discrete category-selective clusters (e.g., face vs. body areas). However, recent results and models suggest that IT cortex could instead be functionally organized by continuous maps corresponding to properties such as animacy or real-world size. It has been difficult to resolve the functional organization of IT cortex with standard techniques; fMRI has insufficient spatial resolution, while electrophysiological sampling is too sparse. To examine the functional organization of IT cortex, we developed an approach for cellular-resolution imaging of neuronal populations in marmosets. The marmoset cortex is smooth, making it accessible with optical imaging. Leveraging this feature, we localized face areas using intrinsic signal imaging. We then recorded calcium dynamics of cortical layer 2/3 neurons in awake, head-restrained marmosets using two-photon microscopy with fields of view up to 3×3 mm2 while presenting faces, non-face objects, and bodies. This approach allowed us to sample thousands of IT cortical neurons in and around face areas. While neurons in face areas responded more on average to faces, the stimulus eliciting the peak response from each neuron revealed a striking diversity across the population. Nearby neurons tended to respond more to similar subsets of face and non-face stimuli. Furthermore, we observed gradients of selectivity between face, object, or body areas rather than discrete boundaries, supporting the conclusion that IT cortex is functionally organized in a continuous map.

NIH R21 EY031486

Poster Sessions

FRIDAY AFTERNOON POSTERS IN BANYAN BREEZEWAY

Visual Search: Memory

Friday, May 16, 3:00 – 5:00 pm, Banyan Breezeway

16.301 ATTENTIONAL GUIDANCE FROM NEGATIVE TEMPLATES IN VISUAL WORKING MEMORY IS IMPLEMENTED INDIRECTLY VIA SPATIAL RECODING

Aditya Prakash[1], Andrew Hollingworth[1]; [1]University of Iowa

Attention can be guided by the content of visual working memory (VWM), often referred to as a search template. These templates can be positive (specifying feature values associated with the target) or negative (specifying feature values associated with distractors). In addition, the implementation of a search template can be either direct or indirect. For example, direct implementation of a cue to ignore red would be executed by filtering stimuli on the dimension of color to reduce the relative priority of red items. In contrast, guidance could be implemented indirectly by marking, and subsequently de-prioritizing, the spatial locations of red items. Beck & Hollingworth (2015) proposed that negative templates maintained in VWM control attention indirectly through a mechanism of spatial recoding. This account predicts that the implementation of a negative template will depend on the ease of spatial recoding, with efficient avoidance when cued items are spatially segregated and less efficient avoidance when cued items are spatially intermixed. Here, we examined this prediction for oculomotor behavior in a search task employing negative color cues. Participants searched through arrays of objects drawn in multiple colors following a negative or neutral color cue. Consistent with previous work, we found that the probability of fixating negative-cue-matching objects was significantly below baseline. Critically, the efficiency of avoidance was influenced by a manipulation of spatial intermixing, with a lower probability of fixating negatively cued items when the level of spatial intermixing was lower. As argued by Beck & Hollingworth, direct implementation of a negative template may not be possible, as maintenance of the color in VWM leads to the obligatory enhancement of matching items. Instead, guidance by VWM-based negative cues appears to be implemented indirectly and reactively through the translation of the color cue into a spatial template based on the observed locations of cue-matching items.

16.302 CONCURRENT SELECTION OF INTERNAL SEARCH GOALS AND EXTERNAL VISUAL TARGETS DURING VISUAL WORKING-MEMORY GUIDED SEARCH

Baiwei Liu[1] (b.liu@vu.nl), Freek van Ede[1]; [1]Vrije Universiteit Amsterdam

When searching for objects, we continuously navigate between internal search goals held in memory and the external sensory world containing the objects we search. Yet, little remains known about the dynamic processes by which internal and external search processes combine to support efficient search. One key open question is whether internal and external search operations necessarily take turns (completing selection of the relevant search goal before searching it) or can proceed concurrently (guiding search from the moment the process of search-goal selection is initiated). To address this, we studied the dynamics that govern internal search-goal and external target selection during working-memory guided search. Participants held two potential search goals in visual working memory and were cued to find one of them in a search display that was presented simultaneously with the memory cue. To isolate internal and external selection, we independently varied the location of the memorised internal search goal and the to-be-found external visual target. We then leveraged spatial modulations in gaze and EEG activity to track the dynamics of internal (search-goal) and external (target) selection through time. Our findings reveal how internal and external selection during working-memory guided search do not necessarily take turns, but can proceed concurrently. This provides unprecedented insight into how internal and external visual operations combine to yield efficient search.

This work was supported by an ERC Starting Grant from the European Research Council (MEMTICIPATION, 850636) and an NWO Vidi Grant from the Dutch Research Council (14721) to F.v.E.

16.303 SEARCHING BETWEEN WORKING-MEMORY VERSUS VISUAL ARRAYS

Dengxinyi Wei[1] (dxywei11@gmail.com), Daniela Gresch[1], Anna C. Nobre[1]; [1]Yale University

Selective attention helps us search for relevant items in the external world and in memory. How external and internal search processes compare and how they are coordinated to guide behavior remain open questions. We developed a new experimental design to compare the properties of visual search in perceptual versus working-memory arrays and to reveal how search across the domains is prioritized. Forty-nine healthy young adults were recruited online to perform a combined perceptual and working-memory search task, looking for a target item that could be present either in a previously encoded array or within a visible array. At the start of a trial, four quadrant placeholders appeared, and two items briefly occupied two of the locations (500 ms) before being replaced by the placeholders again (800 ms). Two other items then occupied the other two locations, and a central word designated the target for search. Participants indicated whether the target item was present in either their WM or the visual array (yes/no) and were provided with feedback (correct/incorrect). Targets were equally likely to be present or absent and, when present, to be in the WM or visual array. Accuracy was systematically higher for targets present in the visual display, whereas reaction times were faster for target items present in WM. Ongoing studies are probing the generality of the effects across different types of cues (verbal, object-based, feature-based) and search items (real objects, abstract colored shapes) and using models to test the functional parameters that guide the competition between external and internal search.

16.304 THE MISSING PIECE OF THE PUZZLE: PREDICTIVE TEMPLATES GUIDE VISUAL SEARCH BY PROMPTING THE GENERATION OF TARGET FEATURES IN VISUAL WORKING MEMORY.

Jennifer Bartlett[1] (jdbartlett@mun.ca), Blaire Dube[1]; [1]Memorial University of Newfoundland

During visual search, a representation of the search target (an attentional template) held in visual working memory (VWM) supports the automatic guidance of attention to feature-matching perceptual inputs. Traditionally, attentional templates represent the exact features of an object; however, sometimes search relies on a template that is not an exact match to the search target but instead predicts target features. For example, when doing a puzzle, the piece you are looking at is not an exact template for the target piece you are looking for. What do we represent in VWM when relying on a predictive template? We evaluated whether participants represent the predictive template in VWM to compare with perceptual inputs or if they generate the target features in VWM to guide search directly. In two experiments, participants first viewed a predictive template (a solid black puzzle piece (E1) with an arrow indicating the position of the connecting piece they were to search for). Most trials then ended in a search for the target (connecting) piece. Critically, 40% of trials instead ended in an unrelated search task that contained a puzzle-piece distractor meant to be ignored. This distractor could either match the predictive template, the target (connecting) piece, or be novel to the trial. This distractor induced memory-driven capture (a slowing of response times when an item matches the contents of VWM) only when it matched the target, suggesting that a representation of the target features was active in VWM in preparation for search. In E2, we replicated this result with puzzle-piece stimuli that had complex designs overlaid on them and thus required pattern extension in order to generate target features. We suggest that predictive templates prompt the generation of target features in VWM to directly guide search, both when the target features are simple and complex.

16.305 POPOUT AND OBJECT RE-IDENTIFICATION

Avery Caulfield[1], Justin Halberda[2], E. J. Green[3]; [1]Johns Hopkins University

The phenomenon of pop-out is commonplace in everyday experience and fundamental to studying the interface between early and mid-level vision (e.g., object-based representation). One question concerns how the visual system prioritizes features following pop-out. Does the feature that caused the pop-out get represented with higher priority or fidelity compared to the object's other features? For instance, one might perceive a colored rectangular object among colored disks and perceptually represent, "that rectangle (pop-out feature) over there just happens to be red (ancillary feature)" Furthermore, given its task of reidentifying objects over time, it makes sense that the visual system would prioritize pop-out features. For example, if an object's color is distinctive but its shape is not, then it can be reidentified more accurately using color. To explore the perceptual priority of pop-out features, 20 adults participated in a visual search and memory task using a staircase procedure. In each trial, participants viewed 9 objects, one of which was unique (a singleton, differing in color, orientation, or size) for 500 ms. Participants were asked to detect the unique object and remember its features. Following a 500 ms mask, participants clicked on the pop-out target's location. Then a probe item appeared in the same location as the target. The probe either matched the original or had one altered feature, with the degree of difference in the altered cases adapting to the participant's performance on each feature dimension. Results revealed a significant memory advantage for pop-out features over non-pop-out features, demonstrating that pop-out features are prioritized in visual working memory, supporting their enhanced role in the visual processing and reidentification of objects.

16.306 REPEATED SEARCH CAN MAKE SEARCH SLOWER AND LESS EFFICIENT: A REPLICATION AND EXTENSION.

Donald (Alex) Varakin[1], Egypt Frye[1], Steven Pesina[1]; [1]Eastern Kentucky University

Some studies, such as those on contextual cuing, suggest that repeatedly searching through the same display leads to faster and more efficient searches. Pesina and Varakin (2020, VSS) demonstrated that in some situations, searching through the same array on consecutive trials makes searches slower and less efficient. The current experiment utilized a repeated search paradigm in which search configurations repeated for eight consecutive trials as participants searched for a rotated T among rotated L's (set sizes of 8, 12, or 16). In the target-replaces-distractor condition (an exact replication of Pesina and Varakin, 2020), the target appeared in a different location within the array on each trial in a repetition series, appearing at a random location on the first trial and a location previously occupied by a distractor on the second through eighth trials. Replicating the prior results, search efficiency decreased from about 39ms/item to 52ms/item from the first to the eighth repetition. One possible explanation for this result is that locations occupied by distractors are inhibited, and the inhibition carries over from trial to trial within a repetition series. To test this idea, we included a target-replaces-empty condition, which was identical to the target-replaces-distractor condition, except that the search target appeared in a previously empty location on each trial in a repetition series, rather than a location previously occupied by a distractor. The inhibition-carryover explanation would not predict a decrease in efficiency in the new condition, since the target is not appearing at a location that could have been inhibited from previous searches. However, search efficiency again decreased from the first to the eighth search, from about 37ms/item to 50ms/item. These results suggest that the decrease in search efficiency is not due to the target replacing a distractor, but to something else about how the displays repeated.

Visual Search: Features, objects

Friday, May 16, 3:00 – 5:00 pm, Banyan Breezeway

16.307 USING THE MATCH TOOLBOX TO EXAMINE GOOD-ENOUGH GUIDANCE FOR REAL-WORLD OBJECTS

Jessica N. Goetz[1] (jessica.goetz@ucf.edu), Mark B. Neider[1]; [1]University of Central Florida

Good-enough guidance proposes that search processes prioritize the most diagnostic feature (e.g., color is prioritized over shape) (for review, see Yu et al., 2023). Accordingly, the target template that guides attention is not a highly veridical representation of the target. For example, when searching for colored circles, target templates will be broadly tuned to a wide range of colors. We examined how good-enough guidance might be employed when searching for real-world objects compared to colored circles. Observers were asked to indicate the presence of a target object and we used our MATCH toolbox (Goetz & Neider, under review) to generate real-world object distractors (critical distractors) that were 15°, 30°, and 45° (color distance) away from the target object's predominant color. On target absent trials, a distractor that matched the target's shape, but was 90-270° away from the target object's predominant color (rotated target distractor), replaced the target. Although we found more initial saccades to the critical distractor at every color distance for colored circles compared to real-world objects (all ps<.001), there was evidence of good-enough guidance for real-world objects. For both real-world objects and colored circles, we found most initial saccades were directed to the critical distractor and the proportion of initial saccades to the critical distractor decreased as color distance increased (all ps≤.001). A significant proportion of initial saccades were also directed to the rotated target distractor for real-world objects compared to colored circles (all ps<.001). Unlike colored circles, shape information can be used to guide search for real-world objects. Critically, for real-world objects at 45° of color distance, more initial saccades were directed to the rotated target distractor compared to the critical distractor (p<.001). Our data suggest that color provides good-enough guidance for real-world objects, but shape becomes more diagnostic when color is sufficiently dissimilar.

16.308 TRIAL-BY-TRIAL FEATURE-BASED SUPPRESSION IS RETROACTIVE AND SLOW

Mark W. Becker[1], Morgan R. Dodd[1], Andrew Rodriguez[1]; [1]Michigan State University

There is controversy about whether feature-based attentional suppression can be applied proactively or only occurs retroactively. To address this debate, we had participants search for a Landolt C with a horizontal break either on the right or left among Cs with vertical breaks. Each C appeared on a colored square. Search arrays were preceded by a cue and participants were informed that the target C would never appear on a square of the cued color. There were three cuing conditions: a negative cue - ¼ of the squares in the array appeared in the cue's color and the target never appeared in these squares; a neutral cue - the cue's color did not appear in the array; and a "reduced set size" cue - the cue's color did not appear in the array but the size of the array was reduced by ¼, providing a boundary condition for complete suppression by a negative cue. In addition, for one set of participants, the colored squares appeared two seconds before the C's appeared within them, which allowed us to investigate retroactive suppression. For other subjects, the colored squares and Cs appeared simultaneously, which allowed us to investigate proactive suppression. Results showed a clear suppression benefit, but only in the retroactive timing condition. In a follow-up, we varied the delay between the onset of the color squares and the Cs across four conditions (0, .5, 1, or 2 seconds). We again showed evidence for suppression, but only in the 2 second condition. These results suggest that suppression can be used on a trial-by-trial basis, but only retroactively. In addition, the process of setting up effective retroactive suppression seems extremely slow – making it unlikely to be used to guide attention in most real-world situations.

16.309 OVERT ATTENTIONAL SUPPRESSION OF HIGHLY SALIENT ONSETS

Esha Brar[1], Han Zhang[2], John Jonides[3]; [1]University of Michigan

Abrupt onsets have been known to capture attention. Is it because abrupt onsets are highly salient? In this experiment, we objectively quantified the level of salience and then examined if more salient onsets lead to greater oculomotor capture. Using the psychophysical procedure developed by Stilwell et al. (2023), we measured the salience levels of abrupt onsets, color singletons, and color-singleton abrupt onsets. In phase one, on each trial, participants (N=54) reported whether the critical item was absent or present, with the display duration dynamically adjusted based on previous trial performance. In phase two, the same participants completed a visual search task in which instead of being the target of detection, abrupt onsets, color singletons, and color-singleton abrupt onsets served as distractors. In phase one, we found that participants needed a longer time to detect abrupt onsets compared to color singletons and color-singleton abrupt onsets, suggesting that abrupt onsets were the least salient among the three. However, only abrupt onsets captured attention in the visual search task of phase two. Abrupt onsets induced a distractor-presence cost in reaction times (RT) and captured initial eye movements. Color singleton and color-singleton abrupt onsets induced a distractor-presence benefit in RT, and initial eye movements to these distractors were suppressed. These results suggest that abrupt onsets can capture attention, but not necessarily because they are physically salient.

16.310 PARALLEL SEQUENTIAL REJECTION IN CONJUNCTION SEARCH: INSIGHTS FROM EYE-TRACKING ON DISTRACTOR PROCESSING DYNAMICS

Andrea Yaoyun Cui[1], Simona Buetti[1], Pengfei Yu[1], Alejandro Lleras[1]; [1]University of Illinois at Urbana-Champaign

This study builds upon the findings of Cui et al. (2022) and Cui et al. (in revision), which proposed that the difficulty in conjunction search arises from a distractor rejection mechanism operating in a parallel-sequential manner, whereby participants first reject distractors based on one feature dimension, then reject distractors based on the second feature dimension. Here, we employed eye-tracking to investigate how participants process and reject color-differentiated and shape-differentiated distractors in conjunction search tasks. We recorded eye movements as participants searched for a target (a red triangle) among two types of distractors: orange triangles (color-differentiated) and red circles (shape-differentiated). Stimuli were arranged in quadrants, each containing only one distractor type to maintain local feature homogeneity. We analyzed fixation counts, durations, and sequences to investigate attentional allocation to different distractor types. Our results revealed a strong bias toward fixating on shape-differentiated distractors as early as the second fixation, suggesting that color-differentiated distractors were efficiently rejected via peripheral vision. Participants made significantly more fixations and had longer total fixation durations within shape-differentiated regions, indicating higher attentional costs associated with rejecting these distractors. Temporal dynamics showed that while initial fixations were unbiased, participants quickly adjusted their strategy to focus on the more challenging shape-differentiated distractors. These findings support the parallel-sequential mechanism, demonstrating that participants employ a strategic allocation of attention during conjunction search—prioritizing the rejection of easily distinguishable distractors via peripheral vision before focusing on more difficult ones. Our study underscores the importance of eye movement analysis in understanding visual search mechanisms and suggests that attentional strategies are dynamically adjusted based on distractor characteristics. In two follow-up experiments, we varied the discriminability along each feature dimension to evaluate the extent to which the parallel sequential rejection mechanism is sensitive to the relative processing ease of one feature discrimination over the other.

16.311 THE (SOMETIMES SLOW AND SOMETIMES FAST) TIME COURSE OF INFORMATION ACQUISITION DURING VISUAL SEARCH

Patrick Wu[1], Howard Egeth, Jonathan Flombaum; [1]Johns Hopkins University

We sought to understand the difference between conjunction and feature search using a limited-exposure localization task: A search display is shown for a variable amount of time, and after it disappears, the participant positions a cursor as close as possible to where they think the target was. Distance from the true location indexes what the participant was able to learn about the location of the target, and over many trials with different exposures, knowledge as a function of time can be traced. Participants were told that the target would always be a unique item, without specific foreknowledge of identifying features. In color-shape conjunction search, two experiments suggested that participants remain maximally uncertain about the location of a target during the first 300 ms of a trial. This is consistent with search being a process of first excluding stimuli while building a template for the target and finally estimating its position. In contrast, experiments with color pop-out showed immediate and rapidly increasing knowledge about the location of a target (following just 17 ms of exposure). This is consistent with a process that only needs to estimate the location of a signal. We also conducted experiments in which participants reported what the target was, following limited exposure and localization. The results suggested that participants often know what the target is—what it looks like—before knowing anything about where it is. Additional experiments demonstrate the utility of the limited-exposure localization approach for tracking the micro-genesis of search in attention capture conditions and in conditions that produce asymmetries.

16.312 LEARNED TARGET-DISTRACTOR SIMILARITY CHANGES THE PRECISION OF THE TARGET TEMPLATE AND VISUAL SEARCH BEHAVIORS

Catherine Halpern[1] (cahalpern@ucdavis.edu), Joy Geng[1]; [1]University of California, Davis

Visual search efficiency is generally lower in contexts with high target-distractor similarity. However, expectations about distractor similarity can improve performance. The current study tested the hypothesis that high distractor similarity contexts drive observers to create more precise target templates, which emphasize the parts of the target that are most distinct from distractors. Eighty participants completed a visual search task using a novel circular mushroom space created from 3D objects rendered in Blender. Participants were eye-tracked during search for a repeating target in one of two groups: the similar-distractor group saw 81.25% similar-distractor trials and 12.5% distinct-distractor trials and the distinct-distractor group saw the reversed proportions. All participants completed 6.25% memory probe trials, on which they selected the remembered target from 75 possible shapes on a "mushroom wheel." Participants in the similar-distractor group were more accurate and faster on similar-distractor trials than those in the distinct-distractor group. They also demonstrated more precise memory representations of the target and had more constrained fixations around parts of the target shape that were maximally differentiated from distractors. However, on distinct-distractor trials they took longer to respond than those in the distinct-distractor group. This difference likely reflected underlying differences in how participants in each group approached the task. Participants in the similar-distractor group serially evaluated each object's identity, demonstrating longer dwell times on targets and distractors during distinct-distractor trials, but relatively shorter dwell times on similar-distractor trials. In contrast, those in the distinct-distractor group rapidly guided attention to the most likely target – a strategy that succeeded on distinct-distractor trials, but failed on similar-distractor trials. These results emphasize the critical role of context in shaping adaptive search strategies - the predominant distractor-context shaped the precision of the target template and search strategy, which resulted in differing speed-accuracy trade-offs in visual search performance.

UCD, JSMF

16.313 INVESTIGATING THE CONTRIBUTION OF UNPREDICTABLE TARGET FEATURES TO ATTENTIONAL GUIDANCE

Zoe (Jing) Xu[1,2] (jingxu9@illinois.edu), Jun-Ming Yu[1], Alejandro Lleras[1], Simona Buetti[1]; [1]University of Illinois, Urbana Champaign, [2]University of Washington

Visual search targets are not always predictable - sometimes observers know one feature of the target (color of a scarf) but might not be able to anticipate other features (the scarf shape). Here, participants searched for a target that differed from distractors along both color and shape, with one feature dimension being known and one varying randomly across trials. Using a model comparison approach, we quantified the contribution of each feature dimension to attentional guidance. Participants were asked to search using the known dimension. In Experiment 1 (color known), the target was red and had one of four possible shapes. Distractors were orange (color-similar) or pink (color-dissimilar) and had one of two possible shapes. In Experiment 2 (shape known), the target was a house in one of four possible colors. Distractors were triangles (shape-similar) or circles (shape-dissimilar) in one of two possible colors. In both experiments, response times for all target-distractor pairs (e.g., red house target among pink diamonds) were predicted by the weighted combination of the search slopes evaluated in simple color (red among pink) and shape searches (house among diamonds). These weights indicate the extent to which each feature dimension was prioritized in attentional guidance. In Experiment 1, the results showed that the priority given to the unpredictable shape dimension depended on the usefulness of color: the shape weight was 0.85 (color weight=1.15) in the color-dissimilar condition, while the shape weight was 1.6 (color weight=0.4) in the color-similar condition. In Experiment 2, the weight for the unpredictable color dimension was 0.47 (shape weight=1.53) in the shape-dissimilar condition, and the color weight was 1.59 (shape weight=0.41) in the shape-similar condition. These results demonstrate that when the target and distractors are too similar along the known dimension, the visual system increases its reliance on the unpredictable dimension.

16.314 IMPACT OF TASK-IRRELEVANT HETEROGENEITY ON VISUAL SEARCH EFFICIENCY AND ATTENTIONAL CAPTURE

Myo-Hee Chang[1] (rarityhappiness@gmail.com), Suk Won Han[1]; [1]Chungnam National University

In the environment, multiple stimuli compete to be represented in our resource-limited perceptual/cognitive system. Such competition intensified when a set of heterogeneous stimuli are presented than when the stimuli are homogenous. The increased competition resulting from increased stimulus heterogeneity was found to affect the efficiency of visual search. According to a seminal study (Duncan & Humphreys, 1989), searching for a specific shape suffered to a greater extent when the target was surrounded by a set of heterogeneous shapes than when it was accompanied by homogeneous stimuli. Expanding this study, we investigated whether stimulus heterogeneity in a task-irrelevant dimension also affects visual search efficiency. Specifically, in the present search task, the task was to look for a 45-degree tilted grating either in clockwise or counterclockwise direction, while non-targets were vertical or horizontal gratings. Hence, the task-relevant dimension was orientation. The search stimuli could also vary in task-irrelevant dimensions, such as spatial frequency (Exp 1) or color (Exp 2). Importantly, the search stimuli were either heterogeneous or homogeneous in terms of these task-irrelevant dimensions. The results showed that when the search stimuli were heterogenous in the task-irrelevant dimension (e.g., spatial frequency or color), search efficiency was lower than when the non-targets had homogeneous spatial frequency or color. Furthermore, for a certain feature dimension, this impaired search efficiency was accompanied by diminished attentional capture by a singleton distractor. These findings are remarkable because the spatial frequency or the color of the stimuli did not have to receive attention. Rather, they affected efficiency of the search process. Based upon these, we suggest that the irrelevant heterogeneity disrupts the formation of coherent percept of visual stimuli, increasing perceptual resource demands. Even though our attentional control could be finely tuned, perfect filtering of a certain visual feature from a multi-dimensional stimulus was found to be impossible.

16.315 EXPLORING THE IMPACT OF TARGET-DISTRACTOR FEATURAL CONTRAST ON FEATURE PRIORITIZATION IN EFFICIENT VISUAL SEARCH.

Jun-Ming Yu[1] (jy87@illinois.edu), Zoe Jing Xu[1,2], Alejandro Lleras[1], Buetti Simona[1]; [1]University of Illinois, Urbana Champaign, [2]University of Washington

The visual attention system is thought to be highly adaptive to the ecology of the environment. Here, we attempted to quantify how the visual system prioritizes various visual features as a function of target-distractor featural contrast. Specifically, we quantified the contributions of color and shape to attentional guidance in conditions where shape and color dimensions varied in terms of their relative contrast. We first measured featural contrasts, indexed by logarithmic search efficiency, between targets and a wide range of homogeneous distractors in unidimensional searches. In Experiment 1, participants searched for a target differing from distractors only in shape (10 shapes). In Experiment 2, they searched for a target differing from distractors only in color (8 colors). In Experiments 3-4, the target differed from distractors along both color and shape and we selected shape and color features so that one dimension had larger contrasts than the other. In Experiment 3 (color contrasts > shape contrasts), the color logarithmic search slopes were 60, 68, 93, whereas the shape slopes were 276, 175 and 150. In Experiment 4 (color contrasts < shape contrasts), the color search slopes were 249, 144 and 140, whereas the shape slopes were 87, 75, 52. Contrary to our predictions, the results of Experiment 3 showed that participants relied more on the shape dimension, even though relying more on color dimension would have led to better performance. In Experiment 4, participants relied to equal extents on both dimensions, even though relying more on shape would have led to better performance. Overall, the results suggests that participants might not always rely on the most efficient feature to search. A follow-up study using more efficient search conditions for the larger contrast dimension was run to investigate at which point participants choose to preferentially attend to the most useful feature.

16.316 IMPACT OF GLOBAL AND LOCAL CLUTTER ON VISUAL SEARCH EFFICIENCY AND ATTENTIONAL GUIDANCE

Enilda Velazquez[1], Nelson Roque[2]; [1]University of Central Florida, [2]The Pennsylvania State University

Visual clutter impairs search performance, interfering with top-down and bottom-up attentional guidance. Previous research suggests localized clutter impairs top-down object segmentation, leading to non-objects being included in the search set, resulting in search errors (Neider and Zelinsky, 2006). Further, global clutter disrupts bottom-up information accumulation across a scene, leading to reduced search efficiency (Wolfe et al., 2002). Quantifying clutter through perceptual grouping (Subband Entropy - SE) and feature covariance (Feature Congestion - FC), allows us to more comprehensively understand the impact of top-down vs. bottom-up guidance mechanisms (Rosenholtz et al., 2007). The present study aimed to disambiguate the effects of clutter on guidance by investigating how SE and FC clutter impact search, as a function of top-down or bottom-up mechanisms, using an ecologically valid multi-target search task. Participants searched for boat targets in marina scenes taken from a satellite image set (DOTA, Xia et al., 2018). Targets were placed according to condition: (1) in marina (in clutter), (2) in water (out of clutter). If search was guided bottom-up by clutter, we predicted that global clutter would decrease search efficiency, while local clutter would not. Whereas, if search was guided top-down by clutter, we predicted that local clutter would decrease search efficiency, while global clutter would not. Linear multilevel models suggest that global SE clutter decreases search efficiency ($b = 1976.3$ ms, $p < .001$), but global FC clutter does not. This effect reverses for local clutter, where local FC clutter decreases search efficiency ($b = 477.6$ ms, $p < .001$), but local SE clutter does not. Findings suggest that clutter may impact guidance at the level of priority map formation (Wolfe, 2021). Specifically, that SE clutter may serve to quickly accumulate target location probabilities globally, then guidance uses local FC around probable locations to select targets for further processing.

16.317 EXPLORING IMPLICIT CATEGORY REPRESENTATIONS USING TARGET-RELATED CLUTTER

Elizabeth Y. Zhou[1], Yelda Semizer[2], Melchi M. Michel[1]; [1]Rutgers University, [2]New Jersey Institute of Technology

Vision researchers have long studied the effect of clutter on visual search performance. We are curious about the influence of the observer's task on image clutter. In Semizer and Michel (2022), participants searched for objects in natural images after they were shown the category name of the objects. In this case, the "category-level" metric for target-related clutter was computed based on the discrepancy between the category representation and the image representation (i.e., the distance between the feature distribution across all exemplars of the target category within the image set and the feature distribution of the image being searched). Zhou et al. (2024) showed that the category-level metric was able to predict search time either when a target is present or absent in the image. Here, we explored whether the category representation is invariant across orientations. When building the feature distribution for the category representation, we rotated each exemplar to eight random orientations within some selected ranges. We observed that as the possible range of rotation increases (30, 60, 90, 180 and 360 degrees), the performance of our category-level metric decreases in predicting the search time. Also, we investigated the effects of richness in the category representation. As the dimensionality of feature increases when building the category representation (12, 50, 100 and 200 components from principal component analysis), the performance of our category-level metric decreases in predicting the search time when the search target is absent, while the trend is mixed when the search target is present. These results potentially reflect limited orientation invariance and limited feature resolution in human category representations that guide visual search.

16.318 BENDING THE TEMPLATE: EVIDENCE FOR RELATIONAL GUIDANCE IN VISUAL FORAGING

Jan Tünnermann[1], Joy J. Geng[2], Anna Schubö[1]; [1]Philipps University Marburg, [2]University of California Davis

Visual foraging tasks offer a dynamic, naturalistic approach to studying selective attention and search strategies. In this study, we investigated how relational guidance—attention driven by the relationship between target and distractor features rather than their exact values—shapes target templates during foraging. Participants foraged for shape-defined targets (octagons) among distractors (hexagons) using stylus clicks on a tablet-PC. Critically, target and distractor colors were drawn from continuous gradients. In two experimental conditions, the distractor gradient connected either to the pink or orange end of the target gradient, while a control condition featured colors from the opposite (bluish) side of the color wheel. Although participants could have relied solely on the constant shape features to guide their search, they incorporated variable color information into their strategy, revealing consistent relational effects: distributions of selected target colors bent systematically away from the distractor color ranges, as demonstrated through Bayesian parameter estimation. This work highlights that searchers include and tune template dimensions, even when the dimension (in this case, color) is not strictly required to find the target. That relational guidance manifests in dynamic selection choices underscores the utility of foraging paradigms in directly visualizing the impact of template adjustments on behavior.

Decision Making: Perception

Friday, May 16, 3:00 – 5:00 pm, Banyan Breezeway

16.319 IMPROVING HEALTH DATA VISUALIZATIONS: EFFECTS OF COLOR AND SUMMARY INFORMATION ON DECISION MAKING

Angelica S. Busciglio[1], Jessica N. Goetz[1], Mark B. Neider[1]; [1]University of Central Florida

The purpose of data visualization is to communicate information. A well-designed visualization allows viewers to quickly extract important patterns or trends in the data (Ware, 2020). Utilizing what is known about the basic properties of human perception, such as color and spatial perception, can enhance information display. The current study explores how the use of different visualizations and forms of risk communication beyond the current standard format type affects responses to health data. We specifically investigated whether the inclusion of overall summary statements and color increased comprehension and the likelihood to engage in health-oriented behavior when viewing hypothetical cholesterol results. Participants ($N = 138$) viewed results in two formats (table and number line), across three risk levels (low, borderline, and high), either with or without summary statements and color. In the table format, participants were more likely to engage in preventative behaviors (following up with a doctor and exercising) at normal risk levels when summary statements were absent, suggesting misinterpretation without additional context (all $ps<.05$). In the table format, color improved the ability to correctly identify values as acceptable at normal risk levels, reducing misinterpretation ($p<.05$). In the number line format, at borderline risk levels, the absence of color led participants to judge values as acceptable. In contrast, the presence of color increased the likelihood of perceiving values as unacceptable, indicating that color may heighten perceived urgency ($p< .01$). Furthermore, color helped participants identify table values as acceptable without summary statements ($p<.05$). Our results demonstrate the importance of effective data visualizations and how summaries and color shape data interpretation.

16.320 MULTIPLE EXCITATORY CELL TYPES IN PRIMARY AND SECONDARY VISUAL CORTEX ARE RECRUITED DURING ACCUMULATION-OF-EVIDENCE DECISIONS

Lillian Wilkins[1], Joao Couto[1], Anup Khanal[1], Anne K Churchland[1]; [1]Department of Neurobiology, University of California, Los Angeles

The mammalian cerebral cortex plays a central role in decision-making and different visual cortical regions have unique functional roles in processing visual information. Further, cortical excitatory pyramidal neurons (PyNs) have rich diversity, but the unique roles of different PyN types across primary and secondary visual areas in decision-making remains unclear. Here, we address this gap by characterizing the activity of two classes of excitatory neurons in mouse anteromedial (AM) and primary (V1) visual cortex. We retinotopically mapped the visual cortex and used 2-photon calcium imaging to record neural activity while mice passively view stimuli or make perceptual decisions in a novel accumulation of evidence task. In the task, visual patches appeared on the left and/or right sides of the screen. Head-fixed mice judged the side with higher rate, held their choice for a brief delay period, and reported by licking one of two spouts. We recorded from ~6,000 cells in 7 mice (~3000 in AM and V1), out of which approximately 60% were responsive during the task. We measured responses in layer 2/3 (L2/3) and layer 5b (L5b) of AM and V1 and, in each population, we found both sharp orientation tuning in the passive condition and stimulus selectivity during decision-making. Interestingly, L5b neurons (that mostly target structures outside of cortex) are highly stimulus selective: individual neurons respond to stimuli corresponding to one choice option but not the other. Overall, we conclude that L2/3 and L5b neurons in both AM and V1 are highly selective to stimulus features during task performance, which invites questions about the roles of these responses in accumulation of evidence decision-making.

This work was supported by the National Institute of Health, award R01EY022979.

16.321 SENSORY INTEGRATION STRATEGIES EXHIBIT DISTINCT STATE-DEPENDENT DYNAMICS

Letizia Ye[1], Hanna Masri[2], Anne Churchland[1]; [1]Department of Neurobiology, David Geffen School of Medicine, University of California Los Angeles

Decisions about visual stimuli are influenced by internal states. Much like how a student's attention can fluctuate throughout the course of a class, internal states can fluctuate within a single experimental session. The inability to sustain engagement is implicated in Attention Deficit Hyperactivity Disorder (ADHD), leading to difficulties in learning, working memory, and sensory processing. Task-engagement directs attentional resources to relevant stimuli, enhancing visual encoding. However, little is known about how engagement can impact the downstream computations necessary to integrate accumulated sensory evidence. We trained wild-type (WT) mice (n=10) and dopamine transporter heterozygous knockout (DAT+/-) mice (n=7), a model of ADHD, on a freely-moving evidence accumulation task. Mice were presented with a sequence of Poisson-distributed stimuli and judged whether the rate was above or below a category boundary of 12Hz. We used maximum a posteriori (MAP) to fit a combined hidden Markov Model and generalized linear model with Bernoulli emissions (GLM-HMM) and characterize internal states. DAT +/- showed lower overall task-performance compared to WT mice. Interestingly, DAT +/- mice in the engaged state outperformed WT mice in the engaged state. DAT +/- mice spent more time in disengaged states compared to WT mice, and had a lower probability of remaining in the engaged state. We then computed a psychophysical kernel that captures the weight of sensory information on the animal's upcoming decision. WT mice exhibit distinct and consistent state-dependent differences in the strategy of evidence accumulation. Engaged decision-makers strongly prioritized early sensory evidence, while sensory weight was reduced across all time bins during disengagement. By contrast, DAT +/- mice weighed sensory information more uniformly over time, and displayed greater heterogeneity across states. These results demonstrate that sensory integration strategies differ with engagement, and highlight how increased fluctuations in engagement may underly behavioral and cognitive differences in ADHD.

16.322 PRE-STIMULUS SHAPE PREDICTIONS FLUCTUATE AT ALPHA RHYTHMS AND BIAS SUBSEQUENT PERCEPTION

Dorottya Hetenyi[1] (dorottya.hetenyi.21@ucl.ac.uk), Peter Kok[1]; [1]Department of Imaging Neuroscience, UCL Queen Square Institute of Neurology, University College London, London, UK

Predictions about future events significantly influence how we process sensory information. However, the neural mechanisms that maintain these predictions in anticipation of incoming stimuli remain unclear. In previous work, we demonstrated that predicted shape representations exhibit oscillatory activity in the alpha band (10 – 11 Hz) throughout pre-stimulus windows. In that work, participants' task was orthogonal to the shape predictions. Here, we extended these findings by having participants perform a shape identification task that was directly related to the shape predictions, to allow us to link the neural correlates of prediction to behaviour. We used magnetoencephalography (MEG) combined with multivariate decoding techniques to examine the content and frequency characteristics of perceptual predictions. Participants performed a shape identification task in which auditory cues predicted which shape would likely appear (75% valid). To make the identification of the shapes challenging, they were embedded in white noise. First, we found that valid prediction cues improved both identification accuracy and reaction times. Signal detection theory analyses revealed that participants were significantly biased towards reported the predicted shape (i.e., reduced criterion) without affecting sensitivity (i.e., similar d-prime). Additionally, frequency analysis of MEG shape decoding signals revealed the oscillatory nature of predicted shape representations, specifically in the alpha band (10–11 Hz) during the pre-stimulus period. Together, these findings suggest that sensory predictions are represented in pre-stimulus alpha oscillations, and the brain may leverage alpha oscillations as a neural mechanism to balance prediction and perception.

16.323 REDUCING THE SAMPLING DEPENDENCY BETWEEN VISUAL NUMEROSITY ESTIMATES IMPROVES AGGREGATED ESTIMATION ACCURACY

Daniil Azarov[1], Adam N. Sanborn[2], Nick Chater[2], Robert L. Goldstone[1]; [1]Indiana University, [2]University of Warwick

According to the "wisdom of the crowd" effect, aggregating multiple estimates leads to a more accurate estimate than most of the individual estimates (Surowiecki, 2004). This effect is observed even when the same person provides multiple estimates (Vul & Pashler, 2008), though it is generally more pronounced when estimates come from different individuals (Dolder & Assem, 2017). Previous research indicates that encouraging individuals to make two highly divergent estimates can lead to greater accuracy when the estimates are averaged compared to providing two similar estimates (Herzog &

Hertwig, 2009). Building on this, we hypothesized that the accuracy of averaged estimates would increase as the dependency between sampled estimates decreases, whether due to temporal separation, involvement of different individuals, or estimation of distinct stimuli. In the present study, 56 participants were tasked with estimating the number of objects displayed on a screen (e.g., 90 crackers). For each stimulus set, participants made two estimates either consecutively (e.g., Trials 1 and 2) or with a larger time gap (e.g., Trials 3 and 90). We then aggregated the estimates for each pair of stimuli. Our results show that averaging two estimates from the same individual was beneficial (compared to individual estimates) only when the estimates were separated in time. Additionally, averaging estimates from different individuals produced more accurate results than from the same subject. Finally, accuracy was maximized when the stimuli for which the estimates were made differed (e.g., estimating 90 crackers and 90 butterflies), rather than when the same stimulus was estimated twice, regardless of whether the estimates came from the same or different individuals. These findings underscore the importance of minimizing the sampling dependence between estimates — whether by separating estimates in time, using multiple individuals, or estimating different stimuli — in order to achieve the most accurate estimates.

16.324 TRACKING THE TIME COURSE OF GLOBAL AND LOCAL PROCESSING

Steven Shofner[1], Paul Dassonville[1]; [1]University of Oregon

Images consist of visual information at multiple scales, from fine details to gestalt. Reaction time (RT) studies of the perceptual mechanisms associated with these different levels typically find that information at the global scale is processed more quickly than information at the local scale (the global precedence effect), and that the impact of irrelevant global information on a local judgment is greater than the opposite (global interference effect). However, reaction times are a blunt measure of these effects across all stages of processing, from early sensory analysis to eventual motor response. The present study is an attempt to more precisely measure the time course of local and global processing, by employing a finger tracking task. Participants began each trial by initiating a tracing movement upward on a touch screen. Shortly after movement onset, a hierarchical C (a large C made of small c's, both of which could be open to the right or left in a mirror-reversed fashion) was presented, with the participant required to quickly and accurately alter the trajectory of the ongoing movement toward one of two target locations on the left or right, as indicated by the orientation of the c's at the attended level (local or global, in separate blocks). Trajectory analysis indicated that global-level orientation information began to be incorporated into the response ~300ms after stimulus onset, and local-level information in ~335ms, yielding a global precedence effect of ~35ms (smaller than the ~60ms global precedence measured in an RT task using the same stimuli). In contrast, the tracking task exhibited a larger global interference effect (~30ms) than in the RT task (~15ms). These findings provide a clear demonstration that a finger tracking task can be used to more precisely determine the time course at which global and local visual information is available to guide movements.

16.325 VISUAL CUES OF INFECTION RISK INFLUENCE IMPULSIVITY AND RISK PREFERENCE

Guangsheng Liang[1], William G. Hayward[1]; [1]Lingnan University, Hong Kong SAR

Recent research has identified a proactive behavioral mechanism that seeks to reduce potential exposure to pathogens; this mechanism has been called the behavioral immune system (BIS). Individuals can accurately detect potential exposure to pathogens through visual perception, and it has been demonstrated that a more aggressive immune response could also be elicited by visually presented infectious disease symptoms. In this study, we further explore the influence of this visually-elicited BIS through multiple cognitive tasks. We elicited the BIS via implementing a learning session with disease symptom pictures (infectious vs. non-infectious disease). Participants were tasked to perform a classic Go/no-go (GNG) task after the symptom presentation to evaluate the immediate influence of visual BIS priming on impulsivity. In addition, before and after the presentation of symptom pictures, participants were instructed to report their current emotional state and complete an Iowa Gambling Task (IGT) to assess their changes in emotion and risk preference in response to the BIS activation. A classic GNG task result was observed among participants viewing non-infectious disease symptoms: accuracy was lower in the response-inhibit condition. Importantly, this pattern disappeared among participants viewing infectious disease symptoms. Moreover, accuracy in both GNG task response conditions after viewing infectious disease symptoms was equivalent to GNG performance in the response-inhibit condition after viewing non-infectious disease symptoms, indicating an impairment of instrumental behavior performance. Although the IGT performance was equivalent between manipulated conditions, participants' risk preferences varied before and after learning. Specifically, participants in both groups improved their decision quality, but only those who viewed infectious disease symptoms showed a trend from a preference for a low-frequency high-penalty to a high-frequency low-penalty. These findings provide extended evidence of the relationship between the visual system and physiological responses to potential vectors of disease.

This work was supported by a Special Topics Grant from the Hong Kong Research Grants Council (STR4/M-701/23-N) to William G. Hayward.

16.326 WHITE MATTER STIMULATION WITH SINGLE ELECTRICAL PULSES MODULATES NEURONAL ACTIVITY IN VISUAL CORTEX

Harvey Huang[1] (huang.harvey@mayo.edu), Kendrick Kay[2], Nicholas Gregg[1], Gabriela Ojeda Valencia[1], Brian Lundstrom[1], Myung-Ho In[1], John Huston[1], Christoph Kapeller[3], Yunhong Shu[1], Gregory Worrell[1], Kai Miller[1], Dora Hermes[1]; [1]Mayo Clinic, [2]University of Minnesota, [3]g.tec medical engineering

Intracranial electrical stimulation can induce phosphenes and distort visual percepts. Relating these perceptual phenomena to mechanistic underpinnings requires understanding how stimulation affects neuronal populations involved in visual processing. The basic unit of stimulation, a single electrical pulse, can travel through white matter to influence connected neuronal populations. Its effects can be directly observed in signals measured by intracranial EEG. Studies have characterized how single pulses affect neuronal activity at rest, but not how they modulate active visual processing. We investigated whether single pulses modulate neuronal responses to images. In two human subjects, we measured intracranial EEG broadband responses to quantify local firing rates and evoked potentials to quantify synchronous inputs. Single pulses were delivered to electrodes in white matter tracts connected to measurement electrodes in visual cortex. Images appeared on-screen at 0, 100, or 200 ms after each pulse. Using finite impulse response modeling, we decomposed each response into components attributed to electrical stimulation and visual processing, and evaluated whether the visual component varied by stimulation interval as evidence for modulation. Single pulses induced transient broadband increase followed by suppression, atop visual broadband responses that did not vary by stimulation interval. In contrast, single pulses elicited prominent stimulation-evoked potentials but also modulated the visual-evoked potentials in an interval-dependent manner: visual-evoked potentials were larger when stimulation occurred closer to visual onset. Control stimulation sites outside connected white matter tracts did not produce such modulation. The stimulation-induced broadband changes resembled single pulse responses previously observed in subjects at rest. The evoked potential changes indicate that even the basic single pulse unit increases the strength or synchrony of visual inputs. These findings suggest that the neurophysiological effects of electrical stimulation in the visual system are two-fold: it adds artificial inputs which excite and inhibit; and it interacts with the processing of natural inputs.

This work was supported by the NIMH (R01MH122258), the NEI (R01EY035533), the NIGMS (T32GM145408), and the American Epilepsy Society (937450).

Action: Navigation and locomotion

Friday, May 16, 3:00 – 5:00 pm, Banyan Breezeway

16.327 VISUAL MODEL OF LOCOMOTION REPRODUCES LANES AND STRIPES IN CROSSING HUMAN CROWDS

Kyra Veprek[1], Sina Feldmann[1], William Warren[1]; [1]Brown University

The visual control of locomotion has been modeled for individual pedestrian behavior; however, this approach has not been applied to collective human behavior, where spontaneous pattern formation is often observed. We hypothesize that an empirical visual model of human locomotion will reproduce the emergent pattern of lanes and stripes observed in crossing flows, when two groups of pedestrians walk through each other at crosswalks or intersections. Mullick, et al. (2022) manipulated the crossing angle between two groups and found an invariant property: stripe orientation is perpendicular to the mean walking direction (i.e. 90˚ to the bisectrix of the crossing angle). Here we determine the combination of model components required to simulate human-like stripes: (i) steering to a goal (Fajen & Warren, 2003), (ii) collision avoidance with opponents (Bai, 2022; Veprek & Warren, VSS 2023), and (iii) alignment with neighbors (Dachner, et al., 2022), together called the SCruM (Self-organized Collective Motion) model. We performed multi-agent simulations of the data from Mullick et al. (2022), using fixed parameters and initial conditions from the dataset. There were two sets of participants (N=36, 38) with 18 or 19 per group. Crossing angle varied from 60˚ to 180˚ (30˚ intervals), with ~17 trials per condition. The minimal model necessary to reproduce stripe formation consists of the goal and collision avoidance components. Mean stripe orientation did not differ from 90˚ to the bisectrix (BF10 < 0.01, decisive). However, the SD of heading during crossing was significantly larger than the human data ($p<0.001$), whereas the SD of speed was significantly smaller ($p<0.001$). Thus, the ratio of heading/speed adjustments is lower than previously found, implying the need to reparameterize model components for walking in groups. In sum, steering to a goal and collision avoidance are sufficient to explain stripe formation in crossing flows, while alignment is unnecessary.

NIH R01EY029745

16.328 GAZE-CONTINGENT VISUAL OCCLUSIONS BIAS STEERING BEHAVIOR: IMPLICATIONS FOR VISION LOSS

Arianna P. Giguere[1] (apg7742@rit.edu), Matthew R. Cavanaugh[2,3], Krystel R. Huxlin[2,3], Duje Tadin[2], Brett R. Fajen[4], 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

Why do some cortically blind (CB) drivers who are missing a quadrant or more of their visual field have trouble maintaining a central lane position (Bowers et al. IOVS 2014), while others do not? A recent driving study in virtual reality showed that most right-sided CBs perform similarly to controls, while most left-sided CBs steer differently (Giguere et al. JoV 2024). Here, we asked whether these steering differences arise from the occlusion of visual information or from noise and/or biases in visual motion processing introduced by CB (Cavanaugh et al. JoV 2015). The steering and gaze behavior of 30 subjects with normal vision (mean age: 19.9 years, SD: 1.35) was recorded in a virtual steering task where artificial occluding “scotomas” were imposed on a quadrant of their visual field. The central five degrees of vision were spared to mimic the sparing present in most CB patients. Turn direction (left/right), turn radius (two non-constant radii), and occlusion quadrant (one of four quadrants or no occlusion) were randomized between trials. Participants were biased in lane position away from the side of the simulated blind field with one exception: lane position was unchanged when turning away from occlusions in the upper quadrants. With respect to eye movements, the occlusion quadrant did not affect average gaze azimuth/elevation, percent fixations on road, or look-ahead distance on road. However, participants made fewer saccades when turning towards an upper quadrant occlusion. This study shows that spontaneous occlusion of visual information in the upper vs lower visual field biases steering, but left vs. right occlusions do not fully account for the differences that arise from left and right-sided CB. Future work should examine the influence of CB adaptation to their visual impairments and the effect of internal processing noise on steering.

Research to Prevent Blindness/Lions Clubs International Foundation Low Vision Research Award (LVRA)

16.329 HUNTING AT THE LIMIT: SENSORIMOTOR INTEGRATION OF VISUAL DIRECTION VARIABLES

GUIDES HEAD AND BODY MOVEMENTS DURING VIGOROUS TARGET PURSUIT IN MICE

Daniel Pollak[1] (dpollak@caltech.edu), Xinghao Li[1], Jasmine Wang[1], Markus Meister[1]; [1]California Institute of Technology

Mice coordinate head and body movements during pursuit in a sequence of dynamic orienting behaviors. These behaviors take place on the scale of tens of milliseconds and are subject to modification from unexpected changes in the environment. To elicit reliable and prolonged hunting bouts, we developed a motorized gantry to maneuver prey items through an open field via a magnet beneath the floor. Mice pursued targets under two hunting paradigms: open- and closed-loop pursuit. In open-loop, the target moves along a pre-determined trajectory. In closed-loop, the target actively avoids the mouse. Food-deprived mice learned to pursue target objects, usually a chocolate chip, within minutes of being introduced to the setup. We developed a transfer function to predict orienting motions of the head and body from directional variables related to a visual target (angular target position, angular target velocity). The predictions of this transfer function accounted for 58% of the variance in head-turning movements during pursuit and 59% of the variance for body turning movements during pursuit. However, the head and the body do not respond instantaneously to visual input. Head movements were delayed by 60 ms while body movements were delayed by approximately 90-100 ms. This phenomenological approach to visual orienting during natural behavior provides a promising avenue for understanding the neural circuits that produce it by condensing a complex behavior into a few informative variables and highlighting details about sensorimotor integration, such as visual feedback delays and coordinated head-body adjustments, that may govern vigorous pursuit behavior across species.

16.330 CEREBRAL VISUAL IMPAIRMENT IS ASSOCIATED WITH ALTERED GAZE DYNAMICS AND CONSERVATIVE WALKING STRATEGIES TO SAFELY NAVIGATE DENSE, DYNAMIC SPACES IN VIRTUAL REALITY

Jonathan K. Doyon[1,2], Madeleine Heynen[1,2], Wei Hau Lew[1,2], Claire E. Manley[1,2], Alex D. Hwang[1,2], Jae-Hyun Jung[1,2], Lotfi B. Merabet[1,2]; [1]Massachusetts Eye and Ear, [2]Harvard Medical School

Cerebral visual impairment (CVI) is a brain-based visual disorder and the leading cause of pediatric visual impairment in the developed world. Typically, CVI arises as a result of early neurological damage and maldevelopment of retrochiasmal visual processing pathways. Individuals with CVI often report an inability to tolerate dense and highly dynamic visual environments (e.g., crowded shopping malls). Consequently, safely navigating these environments becomes nontrivial, requiring attentionally demanding adaptive strategies. To investigate how individuals with CVI safely negotiate such environments, we asked subjects to complete a collision-avoidance task in an immersive virtual shopping mall displayed in a Meta Quest Pro head-mounted-display with integrated eye tracking. Two CVI (25.5 years ±0.71) and 4 healthy control subjects (27.25 years ±3.40) walked unrestricted through crowds of virtual pedestrians (densities of 2, 10, or 20 people) while detecting (via button press) and avoiding potential collisions (e.g., by altering walking speed or path) from target pedestrians approaching from various bearing angles (±20°, ±40°, or ±60°). CVI subjects performed comparably to controls in terms of collision detections, response times, and successful avoidances. However, CVI subjects showed altered gaze profiles with greater time-to-first-fixation of the target (+440ms, $\beta=0.76$, $p=0.02$), greater fixation durations (+588ms, $\beta=0.75$, $p<0.001$), and fewer fixations (-0.67 fixations, $\beta=14.91$, $p=0.003$). CVI subjects also walked slower than controls (-0.05m/s, $\beta=-0.07$, $p=0.049$). For both groups, crowd density slowed response times (+561ms, $\beta=0.29$, $p<0.001$) and resulted in greater time-to-first-fixation (+438ms, $\beta=0.27$, $p=0.003$). These results suggest that the CVI subjects have developed strategies to safely navigate crowded spaces, such as more conservative walking and different gaze behaviors. We suspect that these alterations result from an impaired ability to balance the stimulus-driven and goal-directed aspects of the task, evidenced by the CVI subjects being slower to fixate the target, and maintaining longer fixations on the target to ensure safe passage.

16.331 EXPLORING HUMAN NAVIGATIONAL STRATEGIES IN A DYNAMIC VR SOCIAL WAYFINDING TASK

Jakub Suchojad[1], Samuel S. Sohn[1], Michelle Shlivko[1], Karin Stromswold[1], Jacob Feldman[1]; [1]Rutgers University

How do people navigate through crowded, dynamic environments? We investigated social wayfinding (navigation influenced by other people) in virtual reality (VR). Using a VR headset, we placed our participants in a virtual train station and asked them to physically navigate to one of the two exit gates while under time pressure. The station included several obstacles, and a number of virtual humans ("agents"), some static and others walking. The walking agents passed single file in two rows approximately perpendicular to the participant's path, requiring the participants to navigate through gaps between the agents. In a series of experiments, we manipulated the sitting agents' presence, the walking agents' direction, and the time allotted for subjects to reach the gate. Our analyses focused on the balance between global planning, in which the wayfinder plots the entire path from beginning to end, and local planning, in which the wayfinder continually modifies their path in light of newly encountered obstacles. We found evidence for local wayfinding from several sources. First, though most subjects went directly towards the target gate, following a global plan, some subjects switched midway from the foil to target gate, suggesting a more local strategy. Second, even subjects who consistently headed to the target gate decelerated as they approached the room's midpoint, suggesting local path modifications to pass through a "gap" between agents. In Exp. 3, we collected participants' eye gaze as they traversed the room. Participants fixated almost exclusively on moving agents and other informative elements of the environment (e.g., the sign that identified the target gate or the countdown timer). In sum, participants' movements and eyegaze provide evidence for both global and local decision-making in a social wayfinding task. Our study demonstrates an interplay of both strategies, with the balance depending on the progression of available information.

NRT-FWHTF: Socially Cognizant Robotics for a Technology Enhanced Society (SOCRATES) Grant No. 2021628

16.332 VISUAL CONTROL OF DECELERATION: ATHLETE ADAPTATION IN HIGH-DEMAND TARGET INTERCEPTION TASKS

Dominic Willoughby[1,2] (domwill@unc.edu), Ryan P. MacPherson[2,3], Paula L. Silva[4], Adam W. Kiefer[1,2,3]; [1]Human Movement Science Curriculum, University of North Carolina at Chapel Hill, [2]Matthew Gfeller Center, University of North Carolina at Chapel Hill, [3]Department of Exercise and Sport Science, University of North Carolina at Chapel Hill, [4]Department of Psychology, University of Cincinnati

Efficient pursuit and interception of opponents is fundamental for successful contact-sport performance. However, the underlying perceptual-motor strategies athletes use to adapt deceleration patterns to changing demands remain unclear. Drawing from Lee (1976), Fajen and Devaney (2006) described a visual cue (based on optic flow rate and tau) that specifies "ideal" deceleration (i.e., the deceleration rate needed to stop just short of a target without requiring adjustments). In this study, we investigated whether and how athletes deviate from ideal deceleration—and thus from an adjustment minimization strategy—as target speeds approach maximal effort levels. Using a novel pursuit task in wireless virtual reality, 21 collegiate contact-sport athletes (9 females) ran in a 14×15m space to intercept virtual humanoid targets moving at 50%, 66.6%, 83.3%, or 100% of their maximum sprint speed. Positional data (60Hz) allowed calculation of deceleration onset distance, acceleration profiles, and root-mean-square error (RMSE) between ideal and observed deceleration patterns, normalized per participant. On all trials, athletes successfully intercepted the target before it reached the 15 m boundary edge. Linear mixed-effects models treated participants as random effects to evaluate condition-specific differences. Normalized RMSE values increased with target speed, increasing from 0.84 (50%) to 0.95 (83.3%) and 0.93 (100%), both $p < 0.01$, indicating greater deviation from ideal deceleration. Initial deceleration distances also shifted from 5.53m at 50% to 6.01m and 6.83m at 83.3% and 100%, respectively ($p < 0.001$). This suggests that athletes used a more aggressive deceleration strategy, characterized by a delayed initiation of deceleration and harder braking, as target speed increased. The similarity between 50% and 66.6% conditions demarcates a possible performance threshold where increasing task demands may prompt a shift in control strategies. Overall, these findings elucidate how athletes balance efficient motor execution with success in high-stakes tasks, offering insights for eventual training enhancement and rehabilitation interventions.

The authors would also like to thank Ainsley Mesnard and Max Zawel for their contributions to data collection and participant recruitment

16.333 INFORMATION-BASED VERSUS MODEL-BASED ACCOUNTS OF GAZE BEHAVIOR DURING STEERING

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

The ability to guide high-speed steering along a winding road depends on the skillful deployment of gaze strategies. Recent research on steering behavior concludes that gaze strategies are driven by model-based predictive mechanisms (Lappi and Mole, 2018). By this account, internal models are necessary to identify locations (waypoints) along the desired future path through which the driver intends to travel. The strongest evidence for this hypothesis comes from studies in which subjects drove through a series of waypoints positioned along an S-shaped path in a virtual environment (Tuhkanen et al., 2019, 2021). Occasionally, one of the waypoints was not rendered. Nevertheless, subjects made fixations to the location where the missing waypoint would have been. The authors concluded that gaze is driven by internal models that predict locations of upcoming steering targets based on prior knowledge of the path. First, we explore alternative model-free accounts to capturing the observed gaze behavior based on simple heuristics that predict the positions of future waypoints in the context of the highly repetitive Tuhkanen steering task. Second, the previous studies aimed to demonstrate an influence of prior knowledge under conditions in which the upcoming waypoint was missing (i.e., information was absent). This leaves open the question of whether such knowledge also plays a role when waypoint information is available. To address this question, we replicated the Tuhkanen study with one major difference. Rather than introducing occasional gaps between waypoints, we occasionally shifted waypoints off the path by a small amount. Analyses focused on the accuracy and timing of fixations to incongruent versus congruent waypoints, which would be expected to differ if gaze is driven by model-based predictions but not if they are based on currently available information. We discuss the implications of these findings for theoretical accounts of gaze behavior during steering.

NSF 2218220

16.334 RECALIBRATING TO CHANGES IN ACTION CAPABILITIES WHILE STEERING THROUGH MULTIPLE WAYPOINTS

AJ Jansen[1], Brett Fajen; [1]Rensselaer Polytechnic Institute

Humans perform many locomotor tasks that require swiftly navigating through cluttered environments. To move smoothly through multiple waypoints, actors must maintain the possibility of task success, which is constrained by their action capabilities. In a previous study, we found that when subjects steer through multiple waypoints, they adapt their approach trajectory to the most immediate waypoint (WP_N) in a way that depends on the position of WP_N+1. Specifically, they initially veer away from WP_N in a direction opposite WP_N+1 before turning back, allowing for a smoother trajectory through the series of waypoints. Importantly, such anticipatory steering behavior was more exaggerated when the subject's vehicle was sluggish compared to when it was agile. The present study explored the ability of subjects to recalibrate to changes in action capabilities. Subjects used a game controller to steer a drone through a series of three gates in a virtual environment viewed on a monitor. Two gates were at a fixed distance on the centerline, and the third at a variable distance and angle. The experiment comprised 11 blocks of 18 trials and drone dynamics were manipulated across blocks. The drone was sluggish for Blocks 1-5, gradually transitioned to agile in Block 6, and remained agile for Blocks 7-11. Although subjects seemed unaware of the change in drone dynamics, they adapted their steering behavior following the change, gradually generating more direct trajectories to Gate 2 over blocks. The findings suggest that humans are sensitive to changes in their action capabilities and adapt their behavior to maintain the possibility of task success in accordance with affordance-based control. The study provides a clearer understanding of how and under what conditions humans adapt their movements in anticipation of future goals and informs the development of control strategies for steering through multiple waypoints.

NSF 2218220

16.335 THE ROLE OF SPATIAL KNOWLEDGE IN GUIDING HIGH-SPEED STEERING

Grace Roessling[1], Brett Fajen[1]; [1]Rensselaer Polytechnic Institute

It is well established that online visual information plays a critical role in the guidance of high-speed steering during automobile driving. Visual information can inform the driver of their current speed, heading direction, and lane position, as well as the upcoming road curvature. However, relatively little is known about the role of spatial knowledge of road geometry in the control of high-speed steering. The present study was designed to investigate if spatial knowledge contributes to the control of steering along a winding road by examining how steering behavior differs when drivers follow familiar versus unfamiliar roads. Two groups of subjects (N = 20) used a steering wheel and foot pedal system to guide a car along a set of ten winding roads (~3 km long) in a simulated environment. Each road comprised eight distinct curved segments (90° turns, U-turns, S-turns) interleaved with straight segments. The geometry of the road was identical across trials in the Constant Track condition. In the Variable Track condition, road geometry varied across trials but was the same as that in the Constant Track condition on Trial 10, allowing us to compare behavior across groups on the same road. Patches of fog were also added to some segments to determine whether the role of spatial knowledge depends on visibility. Results showed an effect of visibility on speed, speed variance, and steering variance but no significant effects or interactions involving track constancy. These findings suggest that either the experimental conditions (e.g. the number of trials) were not sufficient for the Constant Track group to learn the road geometry or that spatial knowledge plays a negligible role in the control of steering during everyday driving. Discussion focuses on the implications of these findings for models of control strategies used by drivers when following a winding road.

16.336 VISUOMOTOR COORDINATION DURING WALKING IN A COMPLEX NATURAL ENVIRONMENT

Stephanie M Shields[1] (smshieldsphd@gmail.com), Kathryn Bonnen[1]; [1]Indiana University

During visually-guided walking, a person's view of the environment changes with their body, head, and eye movements. What they see then influences their subsequent movements, forming a sensorimotor control loop. A crucial piece of that loop is the precise coordination between the body, head, and eyes in controlling gaze location. Here, we aimed to study such coordination in unconstrained locomotion, with an initial focus on coordination between the eyes and head. We used previously published data from a study that simultaneously recorded participants' eye and body movements while they walked along a nature trail with varying terrain complexity (Bonnen et al., 2021). Analyzing changes in eye and head direction (azimuth and elevation), we found that the eyes moved more frequently and in a wider variety of directions than the head. The distribution of eye movement directions was fairly even (with a slight cardinal bias) while changes in head direction were more vertical than horizontal. As a result, horizontal gaze shifts were most common. Analyzing the timing of these changes suggests, in line with existing literature, that the temporal relationship between eye and head movements is flexible, with gaze shifts most often preceding or co-occurring with head movements. Additionally, we compared results across terrain complexity and found variation in the direction and timing of changes. Differences in the distributions of change directions are consistent with previous findings suggesting that terrain complexity impacts gaze behavior, seemingly by increasing task difficulty and forcing participants to pay more attention to the ground ahead of them. Future work will aim to test that connection by analyzing eye-head coordination conditioned on the current eye and head directions relative to the walking path. These results will help advance understanding of how eye and movements synergistically interact during natural behavior.

16.337 THE CORTICAL NAVIGATION NETWORK IS ORGANIZED INTO DISTRIBUTED FUNCTIONAL GRADIENTS

Tianjiao Zhang[1], Jack Gallant[1]; [1]UC Berkeley

Actively navigating through the natural world requires close coordination of perception, planning, and motor actions. Multiple regions in the human cerebral cortex have been implicated in representing navigation-related information. Rodent studies suggest that navigation-related information is represented in multiple overlapping spatial gradients that extend across functional regions (Minderer et al., 2019, Tseng et al., 2022). Is this organizational principle shared by the navigation network in the human cerebral cortex? How is the navigation network related to other networks, such as those that process visual inputs and produce motor outputs? To address these questions, we used fMRI to record brain activity from six subjects performing a taxi-driver task in a naturalistic virtual world (110-180 minutes of data per subject). We extracted 28,161 features across 38 visual-, motor-, and navigation-related feature spaces from the experiment recordings, and used banded ridge regression to fit encoding models for all feature spaces simultaneously. To identify the network of cortical regions that represent navigation-related information, model connectivity (MC) was used to hierarchically cluster the voxel weight vectors. MC identified a network of 11 regions in the visual, parietal, and prefrontal cortices that each represent a distinct combination of navigation-related features. To determine its organizational principles, UMAP was used to recover a functional space for the navigation network. In this space, the navigation regions are organized into a continuous functional distribution, and this distribution maps to continuous spatial gradients on the cortical surface. Finally, to relate the navigation network to other networks, UMAP was used to recover a functional space for the whole cortex. In this space, the navigation network appears to be positioned in the middle of a broad gradient extending from visual to motor networks. These results provide a detailed characterization of the functional gradients underlying the cortical network that mediate active, naturalistic navigation.

This work is funded by the NIH, ONR, the NSF GRFP, and Ford

16.338 GAZE IN DYNAMIC NATURAL ENVIRONMENTS

Youjin Oh[1], Nathaniel Powell[1], Daniel Panfili[1], Mary Hayhoe[1]; [1]University of Texas at Austin

In complex natural environments it is unclear how attention is directed to information needed to avoid independently moving obstacles. To examine behavior in such contexts we recorded gaze, head, and foot movements and 3D scene data while subjects walked through crowded sidewalks. Gaze was recorded using a Pupil Labs Core mobile eye tracker which was integrated with a head-mounted Intel RealSense 435i stereo camera used to extract depth relative to gaze location. Objects within the video feed of the RealSense depth camera were classified using a YoloV3 object detection algorithm. The results provide a description of the gaze behavior people exhibit in the presence of moving objects like other pedestrians, and how it varies based on distance from the walker. Preliminary data show frequent gaze shifts to distant objects and regions of space. Pedestrians were fixated when they were around 5 to 8 meters away from the walker, depending on pedestrian density. This distance corresponds to roughly 10 footsteps away, suggesting that walkers adopt a proactive policy for detecting potential hazards. In the context of this strategy, the total time looking at pedestrians was quite short, even with high pedestrian density. About one-tenth of the total time was spent directly fixating other pedestrians. Thus only short time periods of direct gaze on potential obstacles are necessary, and walkers spend most time looking in the direction of the future walking path. The effectiveness of this strategy indicates that the predictability of pedestrian behavior allows efficient use of gaze in these contexts. By remaining sensitive to the statistics of the natural world, subjects can allocate attention efficiently and avoid reliance on bottom-up mechanisms (Jovancevic et al. 2009). Our results add to the understanding of how the visual system extracts the necessary information to adapt to the dynamic environment.

Action: Miscellaneous

Friday, May 16, 3:00 – 5:00 pm, Banyan Breezeway

16.339 SLOWER READING SPEED IN MIGRAINE: EFFECTS OF DISCOMFORT?

Sarah M Haigh[1], Isabella M Durda[2]; [1]University of Nevada, Reno

Uncomfortable stripes can induce illusions, headaches, and seizures, but there is debate surrounding their effects on visual processing. Reading is a visually demanding task that can be slowed by increasing the horizontal autocorrelation of the text (more stripes within words) and by flickering light that induces a phantom array. Individuals who are visually sensitive are more affected by the flicker and stripes than those who are not. The current study investigated if the decline in reading speed was evident with other uncomfortable visual environments. We measured reading speed with text having high or low horizontal autocorrelation, and the text was surrounded by a chromatic grating pattern that varied in its component colors from one trial to the next. Grating patterns that contain large color differences (e.g. red/blue) tend to be more uncomfortable than patterns with a small color difference (e.g. pink/purple). The text comprised three- or four-letter words arranged in a random order to create a 15-line block of text. This reduced the effect of comprehension skills on reading speed and increased the visual demand, even for the low autocorrelation text. We also compared the reading speed of individuals with migraine (a visually sensitive group) with headache-free individuals. Reading speed was slower in the migraine group and slower when reading the high autocorrelation text compared to the low. There were small effects of color difference on reading speed, but these did not depend consistently on color difference. We conclude that uncomfortable visual environments can negatively affect visual processing but may depend on where they are presented in the visual field: the text was presented centrally and reliably affected reading whereas the surrounding chromatic gratings did not. The finding that the migraine group were slower overall highlights the visual demands involved in reading, supporting the need for visually comfortable text.

Nevada Undergraduate Research Award to IM Durda

16.340 NOVEL MOTOR SKILLS RECRUIT VISUOMOTOR ABSTRACTIONS

Samuel McDougle[1,2], Zekun Sun[1]; [1]Department of Psychology, Yale University, [2]Wu Tsai Institute, Yale University

"Motor equivalence" refers to the fact that people can perform complex movements in multiple ways to achieve the same goal (e.g., signing their name with different effectors). This impressive ability suggests that the mind recruits abstract, high-level representations of actions that transcend lower-level muscle commands. However, evidence that these abstract representations actually exist, or contribute to motor skill, is lacking. For example, an overlearned visual template (e.g., a letter in your native alphabet) could help one perform novel visually-guided movements using online feedback control, without any need for abstraction. We reasoned that if abstractions contribute to real-world motor skills (e.g., handwriting), training the abstraction itself – but not lower-level muscle commands or visual templates – should improve skill. We compiled novel symbols from the Omniglot database, and generated videos that demonstrated to participants (n = 30) the hand movements required to write each symbol. In each task block, participants first imitated writing a target symbol a single time after viewing a demo. Next, they completed 3 training trials, where the symbol they attempted to copy was rotated by one of three extreme angles (90, 180, 270 degrees). Thus, training trials did not recruit the same muscle commands or visual outcomes as the initial learning trial. After training, participants again copied the symbol at its initial orientation. Performance was evaluated by computing both copying speed and error (the Procrustes distance, a measure of shape similarity between the symbol and participant copy). Participants' performance improved following training ($P < 0.0001$). In other words, participants' handwriting improvements appeared to be driven by refining high-level, abstract representations of the required movements. Follow-up experiments (n = 200) replicated this result, and ruled out effects of visual familiarity and general motor practice. Our work suggests that the mind represents abstract visuomotor plans to guide skilled behavior.

This work is supported by NIH grant R01 NS13292

16.341 CONTINUOUS PSYCHOPHYSICS DIFFERENTIATES VISUAL AND MOTOR PERFORMANCE ACROSS MOTOR DOMAINS

Veronica Pisu[1] (veronica.pisu@soton.ac.uk), Omer F. Yildiran[2,3], Chloe Lam[1], Saivydas Villani[1], Pascal Mamassian[3], Dominik Straub[4], Constantin Rothkopf[4], Guido Maiello[1]; [1]University of

Southampton, [2]New York University, [3]Ecole Normale Supérieure, [4]Technical University of Darmstadt

'Continuous psychophysics' yields rapid estimates of human visual sensitivity using an easy target-tracking task (Bonnen et al., 2015). The assumption is that tracking performance—e.g., evaluated from the cross-correlation between target and tracking velocities—inversely correlates with perceptual uncertainty: it should be harder to track a target when it is not clearly visible. Estimates of motor variability can also be obtained (Straub & Rothkopf, 2022). Here, across two experiments we tested whether the paradigm can tease apart visual and motor function. Experiment 1 provides a benchmark for comparing eye and mouse tracking. Participants tracked a randomly moving target on a screen using their gaze, with and without concurrent mouse-tracking. Targets were three bivariate Gaussian luminance blobs with fixed total luminance and increasing standard deviation, to manipulate visual noise. Mouse-tracking was more accurate (higher peak cross-correlation, $p < .001$) and had a longer lag than eye-tracking ($p < .001$). Critically, as visual noise increased, performance decreased similarly across tracking modalities: peak cross-correlation decreased and occurred later (peak: $r = 0.39$, $p < .05$; lag: $r = .81$, $p < .001$), suggesting that these differences stem from the motor components of the tasks. In Experiment 2, we compared tracking performance under simulated visual and motor impairment. Participants used a mouse to track a randomly moving concentric Gabor under three levels of simulated visual impairment (using Cambridge Simulation Glasses; Goodman-Deane et al., 2013), using either their dominant or non-dominant hand. Tracking performance decreased with increasing simulated visual impairment (lower peak, $p < .01$; longer lag, $p < .001$) and was consistently worse when using the non-dominant hand (lower peak, $p < .001$; longer lag, $p < .001$). Taken together, our results demonstrate that continuous psychophysics yields reliable measures of visual function, while also providing estimates of oculomotor and upper-limb motor function.

This research was funded by Fight for Sight (Grant Reference: RESSGA2302, awarded to GM).

16.342 PROBING FLEXIBLE ADAPTATION OF INTERNAL PHYSICS MODELS IN HUMANS

Giacomo Aldegheri[1] (giacomo.aldegheri@psychol.uni-giessen.de), Roland W. Fleming[1,2]; [1]Department of Experimental Psychology, Justus Liebig University Giessen, [2]Centre for Mind, Brain and Behaviour (CMBB), University of Marburg and Justus Liebig University, Giessen

Humans are able to successfully interact with objects with vastly different physical characteristics. For example, hard objects, such as boxes, and soft objects, such as cloth, behave in entirely different ways when manipulated. Humans might achieve such flexibility in two ways: either by having a unified internal model that is general enough to make predictions in such diverse scenarios, or a variety of domain-specific models, with the ability to switch between them. Domain-specific models (such as rigid-body and particle-based simulation) are commonly used in video game physics engines: being tailored to solve specific problems, they can make more efficient use of computational resources. Do humans also possess multiple internal models of the physical world, and the ability to switch between them? To answer this question, we designed a task to probe, in a controlled yet ecological way, how human participants interact with objects with diverse physical characteristics. Participants are shown, on a computer screen, an 'arena' including an object, a simple manipulator (a rectangular pusher) and a goal indicated by a shape and location within the arena. Using the keyboard, they can rotate and translate the pusher, and their task is to use it to push the object into the goal. On each trial the object belongs to one of four categories: box, rope, cloth and granular (e.g. a pile of coffee beans). The task varies widely depending on the object category: for example, a box can only be pushed into a specific position, while granular objects can be pushed into a given shape as well. We collect a large-scale online dataset of human responses on this task, and present preliminary results assessing systematic differences in behavior between object categories.

This research was funded by the European Research Council (ERC) Advanced Award 'STUFF' (ERC-ADG-2022- 101098225)

16.343 THE EFFECTS OF ACTION AND STIMULUS VISIBILITY IN EARLY VISUAL CORTEX

Bianca M van Kemenade[1,2], Lars F Muckli[2]; [1]Justus Liebig University Giessen, [2]University of Glasgow

Self-generated stimuli are often perceptually and neurally attenuated. This sensory attenuation is considered the result of predictions made using the efference copy, a copy of the motor command. However, several studies have shown that action can sometimes lead to enhancement instead of attenuation. It is currently still unclear which factors determine which effect action has on perception and neural processing. In the current study, we investigated whether the effect of action depends on the visibility of the sensory action outcomes. Participants completed both a behavioural and an fMRI experiment to test effects on both perception and neural processing. In both experiments, they had to perform an orientation discrimination task on two sequentially presented gratings. In active conditions, participants elicited the presentations by a button press, whereas in passive conditions, the gratings were presented automatically. We manipulated the visibility of the stimuli by embedding the gratings in noise during low visibility blocks. In high visibility blocks, the gratings were clear to see. Behavioural results showed significantly better performance for active compared to passive conditions. The Action x Visibility interaction showed a trend, with better performance for active conditions slightly more pronounced in the low visibility condition. This matched our univariate fMRI results, showing a significant Action x Visibility interaction in visual cortex, where an enhanced BOLD-response for active compared to passive conditions was present in low visibility conditions only. However, preliminary multivariate analyses showed no differences in decoding accuracy between active and passive conditions, neither in high nor low visibility conditions. Overall, these results suggest that action can boost both perception and the amplitude, but not precision, of neural responses particularly when the sensory action outcomes are noisy. This supports the idea that the effect of action on perception is dynamically modulated depending on context.

DFG (SFB/TRR 135, Project A10)

16.344 EFFECTS OF VIRTUAL REALITY USE IN CHILDREN 10- TO 12-YEARS-OLD

Ashley M. Clark[1] (aclark@exponent.com), Karol Silva[1], Nichole Breeland[1], Ilke Oztekin[1], Rachel Kelly[1]; [1]Exponent

Virtual Reality (VR) offers children immersive experiences in various environments, yet the impact of moderate and sustained VR use on visual functioning is not fully understood. This study evaluated the safety of VR exposure on children's visual functioning, with particular focus on several key aspects: visual acuity, stereoacuity, vergence facility, and accommodative response. Further, postural balance and visual/muscular discomfort were also assessed. Methods: Fifty participants, aged 10- to 12-years-old, engaged in 60-minute VR sessions for four consecutive days. Visual assessments were conducted at three time points: baseline (Day-1), after four days of daily VR use (Day-4), and following a break from VR use (Day-5). Visual assessments included visual acuity, stereoacuity, vergence facility, and accommodative response. Balance was evaluated using standardized postural stability tests and participants reported on their visual/muscular discomfort levels (e.g., symptoms of visually-induced motion sickness). Results: No significant changes were observed in visual acuity, stereoacuity, or vergence facility throughout the study, suggesting that moderate VR use over a period of four consecutive days does not show a significant impact on visual functioning in children. Accommodative response showed variability after VR exposure, consistent with previous research indicating that children possess a broader accommodative range compared to adults. These changes are not typically considered clinically significant and remained within normative values for this age group. Balance assessments revealed no significant alterations in postural stability from baseline to post- VR exposure. There were no significant changes in self-reported visual/musculoskeletal discomfort across the study period. Conclusion: These findings suggest that moderate, repeated VR use does not negatively affect visual functioning or balance in children. The study contributes to the growing body of evidence on the safety of VR use in pediatric populations. This research provides valuable insights into the potential use of VR in both educational and recreational contexts for children.

The current study was sponsored by Meta Platforms, Inc.

Attention: Inattention, load

Friday, May 16, 3:00 – 5:00 pm, Banyan Breezeway

16.345 TARGET PROBABILITY, RESPONSE DEMANDS, AND THE VIGILANCE DECREMENT

Henri Etel Skinner[1,2] (henri@ucsb.edu), Isabel Ruacho[1], Janani Aiyer[1], Barry Giesbrecht[1,2]; [1]University of California Santa Barbara, [2]Institute for Collaborative Biotechnologies

Vigilance is required when continuously monitoring for rarely occurring signals in the visual environment over time. When engaged in continous monitoring tasks, one robust phenomenon is a decline in detection rate as subjects continuously monitor for signals. One key parameter to this vigilance decrement may be the probability of target occurrence. We utilized the Continuous Temporal Expectancy Task (CTET) to observe performance across three different target probabilities (10%, 15%, 20%) in a within-subjects design. In this task, subjects (n=56) monitored a stream of gray-scale images appearing for brief durations (800 ms) and were instructed to respond when an image appeared for a slightly longer duration (1200 ms, targets). Furthermore, to discriminate between changes in sensitivity and response criterion, a signal detection theory approach was employed. A robust vigilance decrement, indicated by a decline in detection rate ($b = -0.031$, $se = 0.004$, $p < 0.001$), was observed across time. We also observed a decline in d-prime ($b = -0.093$, $se = 0.015$, $p < 0.001$) and an increase in criterion ($b = 0.052$, $se = 0.008$, $p < 0.001$) across time. There was an effect of probability on response rate, such that higher target probability yielded higher response rates ($b = 36.030$, $se = 1.969$, $p < 0.001$). Additionally, there was an effect of probability on d-prime, such that higher target probabilities yielded higher d-prime ($b = 0.936$, $se = 0.409$, $p < 0.05$). These findings together support the idea that the vigilance decrement is associated with both a decrease in sensitivity as well as an increase in criterion — suggesting attentional processes alone do not fully characterize vigilance performance. Furthermore, our results suggest the vigilance decrement is robust to modest changes in target probabilities and concurrent changes in response rate.

Research was sponsored by the U.S. Army Research Office and accomplished under cooperative agreement W911NF-19-2-0026 for the Institute for Collaborative Biotechnologies.

16.346 STIMULUS-DRIVEN AND GOAL-DIRECTED ORIENTING AND ENCODING INTO MEMORY ACROSS SPACE AND TIME

Carly Chak[1], Elise Kormos[2], Barry Giesbrecht[3]; [1]University of California, Santa Barbara, [2]Institute for Collaborative Biotechnologies

The selection and encoding of stimuli into working memory is dependent on the interplay between stimulus-driven and goal-directed attention. Typically, the contributions of stimulus-driven and goal-directed attention to the selection and encoding of information into working memory are studied separately. Here, we examined how these types of attention contributed to the selection and encoding of information as a function of working memory demands and individual differences in working memory capacity. Subjects (n=49) completed a spatially distributed n-back task where numbers were presented across six locations and prompted to report the 0-back, 1-back, or 2-back item. In the stimulus-driven condition, one number was displayed per frame. In the goal-directed condition, numbers were presented in each of the six locations all at once, but one cued the location of the next number to be attended. Accuracy declined when the to-be-reported number was temporally distant ($\eta2 = .11$, $p < .001$) and when attention was directed by the preceding number in the sequence ($\eta2 = .41$, $p < .001$). The magnitude of the n-back effect did not vary based on working memory capacity in either of the attention conditions. However, participants with low working memory capacity exhibited a speed-accuracy tradeoff in the goal-directed orienting condition such that they were slower in the 2-back condition, but maintained accuracy ($r = .78$, $p = .002$). Altogether, the results suggest that goal-directed orienting is particularly costly while maintaining information in working memory. Additionally, for participants with low working memory

capacity, the cost manifests as a delayed response time to the n-back with preserved accuracy. This suggests that low capacity participants may volitionally encode information in space, but the retrieval of target information may involve filtering out other information (e.g., other encoded numbers or location cues).

This research was sponsored by the U.S. Army Research Office and accomplished under cooperative agreement W911NF-19-2-0026 for the Institute of Collaborative Biotechnologies.

16.347 LAPSES OF SUSTAINED ATTENTION OCCUR WHEN GOALS COMPETE

Matthieu Chidharom[1], Monica Rosenberg[1], Edward Vogel[1]; [1]The University of Chicago

Sustained attention is the ability to maintain focus on a specific goal over time, but lapses in attention are frequent. Many theories have attributed these lapses to a transient failure of cognitive control in maintaining the goal in mind. However, these proposals have been challenged because recent findings have shown greater engagement of cognitive control during states more prone to lapses (Esterman et al., 2013). We hypothesized that lapses occur during periods of high competition between goals, requiring stronger cognitive control. A failure of managing control between competing goals during these periods could explain lapses, reconciling contradictory findings. To test this goal-competition hypothesis, we developed a Switch-Continuous Performance Task (CPT) in which subjects alternated task goals between blocks -either switching or holding the same goal- in an effort to manipulate periods of higher and lower competition between goals. Participants (N=30) viewed a bilateral display showing a scene (indoor/outdoor) and a face (male/female) on each trial. After every 20 trials a cue instructed participants to perform either the scene task (e.g., press for frequent indoor scenes, not infrequent outdoor scenes) or the face task (e.g., press for frequent male faces, not infrequent female faces). Results showed more attention lapses during switch periods than repeats, suggesting that lapses occur during periods of high competition between goals. In a second study (N=30), we monetarily rewarded performance on only one goal (scene or face) to create unequal competition between goals. We found that switching to a rewarded goal did not induce more lapses, whereas switching to an unrewarded goal produced more lapses in sustained attention. In a third study (N=20), we recorded EEG activity to isolate the neural mechanisms underlying goal competition. Together, these findings support the goal-competition hypothesis as an explanation for the occurrence of sustained attention lapses.

16.348 CHANGE BLINDNESS: THE IMPACT OF MOTION AND PERCEPTUAL LOAD

Rachel Pitman[1] (18rp39@queensu.ca), Daryl Wilson[1]; [1]Queen's University

Despite the dynamic nature of real-world environments, previous research on change detection has primarily used static stimuli. However, research exploring the impact of motion on attention (Suchow & Alvarez, 2011) and memory (Blalock et al., 2014; Chung et al., 2023), two necessary components for successful change detection, suggests that motion can impair both processes. Consequently, the objective of the present experiment was to determine whether motion impairs change detection, as well as whether different motion types (i.e., synchronous and asynchronous) have different effects on detection accuracy. Additionally, we sought to determine whether perceptual load moderates the impact of motion. To address these objectives, we conducted a gradual change blindness experiment in which participants were presented with task-relevant (colourful, randomly-oriented isosceles triangles) and task-irrelevant (gray circles) stimuli that were either stationary, moving synchronously, or moving asynchronously. In each trial, one task-relevant stimulus gradually changed while participants attempted to identify the change target. Change detection was examined as a function of Change Type (Color, Orientation), Motion Type (Stationary, Synchronous, Asynchronous), and Load (3, 6, 9, 12 task-relevant stimuli). Interestingly, results showed that Motion Type significantly affected change detection for orientation changes, but not for color changes. Within the Orientation condition, detection accuracy was highest in the Stationary condition, lower in the Synchronous condition, and lowest in the Asynchronous condition. As hypothesized, we observed an interaction between Load and Motion in the Orientation condition, such that the effects of motion were absent when load was low (Load = 3), but emerged at higher loads (Load = 6, 9, 12). It was concluded that motion only impaired orientation change detection because, unlike color, the inherent movement of the orientation changes were masked by the Synchronous and Asynchronous movement.

16.349 BASIC CHARACTERISTICS OF INATTENTIONAL BLINDNESS

Ronald A. Rensink[1] (rensink@psych.ubc.ca); [1]University of British Columbia

Although inattentional blindness (IB) has been the focus of many studies, questions remain as to its nature. First, does the failure to report an item result from a lack of attention, or a lack of expectation? Second, does it correspond to a failure to consciously experience the item, or to remember it—i.e., is it a form of blindness or amnesia? And third, how is functionality affected when visual experience is not reported? Earlier attempts to answer these questions had various limitations. To overcome these, a locked-onset technique had a test item briefly appear while other items around it were tracked (Rensink, VSS 2005). Test items were in 50% of 300 trials given to each observer. If IB results from a lack of expectation, test items should almost always be seen; in contrast, the diversion of attention due to tracking should cause these to be missed. In addition, test items were endogenous cues. If cues reported as unseen differ in their effect from those reported as seen, it would imply that cues reported as unseen were not briefly experienced (and thus acted normally) before being forgotten. Four kinds of cue were used: (i) simple arrow, (ii) arrow embedded in a circle, (iii) triangle embedded in a square, or (iv) reversed arrow; these were shown for 100, 200, or 600 ms. Running 30 observers per condition indicates that (i) IB is due to a lack of attention rather than expectation, (ii) it mostly reflects a failure of perception rather than memory, and (iii) the absence of conscious visual experience during IB slows the extraction of figural elements. These results suggest that much of visual processing occurs in the absence of conscious visual experience, with such experience enabling the control of selective attentional processing, allowing perception to be sped up in nonroutine situations.

16.350 ALPHA TACS OVER RIGHT DORSOLATERAL PFC AND ITS IMPACT ON ATTENTIONAL BLINK

Sahereh Varastegan[1], Weiwei Zhang[2]; [1]Ph.D. Student at University of California, Riverside, [2]Professor of Psychology at University of California, Riverside

Dynamic attentional mechanisms play a critical role in goal-directed behavior by filtering and prioritizing relevant information for further processing, such as working memory encoding. However, during Rapid Serial Visual Presentation (RSVP), these processes may falter, resulting in what is known as the attentional blink. Recent evidence has linked this phenomenon to alpha oscillations within the frontoparietal network which are important for top-down attentional control and working memory encoding. Specifically, reduced alpha synchronization is associated with increased attentional blink. It is thus hypothesized that neuromodulation of alpha activity via tACS would impact the magnitude of the attentional blink . The present study investigated how 10 Hz transcranial Alternating Current Stimulation (tACS) over the right dorsolateral prefrontal cortex (DLPFC) influences attentional blink. While performing the attentional blink task with RSVP of two letter targets among distractor digits, participants received 20 to 25 minutes of 10 Hz tACS at an intensity of 2mA versus a sham stimulation condition (30 seconds of stimulation including a 10-second fade-in/fade-out periods). It was predicted that the active stimulation should reduce attentional blink magnitude at lag 3, but not at lag 1 or 7. Contrary to this prediction, a repeated-measures ANOVA revealed significant main effects of stimulation condition (active vs. sham) and lag (lags 1, 3, and 7) but no significant interaction effect. Overall, the results demonstrated that alpha tACS improved performance on the AB task across all lags. This improvement may reflect reduced working memory demands associated with maintaining the first target (T1), thereby facilitating the processing of the second target (T2).

16.351 CATCHING THE WANDERING MIND WITH REAL-TIME TRIGGERS

Shivang Shelat[1,2] (sshelat@ucsb.edu), Jonathan W. Schooler[1], Barry Giesbrecht[1,2]; [1]University of California, Santa Barbara, [2]Institute for Collaborative Biotechnologies

From moment to moment, our attention fluctuates between optimal and suboptimal states. This flux has been studied using objective indices—such as response time (RT) patterns in laboratory tasks—and subjective measures, like self-reports of task-unrelated thoughts (i.e., mind-wandering). Here, we assessed the predictive utility of a real-time RT-based triggering procedure in capturing both reduced working memory encoding and the phenomenological experience of mind-wandering across various individual differences. Fifty-eight participants completed trait measures of motor impulsivity, mind-wandering propensity, and social desirability bias. Afterwards, they performed a monotonous sustained attention task that demanded a keypress on each trial and lulled them into inattentive responding. When their three-trial RT was one standard deviation above or below their cumulative RT mean, they were interrupted with either a working memory probe to report colors in locations from the previous trial or a mind-wandering probe to report the task-relatedness of their thoughts. Fast RT triggers were associated with reduced working memory encoding compared to slow triggers ($b = 0.19$, $SE = 0.045$, $p < .0001$). Moreover, fast triggers predicted mind-wandering self-reports when controlling for social desirability bias ($b = -0.34$, $SE = 0.14$, $p = .014$). The coupling between these objective and subjective measures of lapsing attention was also weaker among individuals who claimed to have socially desirable traits ($b = 0.058$, $SE = 0.026$, $p = .027$). Together, these findings both 1) reveal a self-presentation bias that obscures attention self-reports and 2) validate the predictive utility of RT-based triggering procedures in anticipating multiple signatures of attentional lapses. We highlight the need for an integrative approach to characterizing the human tendency to decouple from our environment by augmenting subjective techniques with objective ones.

This work is supported by a National Science Foundation Graduate Research Fellowship awarded to the first author under grant number 2139319 and by contract W911NF-19-2-0026 from the US Army Research Office.

16.352 ORIENTING OF INTERNAL ATTENTION BETWEEN SHORT- AND LONG-TERM MEMORY

William Narhi-Martinez[1], Pranava Dhar[1], Kaiki Chiu[1], Anna C. Nobre[1]; [1]Yale University

We investigated how internal attention is directed to contents within or between short-term (STM) and long-term memory (LTM) across two in-person behavioral studies. Many contexts in daily life are associated with both LTM and STM content. We wanted to understand how attention operates when directed to one versus both of these internal domains. For LTM, participants learned the identities and locations of two objects within each of 48 scenes. The next day, the same scenes appeared with two new objects in either different (Experiment 1) or occasionally the same (Experiment 2) locations as the LTM objects. Participants encoded these new items into STM. At the end of the trial, an old item and foil appeared, and participants chose the old item. The target could come from either LTM or STM. During the STM delay, retrocues could guide internal attention. In Experiment 1, retrocues were spatial and indicated the side of two objects (both LTM, both STM, or one from each). In Experiment 2, retrocues were either spatial or neutral, indicating all four objects should continue to be maintained. The spatial retrocues could indicate two items with overlapping or non-overlapping locations, one from each memory domain. Our results showed significant performance benefits of internal attention for both LTM and STM, with retrocue benefits being stronger for STM than LTM probes. Surprisingly, spatially overlapping LTM and STM associations had relatively little effect. Interestingly, performance on LTM probes was significantly worse when retrocues prioritized items from both LTM and STM domains rather than the LTM domain alone. Our study paves the way for studying flexible internal attention for contextual associations in different memory domains. The initial findings challenge the standard view that once a LTM is recalled into working-memory, it is maintained in an equivalent format to new information encoded from sensory stimulation.

16.353 LIMITED SCOPE OF THE FUNCTIONAL FIELD OF VIEW IN CEREBRAL VISUAL IMPAIRMENT

Claire Manley[1], Peter Bex[1], Lotfi Merabet[2]; [1]Northeastern University, [2]Massachusetts Eye and Ear Infirmary, Harvard Medical School

Cerebral visual impairment (CVI) is a brain-based visual disorder associated with early damage and maldevelopment of retrochiasmal visual processing pathways and areas. Individuals with CVI often face difficulties with tasks such as visual search, navigation, scene recognition, reading, and driving, requiring more time and cognitive effort compared to typically developing individuals. A common feature among these tasks is the need to divide visual attention between central and peripheral vision, reflecting the functional field of view (FFV). We assessed the FFV in 10 participants with CVI (meanAge=24.09 years, SD=11.08) and 6 control participants (meanAge=20.00 years, SD=5.00). Two characters ("X"s and "O"s), one foveal and one peripheral (4°, 8°, 16°), were presented at size and duration combinations determined by a QuestPlus algorithm. In a 2AFC task, participants indicated whether the characters were the same or different. The area under the curve (AUC) for target size and duration decreased with eccentricity and was higher for the control group (mean = 0.78, SD = 0.09) (Figure 1.1) compared to the CVI group (mean=0.63, SD=0.14) [F(1,46)=20.84, p= 0.00004] (Figure 1.2, Figure 2). For duration, there was a significant effect of group [F(1,46)=25.37, p=9.459e-06], but not of eccentricity [F(1,44)=0.7204, p=0.4925]. The CVI group required longer durations for accurate identification (meanCVI=0.15, SD=0.08; meanControls=0.05, SD=0.04). For target size, however, there were significant effects of both group [F(1,46) = 15.56, p = 0.0003] and eccentricity [F(1,45)=16.55, p=0.0002]. The CVI group required larger target sizes across eccentricities (meanCVI=0.57, SD=0.32; meanControls=0.29, SD=0.18). This suggests that for both groups, target size must increase with eccentricity for accurate identification. However, processing time does not significantly increase with greater eccentricity. This research highlights the perceptual deficits experienced by individuals with CVI. In the periphery, they require more time and larger stimuli to accurately perceive information in parallel with stimuli in the central visual field.

16.354 FACE PERCEPTION IS RESISTANT TO REDUNDANCY MASKING

Lillian Hauser[1], Timothy Sweeny[1]; [1]University of Denver

Redundancy masking studies have found that observers frequently underreport the number of items (e.g., lines and shapes) in a repeating pattern shown in the periphery. This phenomenon is speculated to compress the complexity of the visual world, simplifying processing demands. The cost of losing awareness of a few items may be tolerable for sets of geometric features, but it may be less desirable when perceiving socially important features like faces. Here we compared the strength of redundancy masking with four types of stimuli: upright faces, inverted faces, rectangles, and lines. We hypothesized that upright faces would be less susceptible to redundancy masking than these other features. On each trial, three to six radially arranged, evenly spaced (edge-to-edge: 1°) identical stimuli were presented in either the left or right periphery, centered 39° from fixation. Observers reported the number of stimuli they perceived in each set, ranging from zero to nine. Observers underestimated the number of rectangles the most (M = -0.39), followed by lines (M = -0.25) and inverted faces (M = -0.01); conversely, observers overestimated the number of upright faces in a set (M = 0.09). This distinction between upright and inverted faces is especially notable because it cannot be attributed to differences in visual complexity. Further, the upright faces' greater resistance to redundancy masking (compared with inverted faces) only occurred when sets were viewed in the left visual field, consistent with the right hemisphere's dominance in face processing. These findings indicate that specialized facial processing systems may help combat redundancy masking.

16.355 INVESTIGATING THE IMPACT OF STRESS ON OVERT ATTENTION DURING AN EYEWITNESS EVENT

Abigail Kortenhoeven[1] (abkern@ttu.edu), Michael Serra[2], Miranda Scolari[1]; [1]Texas Tech University, [2]Texas Tech University Health Sciences Center

Eyewitness testimony is the primary source of evidence in roughly 77,000 criminal cases in the United States each year, but its accuracy is highly variable. Research suggests high levels of stress are linked to impaired accuracy of eyewitness testimony, but the underlying cause of this impairment is unknown. Recent work suggests that individuals in a state of stress show an attentional bias towards threat-related stimuli compared to non-threatening stimuli. The current study seeks to make connections between stress and overt visual attention during a criminal event. The experimental group completed the Maastricht Acute Stress Task (MAST), alternating a mental arithmetic task with placing their hand on an ice pack inside a Ziploc bag. The control group used an upwards counting task and a room temperature icepack. Both groups watched CCTV footage of an armed robbery at a convenience store made publicly available by the Denver Police Department while an Eyelink 1000 tracker captured eye movements across pre-assigned dynamic interest areas. Self-reported stress measures, blood pressure, and pulse readings were used to assess participant stress levels prior to and after viewing the CCTV footage. Current trends within the preliminary self-report data (N = 10) suggest that the experimental group experienced higher levels of stress (M = 41.6) than the control group (M = 26.75) following the protocol. Importantly, the eye-tracking data provide support for our hypothesis: Comparing each interest area across groups indicates that the experimental group spent significantly more time fixating the perpetrators, while the control group spent significantly more time fixating the weapons. This suggests that stress may impact the visual details individuals overtly attend to while witnessing a crime. Future analyses will explore whether these different eye movement patterns during encoding predict accuracy on a subsequent memory task.

16.356 THE TIME COURSE OF AN ERROR: BENEFITS OF EXAMINING PRE-ERROR SPEEDING AND POST-ERROR SLOWING SIMULTANEOUSLY

Sarah B. Malykke[1], Audrey Siqi-Liu[1], Kelvin S. Oie[3], Dwight J. Kravitz[1,2], Stephen R. Mitroff[1]; [1]The George Washington University, [2]US National Science Foundation (SBE/BCS), [3]U.S. DEVCOM Army Research Laboratory

Errors are inevitable in human behavior, and thus it is critical to understand their causes and effects, especially in applied settings. Prior research has shown that individuals often speed up before an error, perhaps due to attentional lapses or diminished cognitive control. Conversely, after an error, individuals typically slow down, which may reflect disrupted processing or an adaptive increase in caution. While much research has examined pre-error speeding or post-error slowing independently, here we examined both phenomena simultaneously to better understand how errors emerge and disrupt performance, with a specific goal of exploring how pre- and post-error behaviors might relate. We examined error performance in a very large dataset (n=17,784) from a mobile game (Airport Scanner), where individuals completed an object-sorting task which involved categorizing items by touching the item on the screen and swiping it up or down into a bin. A linear mixed-effects model revealed consistent speeding leading up to an error (as early as 12 trials before), followed by significant and long-lasting post-error slowing (up to 15 trials after). Correlation matrices focused on trial-to-trial performance revealed strong relationships between adjacent trials leading up to an error, reflecting systematic behavior across participants as they approached the error. However, upon the error, correlations between adjacent trials were disrupted, suggesting a breakdown in the trial-by-trial behavioral consistency. Post-error, the trial-by-trial correlations gradually recovered demonstrating the reestablishment of response time carryover effects. These relationships were absent in a group of matched participants who did not make an error. These findings highlight the importance of studying pre-error speeding and post-error slowing together to better understand how errors influence task performance over time—and how pre-error performance may be indicative of both error occurrence and post-error behavior.

W911NF-23-2-0210, W911NF-24-2-0188

16.357 MAYBE YOU CAN HAVE IT BOTH WAYS: NO STABILITY-**FLEXIBILITY "TRADEOFF" WITH ITEM**-SPECIFIC CONTROL

Rebeka Almasi[1], Emma Wiedenmann[1], Myeong-Ho Sohn[1]; [1]The George Washington University

Cognitive stability refers to remaining on task and ignoring distractors, while flexibility refers to switching between task rules. "Tradeoff" theories of stability and flexibility suggest that if one is engaged, the other is inhibited, like a see-saw. Allocating control to an item, under this view, should result in greater persistence of the task rule upon switching to a new task, yielding higher switch costs. Under other theories that posit independence of stability and flexibility, recruiting stability for an item should not affect future attempts at interacting flexibly with it, or vice versa. By creating item-specific associations with higher switching or higher stability requirements, this research examines the independence of flexibility and stability in visual attentional control. The association phase (804 trials) consisted of task switching (age identification vs gender identification tasks determined by frame color; Experiment 1) or an emotional Stroop task (Experiment 2). Half of the identities shown were associated with the more challenging task (Mostly Task Switching or Mostly Incongruent Stroop, respectively) while half were associated with the easier task (Mostly Task Repetition or Mostly Congruent Stroop). The subsequent transfer phase (160 trials) tested the impact of these associations. For Experiment 1, participants performed an Emotion Stroop task. Crucially, the stimuli were the same faces previously associated with a switching likelihood–so, if participants are less "stable" on faces they learned were Mostly Switching faces, their congruency effects should be larger. In Experiment 2, the transfer phase consisted of task switching, where we examined whether switch costs were associated with the faces' associated proportion of congruency. In both cases, we found no evidence of impact on future engagement of control.

Attention: Individual differences

Friday, May 16, 3:00 – 5:00 pm, Banyan Breezeway

16.358 ARE THERE QUALITATIVE INDIVIDUAL DIFFERENCES IN SPATIAL ATTENTION?

Felipe Luzardo[1], Yaffa Yeshurun[1]; [1]University of Haifa

The ability to prioritize certain regions in our surroundings, known as spatial attention, is critical for human cognition and has been the subject of extensive research. However, studies rarely employ large population samples, resulting in limited exploration of inter-individual variability. Particularly, the question of whether individual differences in spatial attention are qualitative or quantitative remains unaddressed. Quantitative individual differences refer to variability only in the magnitude of attentional effects, while qualitative individual differences imply fundamentally distinct patterns of attention allocation, across individuals. Exploring this distinction is crucial because qualitative differences may reveal diverse underlying mechanisms beyond simple variations in magnitude. We recruited a sizable participant pool (N=514) across three experiments. We used an acuity task paired with valid, invalid, and neutral pre-cues of different types. The inclusion of a neutral cue allowed us to analyze both the averaged attentional benefits—enhanced performance when focusing on the correct location—and costs—diminished performance when attention is directed to the wrong location—and their variability across individuals. The results revealed robust attentional benefits, but not attentional costs, suggesting that facilitatory and inhibitory mechanisms may operate independently. Importantly, comparisons between different neutral cue types revealed no significant differences in performance, indicating that the neutral cue's characteristics are not critical. Hierarchical Bayesian analyses uncovered true qualitative individual differences; while most participants demonstrated effects in the anticipated direction, some exhibited true effects in the opposite direction. This surprising finding highlights the intricacy of attentional allocation, indicating that a thorough understanding of spatial attention must account for multiple underlying mechanisms that may result in attentional effects in the opposite direction, such as individual differences in levels of internal noise or inhibition of return. These results also emphasize the contribution of large sample sizes to uncover and better understand the full spectrum of cognitive profiles among individuals.

16.359 EXAMINING THE RELATIONSHIP BETWEEN INDIVIDUAL DIFFERENCES IN VISUAL SENSITIVITY AND DISTRIBUTION OF COVERT ATTENTION AROUND THE VISUAL FIELD

Zainab Haseeb[1], Anna Kosovicheva[1]; [1]University of Toronto Mississauga

Visual performance varies across the visual field and differs between individuals, potentially contributing to errors in everyday tasks. These variations may stem from anatomical factors, such as photoreceptor density, as well as attentional factors, including the allocation of covert attention. This study investigated the relationship between individual differences in low-level spatial resolution and the distribution of covert attention across the visual field. In addition, we examined participants' awareness of these performance differences across the visual field. Participants performed two distinct tasks: a line bisection task to evaluate low-level spatial resolution, and a visual search task to examine how covert attention is distributed across the visual field, while maintaining fixation at the center of the display. Stimuli were presented at 4º eccentricity across 4 cardinal and 4 diagonal locations. The bisection task measured spatial resolution by having participants judge whether the stem of a T-shape was offset to the left or right across ten levels (-0.12-0.12º). The second task assessed allocation of covert attention with a search task in which participants reported the orientation of a rotated T among Ls, placed in a random location around fixation. All stimuli were presented at the same 4º eccentricity. To assess whether participants were aware of their performance differences around the visual field, participants provided confidence ratings for their responses in both tasks. Results revealed no significant relationship between spatial resolution in the T-bisection task and accuracy in the search task, suggesting that these measures reflect distinct aspects of visual processing. This was further supported by confidence ratings, which tracked accuracy in the search task, but did not correlate with accuracy in the bisection task. Together, these findings point to distinct mechanisms underlying performance asymmetries around the visual field.

This work was supported by an NSERC Discovery Grant to AK

16.360 PRESERVED EXOGENOUS ATTENTION IN THE MYOPIC PARAFOVEA

Katia steinfeld[1] (katia.steinfeld@unil.ch), Micah M. Murray, Marisa Carrasco; [1]University of Lausanne and Lausanne University Hospital, Switzerland, [2]The Sense Research and Innovation Institute, Switzerland, [3]New York University, Department of Psychology and Center for Neural Science, USA

Background: Exogenous attention enhances visual perception across the visual field, often with stronger effects in the periphery, where perception is poorest. Myopes, even with corrected-to-normal vision, exhibit reduced acuity and contrast sensitivity in the peripheral visual field compared to emmetropes. Here we investigated whether the extent of the effects of exogenous attention vary as a function of myopia. Methods: Participants, all with corrected-to-normal vision, performed a two-alternative forced-choice (2AFC) task to discriminate the orientation of tilted Gabor patches presented at 6° eccentricity on the horizontal meridian while maintaining central fixation. Attention was manipulated using cues presented above the Gabors shortly before (50 ms) their appearance. Cues were valid, invalid, or neutral, presented randomly. The tilt angle for each participant was determined using a Weibull staircase, converging at 80% accuracy. Results: Both groups exhibited attentional effects, with valid cues increasing performance (d'), and invalid cues decreasing performance, compared to neutral cues. The magnitude of these attentional effects was comparable between myopes and emmetropes. No speed-accuracy trade-off was observed in either group. Conclusion: Myopia, a condition associated with axial eye elongation and typically developing in late childhood, leads to peripheral visual deficits. Our findings suggest that exogenous attention can partially compensate for these deficits in corrected-to-normal adults. Currently, we are exploring whether this enhancement extends beyond the horizontal meridian to other regions of the visual field.

Swiss National Foundation MD-PhD 323630_207031, Swiss National Science Foundation (169206), NIH NEI R01-EY027401 to MC

16.361 A MACHINE LEARNING APPROACH FOR PREDICTING SINGLE SUBJECT PERFORMANCE FROM EYE METRICS

Alexis Torres[1], Gene Brewer[2]; [1]Arizona State University, [2]University of California, Riverside

Sustained attention declines over time, leading to increased attention failures. Eye metrics, such as prestimulus pupil size, stimulus-evoked pupil responses, and gaze stability, have been identified as indices of task engagement and predictors of attention failures at the group level. This study evaluates the predictive power of eye metrics for impending attention failures at the trial level in single subjects using machine learning algorithms. In this study, 225 undergraduate students completed the psychomotor vigilance task. Eye metrics—including mean prestimulus pupil size and variability, mean stimulus-evoked pupil response amplitude and latency, and preparatory gaze stability—were extracted for each trial. Multiple machine learning algorithms were trained to predict and classify response times based on these metrics at both the group and subject levels. Group-level predictions showed robust accuracy, with models effectively predicting trial-level response times. However, single-subject trial predictions exhibited lower sensitivity, particularly for slower response classifications. These findings suggest that while eye metrics are reliable predictors at the group level, additional refinements are needed to improve their predictive power for individual performance. This highlights the potential of eye metrics as tools for understanding attention dynamics and the challenges of translating group-level insights to single-subject applications.

National Defense Science and Engineering Graduate (NDSEG) Fellowship Program

16.362 NUMERACY AND NAVIGATION: ASSESSING MATHEMATICAL DISPOSITIONS THROUGH EYE TRACKING

Joanna Lewis[1], Caterina Azzarello[1], Gulden Karakok[1], Katherine Morrison[1], Molly Jameson[1]; [1]University of Northern Colorado

This study investigates the mathematical dispositions of pre-service elementary teachers (PSET), focusing on their attitudes toward mathematics and the impact of visual attention. Mathematical dispositions, encompassing attitudes, behaviors, and thoughts about math, significantly influence teaching effectiveness and student learning outcomes. Productive dispositions, characterized by curiosity and persistence, contribute to improved performance, whereas unproductive dispositions, marked by anxiety and avoidance, hinder both teaching and learning (Boaler, 2016; Dweck, 2006). PSET often exhibit unproductive math dispositions, leading to avoidance behaviors and a lack of confidence in teaching mathematics (Alkhateeb, 2014). These negative beliefs can result in reduced classroom time devoted to math and the modeling of unproductive dispositions to students, particularly impacting female students (Beilock et al., 2010; Phillip, 2007). The study involved 41 PSET enrolled in an introductory math for elementary teachers course. Participants completed surveys assessing math anxiety and math self-efficacy. Participants engaged in an arithmetic task while undergoing eye tracking using the Eyelink 1000+ system, which recorded their eye movements as they calculated and assessed responses to basic

arithmetic questions of correct or incorrect calculations. The eye-tracking data provides insights into how PSET visually engages with math tasks, potentially revealing patterns linked to their dispositions. We found that individuals with higher math anxiety tended to have lower fixation rates and durations after a math course. This suggests that PSET with high math anxiety may avoid visually engaging with arithmetic. Given this knowledge, interventions to reduce PSET math anxiety may be useful tools to address their avoidance.

NSF Award #2235588

SATURDAY MORNING POSTERS IN BANYAN BREEZEWAY

Plasticity and Learning: Clinical

Saturday, May 17, 8:30 am – 12:30 pm, Banyan Breezeway

23.301 THE EFFECTS OF EARLY AND LATE ONSET BLINDNESS ON THE STRUCTURE OF **HESCHL'S GYRUS**

Amy Poole[1], Kelly Chang[2], Feiyi Wang[3], Ione Fine[2], Woon Ju Park[4]; [1]University of Minnesota, [2]University of Washington, [3]Boston University, [4]Georgia Institute of Technology

Early blindness results in dramatic changes in the structure and function of the primary visual cortex. Here, we examined the structural plasticity in the auditory cortex, specifically the Heschl's gyrus (HG). Previous work shows increased HG folding in professional musicians than non-musicians, suggesting a possibility that extensive auditory experience resulting from visual loss might alter HG structure in blind individuals. We analyzed T1-weighted images collected from previous MRI studies at the University of Washington (Jiang et al., 2016; unpublished data), University of Pennsylvania, and Oxford University (Aguirre et al., 2016; Bridge et al., 2009). The combined dataset included 6 anophthalmia (individuals born without eyes), 48 early blind, 18 late blind, and 28 sighted control participants. Hand-drawn HG regions of interest for each participant in both hemispheres were created. HG is known to show high morphological variability across individuals. We characterized HG structure by 1) visually categorizing HG morphology (single, partial, or complete duplication of the gyrus) and 2) using continuous metrics including curvedness index, gyrification index, thickness, gray matter volume, and surface area. In contrast to the deprived primary visual cortex, the structure of HG was not altered by blindness. A chi-squared test revealed that the HG morphology was not significantly different across the four groups. Linear mixed-effects models (controlling for age, hemisphere, and scan location) similarly showed no effects of group on the gyrification index, curvedness index, thickness, gray matter volume, and surface area. Our findings suggest that visual loss does not alter the structure of the non-deprived auditory cortex. Structural plasticity following blindness might be limited to the occipital cortex, challenging the idea that extensive experience can shape the anatomical properties of primary sensory cortices. The results in musicians likely reflect a selection bias, where individuals with increased HG folding are more likely to become musicians.

NEI R00 EY034546 to WP; NEI R01 EY014645 to IF

23.302 NEUROPLASTICITY IN A HUMAN ROD-ONLY VISUAL SYSTEM RECEIVING CONE RESCUING GENE THERAPY

Roni Maimon-Mor[1,2] (r.maimon@ucl.ac.uk), Mahtab Farahbakhsh[1,2], Andrew Rider[1], Mohamed Katta[1,3], Andrew Stockman[1], Michel Michaelides[1,3], Tessa Dekker[1,2]; [1]UCL Institute of Ophthalmology, University College London, London EC1V 9EL, UK, [2]Experimental Psychology, University College London, London WC1H 0AP, UK, [3]Moorfields Eye Hospital, London EC1V 2PD, UK

Recent advances in cell- and gene-therapies make it increasingly possible to halt or reverse vision loss for many eye conditions. To develop effective treatments, we need to understand how rescued retinal signals integrate into developing brain function. One promising gene therapy treatment is for the congenital genetic condition achromatopsia, which causes retinal cone dysfunction. In a typically developing brain, cones are responsible for providing fast, chromatic information from across the retina, and all information from the fovea. In research alongside pioneering gene therapy trials, we measured cone recovery in 7 patients with achromatopsia treated between the ages of 8 and 16 years of age. This scenario offers an unprecedented opportunity to untangle the role of early experience in shaping neural specialisation for rod- and cone-mediated functions. Using photoreceptor-specific pRF mapping (fMRI), we found robust and consistent rescue of retinal cone-mediated signals in visual cortex in the majority of children, which was not present before treatment. Crucially, each of these patients could use these rescued signals to detect cone-selective stimuli previously invisible to them. Next, in two patients, we used high-precision psychophysics to determine which additional new visual functions the rescued retinal signals could support, focusing on cone-specific functions not available through rod-only vision: high-speed, foveal, and chromatic (colour) vision. After treatment, visual processing speed was substantially higher. Foveal cortex showed limited cone-mediated function, in line with previous reports of inconsistent benefits for acuity. Several complementary tests for colour vision (lvvCCT, Ishihara, Rayleigh, Stockman) revealed different recovery patterns and perceptual experiences across the two patients. Together this work provides novel insights into how a brain that has developed and organised with one set of retinal signals (rods) can incorporate a whole new parallel set of signals (cones) into its existing structures, and which novel functions can be restored.

NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, the Economic and Social Research Council (ESRC) of the UKRI; MeiraGTx, Retina UK, Moorfields Eye Charity, Foundation Fighting Blindness (USA) and The Wellcome Trust

23.303 CORTICAL AND SUBCORTICAL STRUCTURAL CHANGES IN PATIENTS WITH ARTIFICIAL VISION RESTORATION

Stephanie Saltzmann[1], Noelle Stiles[1], Jeiran Choupan[2], Yonggang Shi[2]; [1]Rutgers University, [2]University of Southern California

The Argus II retinal prosthesis is designed to electrically stimulate the remaining retinal ganglion cells in patients with Retinitis Pigmentosa, a condition that causes progressive vision loss. Research indicates that the onset of blindness, especially at an older age, leads to significant structural changes in the brain, particularly in the cortical and subcortical regions associated with spatial processing, such as spatial navigation, awareness, and memory. These brain regions undergo reorganization to compensate for the loss of vision, and it remains uncertain how visual restoration through devices like the Argus II influences the thickness or volume of these areas compared to sighted individuals. This study measured cortical and subcortical thickness in three groups: ten Argus II patients, ten blind individuals, and thirteen sighted controls. The focus was on regions important for spatial cognition, such as the parahippocampal gyrus and the hippocampus. The findings revealed that, on average, Argus II patients showed a thicker left parahippocampal gyrus than blind patients, suggesting that the device may facilitate structural changes in the brain. In the two case studies, the shorter-using (6.5 months) and longer-using (44.5 months) patients showed an increase in the thickness of the parahippocampal gyrus. In the third case study, there was an increase in the thickness of the left parahippocampal gyrus from 43.5 to 57 months and beyond 106 months of disuse. The right parahippocampal gyrus thickness and right hippocampus volume positively correlated with the duration of experience with the device. The results suggest that using the Argus II device may facilitate structural plasticity in the brain, particularly in spatial memory and navigation regions. This finding underscores the potential for the brain to adapt and reorganize, even in late-blind patients, as they relearn to interact with their environment through restored vision.

The National Institutes of Health, National Eye Institute (1U01EY025864); the National Institutes of Health, the BRAIN Initiative (5K99EY031987 and R00EY031987).

23.304 A COMPUTATIONAL VIRTUAL PATIENT PIPELINE FOR PREDICTING PERCEPTUAL CAPABILITIES WITH VISUAL PROSTHESES

Jonathan Skaza[1], Shravan Murlidaran[1], Apurv Varshney[1], Ziqi Wen[1], Miguel P Eckstein[1], Michael Beyeler[1]; [1]University of California, Santa Barbara

Introduction: Visual prostheses aim to restore sight, but current devices generate only rudimentary phosphenes with limited visual capabilities. Predicting patient perception before implantation is crucial for evaluating device performance, optimizing design, and setting realistic expectations. We present a novel Computational Virtual Patient (CVP) pipeline to predict perceptual performance. Methods: The CVP pipeline simulates prosthetic perceptual experiences using two models: a traditional “scoreboard” approach and an “axon map” model that leverages the spatial layout of retinal ganglion cell axons (Beyeler et al., 2019). These simulations, constrained by psychophysical and neurophysiological data, allow customization of electrode configurations, including tests with 6x10 (Argus II), 6x15, and 12x20 arrays. Sighted participants (n = 18) completed shape and facial emotion classification tasks using simulated phosphenes. Additionally, a ResNet-18 deep neural network (DNN) was fine-tuned through transfer learning, with its final classification layer retrained on a simulated phosphene dataset to predict perceptual performance. Results: The DNN accurately predicted the hierarchical ordering of task difficulty, with binary emotion detection (64.2% with the “axon map” 12×20 configuration) proving more challenging than three-class shape classification (84.3% with the same configuration). Human and DNN results showed similar effects of electrode configuration on perception. For shape classification tasks, increasing electrode resolution from 6×15 to 12×20 improved accuracy in humans (64.2% ± 1.7% to 85.7% ± 1.3%) and the DNN (62.6% ± 3.6% to 84.3% ± 2.6%). Fully trained ResNet-18 and in-house DNNs were not consistent with human performance, suggesting that humans may rely on pre-existing visual processing rather than developing new perceptual mappings. Conclusion: Our computational framework demonstrates the potential to predict prosthetic vision capabilities across multiple tasks and devices, offering a novel approach for evaluating visual prostheses pre-implantation. This tool could accelerate device development and provide more accurate expectations for implant recipients.

23.305 SUBACUTE TRAINING RECOVERS MOTION DISCRIMINATION AFTER EARLY VISUAL CORTEX DAMAGE DESPITE RETROGRADE DEGENERATION OF THE RETINAL GANGLION CELL COMPLEX

Berkeley Fahrenthold[1], Bryan Redmond[1], Rachel Hollar[1], Jingyi Yang[1], Matthew Cavanaugh[1], Duje Tadin[1], Marisa Carrasco[2], Krystel Huxlin[1]; [1]University of Rochester, [2]New York University

Occipital strokes that damage primary visual cortex (V1) and surrounding early visual cortical areas cause homonymous visual deficits, commonly referred to as cortically-induced blindness. These strokes also trigger trans-synaptic retrograde degeneration across early visual pathways. In chronic stroke patients, degeneration at the level of the optic tracts has been shown to predict training-induced visual recovery (Fahrenthold et al., Brain 2021). Here, we investigated whether a similar relation exists in subacute patients (<6 months post-stroke) at the retinal level. Twenty-three patients (49 ±10 years old, 13 males) were recruited 1.2–5.1 months post-occipital stroke as part of an ongoing clinical trial (NCT04798924). Ganglion cell and inner plexiform layer (GCL-IPL) thicknesses on either side of the fovea [corresponding to the blind and intact hemifields] were measured at baseline using optical coherence tomography (OCT) and used to calculate a laterality index (LI). Participants underwent visual training on a direction discrimination task with feature-based attention in two phases. During the subacute period, they were randomized to train in either the intact or blind hemifields, followed by 6 months of training in the blind hemifield during the chronic period. Pre-training, the GCL-IPL LI was small, ranging from -0.012 to 0.032, and it increased with time since stroke ($p=0.0120$, $R2=0.2649$). By the end of the study, blind-field direction discrimination thresholds had significantly improved (paired t-test, $p=0.0001$), irrespective of baseline GCL-IPL LI ($p=0.1805$, $R2=0.0923$). Similarly, training significantly reduced the size of the Humphrey-defined visual field deficit (paired t-test, $p=0.0002$), also irrespective of baseline GCL-IPL LI ($p=0.1410$, $R2=0.1105$). In summary, initiating visual training during the subacute period after occipital stroke provides significant perceptual benefit for all participants. In contrast to chronic patients, this improvement occurs irrespective of the small degree of baseline retrograde degeneration in the retina, which can be reliably detected using OCT.

Supported by NIH funding (R01 EY027314 to KRH, as well as T32 EY007125 and P30 EY001319 to the Center for Visual Science) and

by an unrestricted grant from the Research to Prevent Blindness (RPB) Foundation to the Flaum Eye Institute

23.306 METACOGNITION FOR ORIENTATION DISCRIMINATION: EFFECTS OF OCCIPITAL LOBE DAMAGE AND TRAINING

Matt Cavanaugh[1], Robbe Goris[2], Elisha Merriam[3], Krystel Huxlin[1]; [1]Flaum Eye Institute, University of Rochester, Rochester, NY, USA, [2]Center for Perceptual Systems, University of Texas at Austin, Austin, TX, USA, [3]Laboratory of Brain and Cognition, National Institute of Mental Health, NIH, Bethesda, MD, USA

Occipital strokes damage early visual areas, causing a large visual deficit. Psychophysical training can restore multiple perceptual abilities within this deficit, including direction and orientation discrimination. Here, we assessed metacognitive capabilities at blind-field locations in which orientation discrimination was either restored following training or remained impaired, using a gamified task to collect confidence reports simultaneous with orientation discrimination thresholds in seven patients (4 male; age: 37-71 years; time since stroke: 6.4-25.7 months). Five patients trained on fine direction discrimination (FDD) within their blind-field; two trained on fine orientation discrimination (FOD). Blind-field performance was initially at chance on both tasks. After training, blind-field FDD improved in all FDD-trained participants, with thresholds nearing intact-field levels (blind-field: 17.8±19.4; intact-field: 2.2±1.2; paired t-test: p=0.15). FOD thresholds likewise improved in both FOD-trained patients (blind-field: 3.4deg and 7.9deg; intact-field: 1.56 deg and 0.4 deg). Three FDD-trained participants failed to recover FOD thresholds at trained, blind-field locations, performing at chance on the gamified task (percent correct: 56.5±7.6%), and indicating high confidence on 6.0±9.2% of trials. All remaining patients (two FDD-trained and two FOD-trained) attained significantly improved orientation thresholds (blind-field: 5.0±4.5 deg), reporting high confidence on 49.3±10.8% of trials. Confidence and performance data were then fit with a computational model (CASANDRE, Boundy-Singer et al., 2023) to calculate meta-uncertainty, summarizing the quality of the brain's confidence computation (the higher meta-uncertainty, the noisier the confidence computation). All blind-field locations exhibited low meta-uncertainty (impaired: 0.013±0.005; recovered: 0.48±0.56) indistinguishable from the intact-field (0.87±0.75; intact vs recovered paired t-test p=0.52; intact vs impaired paired t-test p=0.13). Our results show discrimination confidence tracks lawfully with orientation discrimination performance within retrained blind-fields of patients with early visual cortex damage. We conclude that the metacognitive system can correctly process activity from visual circuitry spared by occipital strokes, which supports recovered orientation perception.

Plasticity and Learning: Models

Saturday, May 17, 8:30 am – 12:30 pm, Banyan Breezeway

23.307 RE-TUNING OF ORIENTATION-SELECTIVE RESPONSES IN NEURAL MASS ACTIVITY DURING AVERSIVE CONDITIONING: A MULTI-LABORATORY STUDY

Laura Ahumada[1], Judith Cediel[1], Andrew F. Farkas[1], Faith E. Gilbert[1], Hannah Engel[1], Arash Mirifar[1], Katherine McCain[1], Caitlin M. Traiser[1], Sarah Gardy[1], Christian Panitz[2], Andreas Keil[1]; [1]University of Florida, [2]University of Bremen

Visuocortical changes in orientation tuning consistent with experience-dependent plasticity have been studied in learning paradigms, such as in aversive conditioning. Aversive conditioning consists of pairing an initially neutral stimulus (CS+) with an unconditioned aversive stimulus (US). Here we used an aversive generalization learning paradigm to compare changes in visuocortical tuning in two laboratories—a standard EEG laboratory (standard) and a 3-Tesla fMRI scanner (EEG-fMRI). This paradigm employs a generalization gradient of stimuli to quantify changes in tuning from the CS+ to gradually different orientations (GSs) not paired with the US, which may take the shape of sharpening or generalization. Four Gabor patches were used, where one orientation represented the CS+ and the remaining three the GSs. All patches flickered at 15 Hz and were shown foveally. For the standard study, a white noise was used as the US; for the EEG-fMRI study, an electrical shock was used. The steady-state visual evoked potential (ssVEP) of 31 participants in the standard study and 24 participants in the EEG-fMRI study were separately analyzed using a mathematical model for quantifying tuning functions, the Ricker wavelet. This model identifies the tuning pattern that best characterizes the data. Results showed that in both laboratories, the visuocortical response to the CS+ increased during the acquisition and extinction phases. During acquisition, similar generalization tuning was observed in both laboratories. However, the exact tuning patterns differed during extinction, with sharpened tuning apparent in the standard but not in the EEG-fMRI study. Overall, visuocortical tuning was reliably measured in both settings, demonstrating that aversive learning in combination with Ricker wavelet fitting is suitable for characterizing tuning functions in human EEG/ssVEP recordings. Results also suggest that the laboratory setting, including the nature of the US, may influence the experience-dependent re-tuning of neural mass activity as measured by EEG signals.

23.309 THE EFFECTS OF LEARNING ON THE REPRESENTATIONAL GEOMETRY OF SKILLED CHESS PLAYERS

Andrea Ivan Costantino[1] (andreaivan.costantino@kuleuven.be), Esna Mualla Gunay[1], Emily Van Hove[1], Laura Van Hove[1], Felipe Fontana Vieira[1], Merim Bilalic[2], Hans Op de Beeck[1]; [1]KU Leuven, [2]Northumbria University

Exploring the neural and behavioral correlates of expertise offers a window to understand how cognitive and neural representations may change due to domain-specific training. Previous studies suggest that experts undergo a representational re-organization compared to non-experts, resulting in more linearly separable representations, particularly for task-relevant, high-level dimensions. Chess, with its rich history as a metaphor for human intelligence and strategic thought, serves as an ideal domain for probing expertise effects. Studies on chess expertise suggest expert players analyze board setups differently from novices, emphasizing piece relationships over visual traits. However, prior studies did not explore representational

structure and information processing changes in expertise, and in what brain areas these changes may occur. Our work bridges this gap by employing computational, behavioural, and neuroimaging methodologies to uncover representational changes in expert biological and artificial systems. In this study, 40 participants (20 chess experts and 20 novices) performed a chess-related task during fMRI scanning. We applied multivariate decoding techniques (MVPA, RSA) to analyze representational changes in human brain activity and artificial deep neural networks (DNNs). We hypothesized that (i) experts would exhibit distinct neural patterns associated with chess-relevant features in high-level brain regions, reflecting their advanced understanding, and (ii) there would be alignment in representational and behavioral patterns between expert biological and artificial systems. Our findings reveal a striking similarity in information processing between humans and DNNs, highlighting representational and behavioral alignment among expert systems. Additionally, experts—both biological and artificial—show a representational re-organization, resulting in more linearly separable representations for task-relevant high-level features at later processing stages.

23.310 HOW DOES CATEGORIZATION RESHAPE THE REPRESENTATIONAL GEOMETRY OF VISUAL CORTEX?

J. Brendan Ritchie[1], Maleah J. Carter[1], Hector Sanchez Melendez[1], Peter Molfese[1], Micah Holness[1], Vinai Roopchansingh[1], Chris I. Baker[1]; [1]National Institute of Mental Health

When we look for something of a particular category (e.g. a lost pair of keys), we differentially attend to some features of objects at the expense of others. What kind of influence does this selective attention have on the neural representation of stimulus features? One hypothesis is that categorization causes us to attentionally weight whole stimulus dimensions. Another is that categorization modulates perception resulting in greater local discriminability between stimuli spanning the categorical divide. Previous findings favoring one or other of these hypotheses have tended to focus on different brain regions (e.g. medial temporal lobe vs early visual cortex) without directly comparing the two alternatives or assessing how the effects of selective attention unfold over time. We conducted a simultaneous fMRI-EEG (3T and 256 channel) study in which participants (N = 15, two sessions each) carried out two categorization tasks of simple stimuli that varied in their orientation and spatial frequency. We then isolated a variety of regions of interest (ROIs), focusing on early and high-level visual areas in the ventral visual stream, as well as areas associated with conceptualization and memory, including the anterior and medial temporal lobes. Representational similarity analysis was used to compare the neural dissimilarities in these ROIs to the EEG signal, as well as a large range of categorization models of behavior that varied in the form of selective attention. We found that different attentional weighting models better captured the effect of categorization task with respect to behavior and spatiotemporal characteristics of the neural responses, both within and between participants.

23.311 DECODING RETINAL RESPONSES: A TRANSFORMER-BASED MODEL FOR VISUAL STIMULUS PREDICTION

Maria Jose Escobar[1], Francisco Miqueles[1], John Atkinson[2], Adrian G. Palacios[3]; [1]Universidad Técnica Federico Santa María, Department of Electronic Engineering, Valparaiso, Chile, [2]Universidad Adolfo Ibañez, Santiago, Chile, [3]Centro Interdisciplinario de Neurociencia de Valparaíso, Facultad de Ciencias, Universidad de Valparaíso, Valparaíso, Chile

The retina encodes visual information into spike trains using diverse functional cell types. While the encoding principles remain unclear, evidence highlights the complementary roles of individual and collective activity. This study explores a transformer-based neural model (POYO) to decode retinal structure and responses. POYO analyzes retinal networks by tokenizing neuronal activity into spikes and processing them through attention mechanisms, accounting for inter-animal variability. Using data from diurnal rodents (Octodon degus), the model was trained on five retinas over 400 epochs, achieving a loss of 0.019 and an R^2 of 0.921. Fine-tuning on a sixth, unseen retina led to rapid convergence in 30 epochs, reaching an R^2 of 0.998. To examine the role of different cell types, the model was trained on the entire population (496 units) and subsets: ON (46), OFF (256), and ON-OFF (194) cells. The R^2 scores were 0.997, 0.910, 0.882, and 0.937, respectively. While single-cell-type subsets achieved good performance, they could not fully recover the stimulus, which was only recovered using all retina cell types. These findings demonstrate that transformer-based models effectively predict visual stimuli from retinal responses and reveal the inner structure of retinal populations. Neuronal diversity plays a critical role in model convergence, as restricted inputs fail to capture the full stimulus. Moreover, trained models cannot generalize across tissues without fine-tuning. Rapid convergence on new retinas indicates that POYO captures generalizable retinal features, facilitating efficient adaptation to new datasets. These results underscore the importance of retinal heterogeneity in visual encoding and the potential of transformer models for advancing the understanding of sensory information processing.

ANID FONDECYT 1230170 and 1200880

23.312 DIFFERENT INTENSITIES OF TRANSCRANIAL MAGNETIC STIMULATION RESULT IN DISSOCIABLE PUPIL DILATIONS

Phivos Phylactou[1], David A. Seminowicz[1], Siobhan M. Schabrun[1]; [1]University of Western Ontario

The application of transcranial magnetic stimulation (TMS) most commonly relies on the estimation of the resting motor threshold (rMT), which serves as a proxy measure of cortical excitability. However, the rMT cannot always be estimated, as it relies on an intact pathway, travelling from the primary motor cortex (M1) to the periphery. In order to broaden the application of TMS, as well as to better understand the effects of TMS on the nervous system, there is an evident need for additional measures of cortical excitability. In this preliminary study, we provide evidence that pupil size dilation may serve as a potential measure of cortical excitability. In detail, 11 participants received 200 single pulses of either active or sham TMS at various intensities, while their pupil size was recorded. Bayesian evidence from a repeated measures ANOVA suggested that the maximum pupil size dilation as a response to TMS, was dissociable across intensities and between sham and active TMS (BF > 6). Post-hoc Bayesian paired t-tests provided further evidence of the different pupil responses between the various intensities and in sham versus active TMS. These findings

show promise for the introduction of a new potential estimate of cortical excitability through measuring pupil responses.

This work is supported by the Mary Elizabeth Horney Fellowship in Rehabilitation by the St. Joseph's Health Care Foundation

23.313 EFFECT OF SESSION DAY AND ATTENTIONAL STATE ON EXPERIENCE-DEPENDENT VISUAL CORTICAL NEUROPLASTICITY

Patricia Limon[1], Anthony Norcia[1], Ryan Ash[1]; [1]Stanford University

Passively viewing a 2 Hz contrast-reversing stimulus potentiates visual evoked responses in humans and rodents, enabling quantification of experience-dependent visual cortical neuroplasticity coined stimulus-specific response potentiation (SSRP). SSRP demonstrates input-specificity, retinotopic-specificity, and accumulates across days. While attention is known to modulate learning and plasticity, its impact on SSRP remains under-investigated. We hypothesized that directing spatial attention toward an SSRP-induction stimulus would enhance response potentiation within its retinotopic extent, while diverting attention would diminish it. Stimuli consisted of 2 semicircular hemifield checkerboards (10° eccentricity), each frequency-tagged to measure EEG steady-state visual-evoked potentials, using a 128-channel system. Participants (n=26) observed contrast-sweeps of the checkerboards before and after SSRP-induction, with sign-reversal rates of 6 Hz and 7.5 Hz (non-plasticity inducing), maintaining fixation via a letter-detection task. During SSRP-induction, the left checkerboard (100% contrast, 2 Hz reversal) served as the potentiation stimulus, while the right checkerboard (2% contrast, 3 Hz reversal) was the non-potentiated control. Via a contrast change detection task, participants directed peripheral attention to the potentiating stimulus (Potentiation-Attention Congruent) or the control (Potentiation-Attention Incongruent) on separate sessions. We observed a modest increase in response amplitude following SSRP that did not initially appear to depend on the attention condition. Interestingly, we observed that potentiation occurred exclusively on the second experiment day, suggesting that the SSRP effect accumulates across days. The enhancement was more pronounced at higher contrasts. Contrary to expectation, SSRP was not retinotopically specific, occurring in both hemifields. Second-day SSRP response enhancement occurred in both attention conditions; however, a significant three-way interaction (P=0.0005, VEP response x session day x attentional deployment) indicated that the sequence of attentional conditions (Day 1 congruent, Day 2 incongruent vs. Day 1 incongruent, Day 2 congruent) impacts SSRP development. Our findings demonstrate stimulus-induced neural plasticity in the visual cortex that dynamically evolves across sessions.

National Eye Institute

23.314 SHORT-TERM MONOCULAR DEPRIVATION CHANGES THALAMO-CORTICAL CONNECTIVITY MEASURED WITH ULTRA-HIGH FIELD FMRI DURING VISUAL STIMULATION

Miriam Acquafredda[1], Cecilia Barachini[1], Laura Biagi[2,3], Michela Tosetti[2,3], Maria Concetta Morrone[1], Paola Binda[1]; [1]University of Pisa, [2]IRCCS Stella Maris, Calambrone, Pisa, Italy., [3]IMAGO Center,Pisa, Italy

In adult humans, brief periods of monocular deprivation induce ocular dominance plasticity, transiently enhancing responses to stimuli in the deprived eye in the primary visual cortex V1 (Binda et al., 2018) and the ventral pulvinar (Kurzawski et al., 2022). Prior work (Acquafredda et al., VSS2023) suggests that these changes are accompanied by a functional reorganization of visual processing circuits, particularly cortico-pulvino-cortical loops, measured in resting-state. Here we investigated how thalamo-cortical connectivity is modulated by visual stimulation of either eye, before and after deprivation. We acquired ultra-high-field 7T fMRI EPI sequences from 22 normally sighted adults, pre- and post-2 hours of monocular deprivation. Participants were presented with monocular band-pass noise stimuli with five contrast levels. BOLD responses were analyzed for three regions of interest: V1, lateral geniculate nucleus, and pulvinar. Their effective connectivity was assessed by Dynamic Causal Modeling (Tapas toolbox, Frassle et al., 2017). Post-deprivation, V1 evoked responses increased for the deprived eye and decreased for the non-deprived eye, and the modulation was stronger for higher contrast stimuli, consistent with a response gain change. During deprived eye stimulation, pulvinar-to-V1 effective connectivity decreased post-deprivation (like in the resting state) while the V1-pulvinar connectivity remained constant. During stimulation of the non-deprived eye, the opposite connectivity changes were observed. These results suggest that monocular deprivation produces a disconnection between pulvinar and V1, probably decreasing V1 top-down modulation. The mirror-symmetric effect for the non-deprived eye recapitulates the symmetric modulation observed psychophysically. Overall, resting state and visual-evoked results suggest that brief periods of monocular deprivation produce changes in visual processing that extend beyond local V1 processes and may reflect a reorganization of top-down visual influence in V1.

Research funded by: grant PREDACTIVE (n. 101170249); grant THE (CUP I53C22000780001); grant PRIN 2022 (Project 'RIGHTSTRESS—Tuning arousal for optimal perception',n.2022CCPJ3J, CUP I53D23003960006); grant SMILY (by Italian Ministry of University and Research FARE-2).

Visual Memory: Neural mechanisms

Saturday, May 17, 8:30 am – 12:30 pm, Banyan Breezeway

23.315 ANTERIOR SHIFT FOR VISUAL RECALL VS. PERCEPTION IS SPECIFIC TO SCENES

Deepasri Prasad[1] (deepasri.prasad.gr@dartmouth.edu), Adam Steel[1,2], Caroline E. Robertson[1]; [1]Dartmouth College, [2]University of Illinois Urbana-Champaign

Although prior work proposed that visual recall reactivates visual category-selective brain regions, recent research shows that recalling familiar scenes recruits areas anterior and adjacent to visual scene-selective regions (Steel et al., 2021). Here we examined whether the anterior shift for visual recall vs. perception is unique to scenes, or if this occurs for other visual categories (e.g., faces, bodies, objects) using fMRI. We scanned participants (N=18) performing a visual recall and a visual perception paradigm. Before the visual recall task, participants provided a list of personally familiar faces, places, objects, and body parts (5 examples per category), which they vividly recalled during the scan. During the visual perception task, participants viewed dynamic videos from each category. For both paradigms, we contrasted activity during place-, face-, and body trials against object-activity to identify category-specific activation clusters and intersected this activity with locations of eight functional areas defined from publicly available atlases: parahippocampal (PPA), occipital (OPA) place areas, occipital (OFA) and fusiform face areas (FFA1, FFA2), and extrastriate body area subregions (ITG, LOS, MTG). Scene areas had a higher percentage of significant voxels during recall compared to body and face areas (all p-values ≤ 0.001); these recall activity clusters were larger and more contiguous in scene areas than in other areas. As predicted, all scene areas showed a significant anterior shift in peak activity for recall versus perception (all p-values < 0.001). Crucially, no anterior shift was observed for any face or body areas (all p-values ≥ 0.05). Taken together this suggests that scenes have distinct yet adjacent areas for perception and memory, but other categories do not. The unique anterior shift for scenes may reflect the need to represent contextual information outside the current field of view during navigation.

NSF Career Award 2144700

23.316 HOW DOES CHUNKING IMPROVE VISUAL WORKING MEMORY?

Yin-ting Lin[1] (lin.3913@osu.edu), Lisa M. Heisterberg[1], Mackenzie J. Siesel[1], Andrew B. Leber[1]; [1]The Ohio State University

Visual working memory (WM) is limited, but we can strategically use grouping cues to improve performance. Work has revealed better performance for items forming an illusory shape (Kanizsa illusion) versus randomly oriented items (Allon et al., 2019). However, the mechanism behind this benefit is still unclear. Some have proposed that Kanizsa illusions improve memory by reducing storage demands, but existing studies have found mixed evidence (Heisterberg, 2021; McCollough, 2011; see also Diaz et al., 2021). Notably, the studies that found reductions in storage used longer stimulus encoding intervals. We propose that participants need these long durations to integrate the illusory shape and thus reduce their memory load, and we test this in the present study. Participants performed an orientation change detection task with EEG recording. The memory array consisted of either one “pacman” stimulus (single), three randomly oriented stimuli (proximity), or three stimuli that form an illusory triangle (Kanizsa). Critically, we manipulated encoding time (short vs. long) between subjects. Results show a Kanizsa benefit in accuracy for both groups. Importantly, we used both the contralateral delay activity (CDA) and a multivariate classifier to examine the amount of information stored in WM. There was a greater CDA amplitude for proximity versus single condition for both groups, showing that CDA is sensitive to WM load. However, we did not see a statistically reliable difference between Kanizsa and proximity conditions for either group. The multivariate classifier trained on single and proximity conditions was able decode WM load reliably above chance. Interestingly, the multivariate analyses suggest that there is a reduction in WM load for the long group but not for the short group. Our findings raise the possibility that Kanizsa benefits do not always result from reduced storage demands. Rather, chunking may reduce WM load only when there is sufficient encoding time.

23.317 LOGARITHMICAL CHANGES IN NEURAL LATENCY WITH INCREASING LIST LENGTHS IN VISUAL MEMORY

Chong Zhao[1], Edward K. Vogel[1]; [1]University of Chicago

Remembering a list of 128 real-world objects is much more challenging than memorizing a list with just a few objects. Previous behavioral studies had suggested that visual long term recognition tasks produce response times that scale logarithmically with list length. Here, we investigated the neural mechanisms underlying this phenomenon. Sixty participants studied lists of real-world objects with lengths of 1, 2, 4, 16, 32, and 128, followed by a recognition memory test while electroencephalograms (EEG) were recorded. Event-related potential (ERP) analyses revealed that the amplitudes of the parietal and frontal old-new effects significantly decreased as list length increased, particularly for lists of 32 and 128 items. Additionally, the latencies of these effects shifted systematically: shorter lists elicited earlier latencies, while longer lists exhibited progressively delayed responses. Critically, the latency increases were well-modeled by a logarithmic function of list length. These findings demonstrate that memory load modulates both the strength and timing of recognition-related neural activity, aligning with logarithmic memory search processes reported in hybrid search paradigms.

This research was supported by the funding from National Institute of Mental Health (grant ROIMH087214); Office of Naval Research (grant N00014-12-1-0972).

23.318 NO SPONTANEOUS REPRESENTATIONS OF SPATIAL LOCATION IN FEATURE-SELECTIVE RETINOTOPIC CORTEX DURING WORKING MEMORY FOR COLOR AND MOTION

Daniel Thayer[1], Yanming Li[1], Thomas Sprague[1]; [1]University of California, Santa Barbara

Per the ‘sensory recruitment’ model of working memory (WM), frontal and parietal cortex engage sensory regions which have precise feature selectivity to maintain information in WM during delay periods (Curtis & D’Esposito, 2003; Postle, 2006). Since this model predicts that neural populations with specialized tuning should be recruited during WM maintenance, regions with selectivity for non-spatial features, such as color (hV4/VO1/VO2) and motion (TO1/TO2), should be recruited to maintain color and motion information, respectively, and the populations best suited for robust recruitment are those spatially aligned with the sample stimulus location. Some previous studies have shown that when remembering specific object features, the associated location is automatically encoded (Foster et al., 2017; Pratte & Tong, 2014), while others suggest that features are maintained in a spatially global manner (Ester et al., 2009). Here, we tested whether encoding features in WM results in recruitment and maintenance of activation of spatially tuned populations, or instead if feature representations are encoded globally. Participants viewed a colorful moving dot stimulus at a random location on each trial and were postcued to remember its color or motion. After a 12s delay, they adjusted the relevant feature of a probe stimulus, presented at the sample location, to match the sample. We used multivariate spatial inverted encoding models to quantify multivariate activation patterns. Strikingly, in retinotopic color- and motion-selective regions, reconstructed maps contained no representation of the sample stimulus location during the delay period.

In sharp contrast, parietal cortex had a robust delay-period representation of the sample stimulus location, even though location was irrelevant for the task. Together, these results indicate that space is not obligatorily encoded in feature-selective cortices. Rather, it appears that parietal cortex encodes the location of the sample stimulus, potentially instantiating feedback signals to feature-selective regions to bind features to a location.

Funding: Alfred P Sloan Research Fellowship, National Eye Institute R01-EY035300

23.319 SPATIAL WORKING MEMORY IMPAIRMENT PREDICTS THE SEVERITY OF SPATIAL NEGLECT SYMPTOMS OVER TIME.

Branden T. Otte[1] (otteb@mymacewan.ca), Stephen H. Scott[2], Sean P. Dukelow[3], Christopher L. Striemer[1,4]; [1]MacEwan University, Edmonton, Alberta, Canada, [2]Queen's University, Kingston, Ontario, Canada, [3]University of Calgary, Calgary, Alberta, Canada, [4]University of Alberta, Edmonton, Alberta, Canada

Damage to the right temporal-parietal cortex often leads to spatial neglect – a disorder in which patients are unable to attend to people or objects on their contralesional (i.e., left) side. Previous research suggests that patients with spatial neglect may have impairments in spatial working memory (SWM). This has led some to suggest that a SWM deficit is a core component of the neglect syndrome that may exacerbate the severity of neglect symptoms. To further examine this hypothesis, 28 right hemisphere stroke patients (n=5 with spatial neglect) completed the Behavioural Inattention Test (BIT) – a standardized assessment for spatial neglect – as well as Spatial Span, Trails-A, and Paired-Associates Learning tasks using the Kinarm exoskeleton at baseline (median=31 days), and at ~2 months (n=21; median 65 days), and ~3 months (n=8; median=97 days) post-stroke. Our results demonstrate that, at each time point, patients with spatial neglect performed more poorly than patients without neglect on the Spatial Span task that measures SWM, as well as Trails-A, which measures visual search and visuomotor processing speed. In addition, at baseline, poorer performance on Spatial Span and Trails-A tasks were significantly correlated with the severity of spatial neglect symptoms (as measured by the BIT) across the entire patient group (n=28). More importantly, performance on the Spatial Span and Trails-A tasks at baseline were significant predictors of neglect symptom severity across the entire patient group at both the 2-month and 3-month follow-ups. Overall, these results provide further evidence that spatial neglect is accompanied by deficits in SWM and demonstrate, for the first time, that SWM deficits are associated with the severity of neglect symptoms over time.

This work was supported through a Natural Sciences and Engineering Research Council of Canada Discovery Grant awarded to C.L.S., an Ontario Research Foundation grant to S.H.S., as well as Canadian Institutes of Health Research and Heart and Stroke Foundation of Canada grants to S.P.D.

23.320 THE MULTIPLE ROLES OF VISUAL CORTEX IN SINGLE-SHOT EPISODIC MEMORY: ENCODING, RECOGNITION, AND RECALL

Robert Woodry[1], Jonathan Winawer[1], Serra Favila[2]; [1]New York University, [2]Brown University

Sensory cortex exhibits activity patterns during long-term memory retrieval similar to those during encoding. This memory reactivation in primary visual cortex is especially robust for items that were highly practiced in an associative learning task requiring repeated, precise judgments (Woodry, Curtis, Winawer, 2024). Here we examined the role of visual cortex in a more natural, one-shot memory paradigm. Building on a pilot fMRI study (n=5) by Favila and Winawer (VSS 2020), we increased the sample size (n=13) and conducted new analyses of an experiment in which participants encountered 480 object images across three interleaved tasks: encoding, recognition, and spatial recall. During encoding scans, peripheral objects appeared at one of four polar angles while subjects maintained central fixation and judged whether the object was bigger/smaller than a shoebox. During recognition scans, either the same objects or lures were presented centrally while subjects made old/new judgments. During spatial recall scans, subjects viewed old objects centrally and indicated the polar angle of the object at encoding. We observed spatially tuned responses in visual cortex, including V1, that were relevant for encoding success, recognition, and recall. First, at encoding, V1 responses were larger and more sharply tuned for objects whose location was successfully recalled later. Second, spatially tuned responses were evoked in visual cortex during successful recognition, even though the items had only been viewed once and location wasn't probed. Moreover, spatial tuning in visual cortex during recognition was poor for objects later forgotten (incorrect spatial judgment), indicating that the spatially tuned recognition responses were behaviorally relevant. Third, when aligning to the reported location during recall, we observed spatially tuned responses for both remembered and forgotten objects. These findings show how the properties of visual cortex activation track memory behavior during encoding, recognition, and recall, even for single shot memory.

23.321 TRANSIENT GLUTAMATERGIC BOOSTS IN EARLY VISUAL AREAS ACCOMPANY SUCCESSFUL TARGET RECOGNITION WITH INCREASING WORKING MEMORY LOAD

Zhiyan Wang[1], Antonia Wittman[1], Sinah Wiborg[1], Markus Becker[1], Mark W. Greenlee[1], Sebastian M. Frank[1]; [1]Institute for Psychology, University of Regensburg

Functional magnetic resonance imaging (fMRI) studies have provided compelling evidence that the content of working memory can be successfully decoded in early visual areas (EVA). However, it remains elusive whether fMRI-measured hemodynamic response patterns following neural activity in EVA support the successful recognition of targeted information, during which a rapid comparison of current sensory input with representations held in working memory is necessary. Using event-related functional proton magnetic resonance spectroscopy (1H-fMRS), it is possible to measure transient changes in excitatory glutamatergic processing in EVA. Therefore, to characterize the transient processes underlying the recognition of targets, we measured participants' (n = 30) glutamatergic responses in EVA while they performed a working memory task with low (1-back) and high (2-back) working memory load. Participants were presented a series of letters in the center of the screen and responded by button press when the letter in the current trial matched the previous trial (low

load, 1-back) or when it matched the letter two trials before (high load, 2-back). Low and high load conditions were presented in blocks using a pseudorandom interleaved order. Each block contained 16 trials, where on each trial a letter was presented for 700 ms, following by an intertrial interval jittered between 1.05-1.25 s. On average, 4 targets were presented during each block. The glutamate response was calculated for the time periods corresponding to target and non-target presentations. We found a significantly greater glutamate-response during target compared with non-target presentations. Furthermore, as mental loads increased from 1-back to 2-back, the target-related glutamate-response also increased. No such effects in glutamate response were found in a separate control experiment with 1H-fMRS in dorsolateral prefrontal cortex. Our study suggests that a transient boost of excitatory neuronal processing in EVA, modulated by working memory load, may underly the dynamic recall of target information.

Alexander von Humboldt Foundation. Julitta und Richard Müller Stiftung; Deutsche Forschungsgemeinschaft (DFG): Emmy Noether Grant (Project Number 491290285)

23.322 VISUAL-MEMORY-GUIDED DRAWING: INVESTIGATING CEREBELLAR COGNITIVE ARCHITECTURE AND CEREBRO-CEREBELLAR INTERACTIONS

Lora Likova[1], Zhangiyi Zhou[1], Christopher Tyler[1], Kristyo Mineff[1]; [1]Smith-Kettlewell Eye Research Institute

Traditionally associated with motor control and coordination, the cerebellum is increasingly recognized for its contributions to diverse cognitive processes. However, its specific cognitive roles and interactions with the cerebrum remain poorly understood and underexplored in the context of vision and cognition, highlighting the need for further investigation. Memory-guided drawing—a complex task engaging the full "perception-cognition-action loop"—provides a powerful framework for probing these mechanisms. Methods: Participants underwent whole-brain fMRI (Prisma 3T scanner) while drawing from visual-memory and performing additional tasks to isolate its key components: 1) Drawing from visual memory: Creating line drawings of complex spatial structures based on immediate visual memory; 2) Visual-memory recall: Visualizing images from immediate visual memory without drawing, to isolate memory and visualization processes; 3) Non-visual scribbling: Scribbling freehand blindfolded, to isolate visual motor control for drawing while excluding memory, cognition, and visual mechanisms; 4) Drawing from haptic memory: Drawing blindfolded from memory, to isolate haptic memory and coordination while excluding visual mechanisms. Granger Causal Connectivity Analysis (GCA) was used to investigate intracerebellar networks and their interactions with large-scale cortical networks. Results and Conclusions: The findings illuminate the cerebellum’s role in higher-order cognitive functions and its intricate functional architecture. Distinct cerebellar and cortical networks were revealed across tasks. Visual-memory-guided drawing elicited the most widespread and strong brain activation; notably, all cerebellar lobules—including non-motor regions—showed significant activity. Comparative analyses with the secondary component tasks provided novel insights into the cerebellum's perceptual-cognitive organization. GCA revealed causally directed cerebro-cerebellar interactions, involving both facilitative and suppressive feedback mechanisms. Particularly noteworthy were interactions with the Default Mode Network and other large-scale cortical networks. These results enhance our understanding of the cerebellum's integrative roles in perception, memory, cognition, and action, elucidating how its contribution extends far beyond motor coordination to encompass complex perceptual-cognitive operations.

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

23.323 DIRECT COMPETITION BETWEEN HIPPOCAMPAL PATTERN SEPARATION AND SHORT-TERM MEMORY PRECISION

Inik Kim[1], Bo-Yeong Won[2], Ana Acebo[1], Weiwei Zhang[1]; [1]University of California, Riverside, [2]California State University, Chico

The ability to remember detailed information is critical not only for daily functioning but also for understanding memory declines with aging and neuropsychiatric disorders, such as schizotypy. Given the hippocampal pattern separation mechanism for Long-Term Memory (LTM) precision, the interactions between Short-Term Memory (STM) and LTM by Cowan’s embedded process model, and the growing literature on the hippocampus’ roles in STM, we hypothesized that hippocampal pattern separation as a neurocomputational mechanism for mnemonic precision in visual STM would limit the task performance when maintaining precise visual information concurrently in LTM and STM. Specifically, the high task demand on pattern separation in LTM would degrade mnemonic precision for a concurrent visual STM task. To test this prediction, we embedded a Change Detection (CD) task into the Mnemonic Similarity Task (MST) test phase. Participants were required to perform the MST test phase during the delay interval of the CD with line orientations. Experimental manipulation (high-capacity load versus high-precision load, Experiment 1) and ROC method (Experiment 2) were used to tease apart WM quantity and precision. The results from both experiments indicated that memory precision was selectively impaired in the MST lure trials that required the ability to remember the rich details of the item, whereas memory quantity remained intact. Experiment 1 specifically demonstrated that CD accuracy was notably impaired during the MST lure trials under the high-precision load while remaining intact for other MST trial types and under the high-capacity load of the CD task. Experiment 2 revealed that the ROC-based computational estimate of mnemonic precision, but not capacity, was impaired during the MST lure trials. These findings support the hippocampal pattern separation hypothesis by demonstrating that hippocampal pattern separation is pivotal for visual STM precision.

Visual Memory: Encoding and retrieval

Saturday, May 17, 8:30 am – 12:30 pm, Banyan Breezeway

23.324 COMPLEX EFFECTS OF DIVIDED ATTENTION ON SACCADIC EXPLORATION DURING MEMORY ENCODING

Chloe Kindell[1] (ckinde2@lsu.edu), Heather Lucas[1]; [1]Louisiana State University

Humans use eye movements to encode information into memory, and memory difficulties are often linked to altered viewing behaviors. For instance, hippocampal amnesia has been associated with reduced or less effective viewing patterns during memory tasks, indicative of altered communication between memory and vision. This study examined whether taxing attentional resources would produce similar effects in healthy young adults. Eye movements were recorded while participants (n=44) viewed eighty displays of six abstract line drawings, each followed by a spatial reconstruction test. Attentional resources during the study phases were manipulated within-subjects by varying the difficulty of a concurrent auditory task in a block design. Block order was counterbalanced across participants, with half completing the "easy" blocks first and half starting with the "hard" blocks. Overall, greater secondary task difficulty was associated with: 1) worse memory performance, and 2) constrained visual exploration, as evidenced by fewer and longer fixations during study. Unexpectedly, these effects interacted with the order in which participants were exposed to the two levels of secondary task difficulty. Participants who received the hard blocks first showed improved memory performance when tasks became easier, despite no changes in viewing behavior. Conversely, participants who received the easy blocks first exhibited stable memory performance across task difficulties but made fewer and longer fixations when the task became harder. These findings suggest that the relationship between attentional resource availability and memory-related viewing behaviors can be influenced by task experience. Strategically constraining one's viewing may help to preserve memory performance when attention is suddenly taxed mid-task, while a mid-task increase in attentional resources may enhance encoding effectiveness without altering viewing patterns. Follow-up studies are in progress to clarify the mechanisms driving these patterns.

Louisiana Board of Regents

23.325 CONTRIBUTIONS OF FAMILIARITY AND RECALL TO RECOGNIZING NON-SEMANTIC VISUAL STIMULI

Lotta Pesonen[1] (pesonen_lotta@phd.ceu.edu), Jozsef Fiser[1], Máté Lengyel[1,2]; [1]Central European University, [2]University of Cambridge

Understanding the relationship between statistical learning and recognition memory is essential for explaining how environmental input is encoded and recalled. While statistical learning has been extensively studied in visual paradigms, recognition memory research has typically involved semantically rich stimuli such as words and images of natural scenes. The present study bridges these fields by examining how previous findings in recognition memory generalise to paradigms where simple, non-semantic visual stimuli is used. In seven experiments, participants completed recognition tasks with shape stimuli. In the item recognition task (n = 150), participants distinguished between previously viewed shapes and novel lures. In the association recognition task (n = 200), participants judged whether pairs of shapes were previously viewed or recombined from familiar shapes. Unlike traditional paradigms, our study used a small stimulus set and a single familiarization-test block. The stimulus presentation times during familiarization and inventory sizes were varied between participants. Prior research suggests item recognition depends on both familiarity and recollection, while association judgments rely mainly on recollection. Our findings confirm this for item recognition but reveal that associative memory judgments are primarily driven by familiarity, contrary to prior studies. This challenges earlier models, indicating that recognition processes are influenced not only by task type but also by stimulus characteristics and experimental design. Moreover, reducing the study list length in the association task further amplified familiarity's role, contradicting previous claims linking list length solely to recollection. These results underscore the impact of contextual factors on recognition memory and suggest a more prominent role for familiarity in tasks involving non-semantic stimuli.

23.326 EFFECTS OF SALIENCE ON VISUAL WORKING MEMORY DISAPPEARED! CONTEXT DURING RETRIEVAL MATTERS.

Martin Constant[1] (martin.constant@unige.ch), Dirk Kerzel[1]; [1]University of Geneva, FPSE

We recently demonstrated that salience can have a massive impact on visual working memory in a delayed continuous recall task (Constant & Liesefeld, 2021; https://doi.org/gjk9jh). We also showed that this effect is hard to erase or overrule with pre-/retro-cues or with manipulations of relevance, even at long encoding times (Constant & Liesefeld, 2023; https://doi.org/gr6xzr; Constant & Kerzel, 2024; https://doi.org/mvnk). Here, participants memorized three tilted bars (12°, 28° and 45°) presented for 500 ms among 33 vertical bars. After a delay of 1000 ms, a response display appeared and memory for one of the tilted bars was probed. In previous studies, participants adjusted the color of the probed bar in a grayscale version of the original display. Here, the probed bar was shown in its original tilt in the center, which had unexpected effects on memory performance. In the first experiment, participants first recalled the original position of the probed bar and then recalled its color on a colorwheel. As expected, the most salient bar with 45° tilt was remembered more accurately than the less salient ones, but in contrast to previous studies, there was no difference between the two less salient bars with tilts of 28° and 12°. The same pattern of results was observed when participants judged only color, ruling out dual-task load as an explanation. Finally, restoring a context of vertical bars around the tilted bar in the center of the response display did not change the results. These results may suggest that the most salient bar is maintained in an active state in VWM, whereas the less salient bars are stored in a silent state which is less resilient to sudden death or swapping.

23.327 ERP MARKERS OF HYBRID VISUAL AND MEMORY SEARCH FOR FAMILIAR AND NOVEL OBJECTS

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

Previous research investigated the neural dynamics of memory search using electrophysiological markers of old-new recognition and working-memory load (Rugg, 2007; Williams et al., 2024). All those studies involved situations when only one test item, either old or new, was present per trial. However, in many cases, observers deal with several simultaneously presented items and have to decide which are familiar or novel; this is often termed a hybrid visual and memory search or forced-choice memory test. Here, we present a method to reveal ERP markers of hybrid search. In each block of our experiment, participants studied 20 visual objects. At the test, participants were bilaterally shown two objects, old and new. We manipulated search instructions across blocks, asking participants to localize either the old or the new object. The bilateral presentation allowed us to analyze ERP signals from the old and new items as the differences between lateralized ERP's. We found two lateralized ERP effects. First, we found a sustained negative activity in the posterior electrodes contralateral to the location of the old item regardless of search instruction. It resembles an established ERP component known as contralateral delay activity (CDA) (Vogel & Machizawa, 2004) associated with encoding and active storage of information in working memory. It may indicate that familiar items get prioritized access into working memory. We label this effect r-CDA (retrieval-CDA). Second, we found a frontal negativity contralateral to the new item when participants looked for those items. The onset latency of this frontal effect was later than the r-CDA, which can suggest a specific novelty-related signal following the initial familiarity check developing only if novelty is relevant. The dissociation between these two ERP markers can indicate an essential separation between two components or pathways used by recognition memory to evaluate the familiarity and novelty of multiple objects.

Project funding: ONR N00014-22-1-2123

23.328 INVESTIGATING THE EVOLUTION OF THE SOCIAL ENCODING BENEFIT ACROSS HEALTHY HUMAN DEVELOPMENT

Salma Ben Messaoud[1] (sbenm025@uottawa.ca), Gray Koudinova[1], Cordula Hunt[2], Bozana Meinhardt-Injac[3], Boutet Isabelle[1]; [1]University of Ottawa, Ontario, Canada, [2]Johannes Gutenberg University of Mainz, Rhineland Palatinate, Germany, [3]Catholic University of Applied Social Sciences, Berlin, Germin

Introduction: Facial recognition is crucial to the unfolding of our social lives. In younger adults (YAs), two decades of research provide support for a 'social encoding benefit', wherein individuals are better at recognizing new faces learnt using social encoding strategies as opposed to perceptual ones (e.g., Mueller et al., 1978; Winograd, 1981; Schwartz & Yovel, 2018). Healthy ageing is associated with a generalized decline in facial recognition abilities (e.g., Boutet et al., 2015). To our knowledge, only one study has sought out to investigate how different encoding strategies could mitigate this decline (Winograd, 1978). However, this investigation presents various methodological limitations, including usage of the same images during encoding and testing. The present study aimed to investigate the effectiveness of social and perceptual encoding strategies in older adults, whilst varying the presentation format of faces in between encoding and testing to account for such real-life variations (Burton, 2015). Methods: Younger (YAs; n=34, aged 18-35) and older adults (OAs; n=16, aged 65+) have been tested so far. In the encoding phase, participants were presented with faces of different individuals and asked to remember them. Faces were shown under three encoding conditions: social, perceptual and control. In the testing phase, participants were shown pictures of the same individuals under a different illumination or viewpoint, as well as faces of new distractor individuals. Participants indicated if they recognized the shown individual. Results: Preliminary results suggest that irrespective of participant age, performance was better in the social encoding condition than control and perceptual conditions. Additionally, OAs displayed a tendency for lower facial recognition accuracy overall. Conclusion: This study suggests that OA could benefit from using a social encoding strategy to facilitate recognition of new people they meet. We are currently testing additional OA participants and exploring the effectiveness of other encoding strategies.

23.329 IS AFFECTIVE INFORMATION BOUND TO VISUAL OBJECTS OR ACTIONS? THE INTEGRATION OF POSITIVE AND NEGATIVE FEATURES IN STIMULUS-RESPONSE BINDING

Sihan He[1] (sihan.he@mail.utoronto.ca), Jay Pratt[1]; [1]University of Toronto

As humans, we frequently evaluate the affective value (positive/negative) of things we see daily. Although such affective features are psychological rather than physical features of objects, abundant research has demonstrated their critical roles in visual processing (e.g., enhanced perception, biased attention, more durable memory) and action choices (e.g., approach/avoidance). Intriguingly, our brain integrates visual stimuli and corresponding responses into an episodic representation, an "event file," that can be stored and retrieved holistically. Thus, an important question is: Whether affective features, which can be relevant to both visual stimuli and actions, are bound to this event file and, if so, to which components they are bound (stimuli, action, or both)? To investigate this question, we used evaluative conditioning, validated with a priming task, to associate valences to simple visual geometric shapes with different colors. These affective stimuli were then used in a partial repetition task in which participants responded to two affective stimuli sequentially, with responses and stimuli features (color, valence) either repeated or switched. Reaction time differences between the partial repetition conditions where only one component was switched or repeated and the baseline condition where all components were repeated (the partial repetition cost) was the test of integration of that feature into the event file. Our findings revealed a significant cost when only valence or response was repeated. Surprisingly, no cost was found when both valence and response were repeated, indicating no cost from a switch of color. These findings suggest that the affective features of stimuli can be integrated into event files and might be bound to responses rather than stimuli. This study provides new insights into the integration of affective features with perceptual features and actions.

23.330 IS REPULSIVE SERIAL BIAS IN VISUAL PERCEPTION DRIVEN BY LOW-LEVEL ADAPTATION?

Scott Janetsky[1], Kuo-Wei Chen[1], Kenzy Moustafa[1], Gi-Yeul Bae[1]; [1]Arizona State University

Visually-guided behaviors can be systematically influenced by recent perceptual history. For example, reports in a location working memory task exhibit repulsive biases away from task-irrelevant prior stimuli. The prevailing theories suggest that repulsive serial bias is best explained by low-level adaptation that directly alters the perceptual representation during the stimulus encoding. However, empirical evidence for this low-level adaptation account is limited. Here, we sought to find empirical evidence for the low-level adaptation account via a correlation approach. In the experiment, 40 participants performed both a location estimation task with an inducer designed to measure adaptation-induced repulsive bias (i.e., visual aftereffect) and a typical location delayed estimation task designed to measure repulsive serial bias. In the analysis, we correlated the adaptation-induced repulsive bias with the repulsive serial bias. We hypothesized that if the repulsive serial bias is driven by low-level adaptation, then it should be positively correlated with adaptation-induced repulsive bias. We confirmed that both tasks exhibited repulsive biases. However, the two biases were not correlated (BF for the null = 2.8). In subsequent analyses, we investigated whether the biases were associated with post-perceptual decision processes by examining response time as a function of the stimulus difference between the current and the prior stimuli, and the difference between the target and the inducer. We found that response time was slower when the current stimulus was more similar to the prior stimulus in repulsive serial bias, indicating the decision was more difficult in such trials. However, this was not the case in adaptation-induced repulsive bias, indicating that the inducer directly altered the perception of the target stimulus. Together, these results suggest that repulsive serial bias is driven by decision-related processes that integrate the prior stimulus rather than by low-level adaptation that directly impacts the representation of the stimulus itself.

23.331 MEMORY SPECIFICITY THROUGH VISUAL PRODUCTION: MULTIMODAL RECOGNITION AND SOURCE MEMORY MISATTRIBUTIONS

Hala Rahman[1] (13har1@queensu.ca), Keanna Rowchan[1], Jeffrey Wammes[1,2]; [1]Queen's University, Department of Psychology, [2]Queen's University, Centre for Neuroscience Studies

Successfully retrieving information from memory is critical to our ability to learn from experience to better inform our decision-making and interactions with the environment. As such, various strategies have been developed to aid memory encoding, from verbal mnemonics to visualization. Among these, drawing, or visual production, has emerged as a powerful tool. Beyond some traditional approaches, visual production engages several forms of cognition (visual, motoric, and elaborative processes) simultaneously. Therefore, drawing provides an excellent test case for studying how these multisensory components interact to improve memory. Here, we tested the relative mnemonic impact of the various cognitive features that are involved in the act of drawing. Participants (N = 60) completed four encoding tasks - Drawing, No Ink drawing, Tracing, and Visualizing - designed to differentially engage the multi-sensory 'components' thought to underlie drawing. We found that engaging all three components during Drawing was associated with the best memory, while engaging only one component during Visualizing was the worst. This is consistent with prior work, and we demonstrated it in both classic old/new recognition (Experiment 1), and source memory (Experiment 2). Despite lacking the visual component, recognition memory for No Ink drawing items was as good as Drawing, but interestingly, resulted in the most source memory confusion: No Ink items were frequently mistaken as Drawn items. We replicated this finding in Experiment 3, which only compared drawing with and without ink. Across these three experiments, lack of visual feedback did not influence recognition memory overall, but clearly undermined the precision of the source memory. Ongoing fMRI work is investigating the underlying neural mechanisms supporting memory in each of these tasks to clarify how shared visuomotor processes might lead to confusable representations. Together, our results extend our current understanding of the cognitive processes and visuomotor interactions underlying successful retrieval from memory.

NSERC

23.332 REPRESENTATION OF THE COLOR-SHAPE CONTINGENCY OF OBJECT CONCEPTS IN MACAQUE VISUAL CORTEX USING FMRI

Helen E. Feibes[1], Spencer R. Loggia[1,2], Karthik Kasi[1], Bevil R. Conway[1]; [1]Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, [2]Department of Neuroscience, Brown University

Object concepts defined by color-plus-shape are important tools of cognition. For example, they enable inferences about missing or mismatching perceptual data, providing a mental picture that guides visual search. The brain regions and associated processes responsible for storing object concepts remain poorly understood. One hurdle has been that concepts and patterns of brain activity vary among individuals; moreover, the subjective value associated with different concepts varies. Here we overcome these challenges by training two macaque monkeys to learn about 12 objects, each defined by a unique color and shape. After four years of experience with these objects, we pairwise swapped their colors, generating 12 'incongruent' objects. We then took fMRI scans while the monkeys viewed blocks of congruent objects (the learned color-shape pairs) and incongruent objects. All colors and shapes were the same between incongruent and congruent objects, as was the reward associated with them. This enabled us to isolate brain regions that are sensitive to the specific conjunctions, providing a way of determining where in the brain color-shape knowledge is stored. Prior work has generated two main hypotheses: that knowledge is stored in sensory processing regions, or it represented abstractly elsewhere. Incongruent objects preferentially activated large parts of V2, V3, V4, and posterior-central IT (average selectivity .046; 95% CI (.036, .056)), and trended upwards along the posterior-to-anterior axis, while a consistent bias was absent in temporal pole and prefrontal cortex. A subset of color-biased and object-biased regions defined with an independent fMRI localizer experiment were selective for incongruent objects. Within these regions, voxels most selective for colors or objects were largely separate from voxels most selective for incongruency, (median overlap 20% with color-biased regions, 38% with object-biased regions). These results show that extrastriate retinotopic and posterior IT cortex are important in the formation of and access of color-shape conjunctions.

23.333 RETHINKING FORGETTING: THE INFLUENCE OF OVERT ATTENTION ON MEMORY TRACE DECAY

Marika Mauti[1,2] (marika.mauti@uniroma1.it), Moreno I. Coco[1,2]; [1]Sapienza University of Rome, Rome, Italy, [2]IRCCS Fondazione Santa Lucia, Rome, Italy

Our memory processes undergo a continuous change, shaped by a dynamic interplay between the formation and decay of episodic representations. While memory formation has been extensively investigated, the forgetting mechanisms underlying memory degradation remain elusive. More importantly, forgetting has always been derived as changes in the rate of successfully recalled memory over time. However, as memory processes are intimately related to attentional mechanisms, it is conceivable to derive an explicit understanding of forgetting processes directly from them. Unlike most studies on forgetting, we employed naturalistic scenes rather than letters or isolated objects, which provide rich and meaningful contexts,

returning a more ecologically valid explanation of its mechanisms. Participants (N = 15) studied 132 scenes encompassing diverse indoor and outdoor settings in preparation for three recognition sessions (30 minutes, 4 hours, and 8 hours) while being eye-tracked. The spatial distribution of attention during the study phase, measured as root-mean-square distances weighted by fixation duration, positively predicted memory accuracy independently of recognition sessions (i.e., no significant interaction). While the forgetting rate tends to decrease across recognition sessions, we observe it not to be perfectly linear. Specifically, we observed a rapid decline in memory accuracy between the initial two intervals (30 minutes and 4 hours), followed by a marked stabilization and even a slight improvement between the second and third intervals (4 and 8 hours). This pattern challenges the traditional linear forgetting curve and suggests a more complex decay process. Overall, these findings provide novel insights into the dynamic nature of forgetting and the role attention plays in preventing memory decay.

This work has been supported by NextGeneration EU under Piano di Ripresa e di Resilienza(PRIN 2022) with grant agreement (2022APAFFN; CUP B53D23014480001) awarded to MIC

23.334 RETRO-CUE REDUCES THE CATEGORICAL DISTORTION OF VISUAL SHORT-TERM MEMORY

Tianye Ma[1] (tma039@ucr.edu), Weiwei Zhang[1]; [1]University of California, Riverside

Visual short-term memory (VSTM) within the internal focus of attention (FoA) can be better recalled. The FoA of VSTM can be manipulated by cuing specific memory items after the offset of the memory stimuli, i.e., the retro-cue. However, the mechanisms for the often-observed retro-cue benefits are under heated debate. The current study aims to investigate whether the mitigation of the categorical bias of VSTM (e.g., shifts towards/away from the prototypical colors) can be a potential mechanism for the retro-cue benefit. Specifically, we tested if the memory item within the FoA shows a weaker categorical bias. In this study, the prototypical colors are defined by an independent color-naming task. We then measured the weight of categorical versus continuous encoding by fitting a mixture model to the responses from a delayed estimation task with colors. The to-be-recalled color has a 70% chance to be retrospectively cued after the offset of the memory items. Participants showed a smaller weight of categorical memory for the cued colors than the other colors. To investigate the temporal dynamics of categorical memory retrieval, we modeled continuous mouse cursor trajectories using a method previously developed by Ma and Zhang (2024, Journal of Vision). Our analysis revealed that the model-estimated weight for categorical memory remained consistently low throughout the retrieval process. This finding contrasts sharply with the phasic ramping observed in categorical retrieval in prior studies that did not employ retro-cue manipulation. These results suggest that visual short-term memory (VSTM) within the focus of attention (FoA) can be effectively insulated from biases introduced by prototypical feature categories.

23.335 THE INFLUENCE OF MOTOR REINSTATEMENT AND DRAWING QUALITY ON REMEMBERING

Tasha Ignatius[1], Gerome Manson[1], J. Randy Flanagan[1], Jeffrey D. Wammes[1]; [1]Queen's University, Kingston, CA

Creating drawings of information can provide elaborative, pictorial, and motor cues that facilitate later retrieval from memory, but the contribution of motor information remains unclear. To test this, we had participants encode words via drawing using a robotic manipulandum. 10 mins (E1) or 1-2 days (E2) later, they completed a visual recognition task while the robotic manipulandum guided their arm through predetermined motor paths. Unbeknownst to participants, the paths were either congruent (motor reinstatement) or incongruent (interference) with their drawing of the current target word. Response time was consistently fastest with motor reinstatement, indicating that reactivating an encoded motor path made visual recognition more efficient regardless of the delay between encoding and reinstatement. However, while E1 revealed that passive reinstatement improved recognition accuracy, this benefit disappeared with the longer delay in E2. Drawings were submitted to a pretrained neural network (NN), and E1 revealed that higher-quality drawings (i.e. those more easily identified by the NN) were better recognized, regardless of reinstatement condition, but again, this pattern disappeared with the longer delay. The collective findings demonstrate that passive motor reinstatement reliably increases the speed of memory retrieval, but that any influence on accuracy is transient. Using principal components analysis (PCA) we identified two components in NN features that were associated with better memory, and these were consistent across experiments, indicating that there were predictable features that lead to more memorable drawings. Ongoing work will determine whether changes in the level of motor engagement, quality and timing of reinstatement change the observed benefit to memory. Together, these results emphasize the influence of the quality and type of encoded features as well as the timing of reinstatement in influencing memory performance.

The research was undertaken thanks in part to funding from the Connected Minds Program, supported by Canada First Research Excellence Fund, Grant #CFREF-2022-00010

23.336 THE NOT-SO-DARK SIDE OF EXPERTISE: SOURCE MEMORY EFFECTS IN EXPERT MEMORY

Benjamin Johnson[1], Dyllan Simpson[1], Timothy Brady[1]; [1]University of California San Diego

Recognition memory is typically enhanced by expertise. Chess masters better remember chess positions and radiologists better remember mammograms. However, in some contexts, expertise appears to impair memory through increased source confusion. For instance, when asked to recall a subset of NFL teams, football experts are more likely than novices to falsely "remember" unstudied teams (Castel et al., 2007). Does this reflect a genuine cost of expertise? In our first experiment, we tested recognition memory for sports logos in NFL and NHL fans, comparing experts remembering logos from their domain of expertise versus outside their expertise. We replicated previous results showing increased false alarm rates on logos within participants' domain of expertise. However, in strong contrast to previous findings, we show this does not reflect a genuine memory impairment. ROC analyses revealed only a criterion shift: experts showed the same memory strength for items in and outside their domain but struggled to distinguish whether their sense of familiarity came from the experiment prior exposure. This led to more 'old' responses for both studied and unstudied logos in their domain of expertise. Our second experiment used novel game photos instead of logos. Here, experts demonstrated better recognition memory for photos from their domain of expertise. Unlike logos, these game photos were less familiar, reducing source memory interference while still allowing experts to leverage their domain knowledge. These findings reveal that expertise effects on recognition memory depend critically on stimulus familiarity. While expertise provides no advantage for highly familiar stimuli due to source confusion, it enhances memory for novel domain-relevant materials where prior exposure cannot interfere. Rather than demonstrating a "dark side" of expertise, these results show that controlling for stimulus familiarity eliminates apparent expertise costs, allowing experts to demonstrate their enhanced encoding and retrieval capabilities.

Visual Memory: Objects and features

SATURDAY, MAY 17, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

23.337 COMPARISON OF SHORT- AND LONG-TERM RECALL AND RECOGNITION MEMORY FOR HAPTICALLY AND VISUALLY EXPLORED OBJECTS

Savannah Waymer[1], Susan Wardle[1], Elanor Chang[1,2], Chris Baker[1]; [1]National Institute of Mental Health, [2]Cornell University

We spend much of our everyday life physically interacting with real objects, yet most research on visual memory is conducted on words and images. In this study we examined short- and long- term memory for everyday objects as a function of the type of real-world experience (haptic & visual, or visual only). Participants explored 24 real objects (e.g. sneaker, clock, wooden birdhouse) for 30 seconds each, in random order. Participants in the visual group viewed the objects on a turntable and controlled the rotation speed and direction; participants in the haptic group held and freely explored the objects. Following the exploration phase, participants completed a distractor task (digit span). Participants were videorecorded during both exploration and memory phases. The first memory task was free recall of the explored objects. Next, participants were given verbal prompts of the object categories (e.g. "Do you remember any jewelry?") to help remember any additional objects; two prompts were foil categories. Lastly, participants performed a computer-based recognition task. The task consisted of photographs of the explored objects in 3 different orientations (72 images) and similar unexplored objects (foils) in 3 orientations (72 images). Participants returned 2 months (+/- 2 weeks) later to repeat the memory tasks. During immediate recall, participants recalled the objects with varying amounts of visual detail, with an advantage for the haptic group. Short term recognition memory performance was higher than recall for both groups, with a modest decline at the 2-month follow-up. Surprisingly, during the delayed 2-month recall, participants in both groups generated many false memories of objects that were not presented. In conclusion, we find that memory for real-world objects is modulated by the type of experience, and that real-world experience with an object does not impact the generation of false memories.

This research was supported (in part) by the Intramural Research Program of the NIMH (ZIAMH002909).

23.338 FORAGING FOR FRUIT: LEARNING OF COLOR SHAPE CONCEPTS IN MACAQUE MONKEYS

Spencer Loggia[1,2], James Cavanaugh[1], Bevil Conway[1,3]; [1]National Eye Institute, [2]Brown University, [3]National Institute of Mental Health

Object concepts are cognitive tools that reflect an interaction of features, such as shape, color, and value. A "Banana" is a yellow crescent that promises sweet nutrition. The neural computations by which visual input allows object concepts to be learned and deployed are poorly understood. A challenge is that concepts differ between people, and likelihood functions and priors are not precisely known. Monkeys could be a useful model – their visual diet can be curated, and tools exist to probe mechanisms at cellular resolution. But do monkeys form object concepts? We developed and deployed a foraging paradigm using in-cage touchscreens to address this question. We used an innovative color-shape space to define 2-dimensional objects and evaluated behavioral data using a reinforcement-learning framework. Smooth maps of the objects' reward values and frequencies of presentation were defined over a circular parameter space. Importantly, the paradigm was constructed so that the resulting marginal reward values for color and shape are useful but not sufficient for determining reward, color-shape contingencies are also informative. Four macaques performed a four-alternate-choice task daily for two years, achieving near-optimal performance. At plateau performance, the monkeys performed significantly better than a model that linearly integrates separable color and shape information (paired t-test, $p<1e-8$) - they can exploit interactions between shapes and colors. Moreover, the monkeys assigned higher value to exemplars further from reward boundaries (linear regression; $p=0.0$), predicted if they employ a cognitive model of object concepts. A Recurrent Neural Network uncovered the process through which the animals acquired concepts. The RNN suggests that the monkeys' default object concept is warm-and-blobby, and the reward marginals are learned before contingencies. These results show that monkeys, like humans, exploit object concepts that emerge as conjunctions of color and shape, conditioned predominantly by reward and less so by frequency of exposure.

23.339 REPRESENTATIONAL MOMENTUM FOR SPATIAL SCALING

Dominique Lopiccolo[1] (dominique.s.lopiccolo@gmail.com), Alon Hafri[1]; [1]University of Delaware

Representational momentum (RM) refers to the phenomenon where people reliably misremember an object's location as further along its implied motion path. While RM has been observed for a variety of transformations—ranging from rotating rectangles to melting ice to leaping ballet dancers—little is known about whether RM holds for spatial scaling, a non-rigid transformation that involves the resizing of an entity while retaining its shape. Scaling is critical to many spatial reasoning tasks, including map-reading, physical model construction, and geometric problem-solving. For this reason, it is important to know whether scaling is prone to the same distortions in visual memory observed for other transformations. This study sought to answer this question with a special focus on scaling in the context of geometric diagrams. Participants viewed dynamic displays involving two circles and a point in different spatial relations (e.g., inside, outside, intersecting). One of the circles either grew or shrank before the display was masked. Participants were tasked with selecting the exact last frame they saw from two options (neither of which was actually correct): one slightly later and one slightly earlier than the true target frame. We observed a significant RM effect: participants selected the later probe more often than the earlier one, with the effect more pronounced at the longer durations tested. Taken together, these findings suggest that spatial scaling induces RM effects in visual memory, paralleling other types of transformations. Our results contribute to a more general theory of RM and suggest broader implications for understanding geometric reasoning over spatial diagrams and memory for dynamic spatial relations.

23.340 DOES WHAT WE REMEMBER INFLUENCE WHAT WE SEE? EFFECTS OF VISUAL WORKING MEMORY ON PERCEPTUAL DISCRIMINATION

Zexuan Niu[1] (zexuan-niu@uiowa.edu), Andrew Hollingworth[1]; [1]University of Iowa

Visual Working Memory (VWM) has been proposed to be maintained by sustained activity in early visual cortex. This sensory recruitment hypothesis is supported by evidence of bidirectional influences between VWM and perception. VWM report can be biased systematically by new perceptual information. In addition, VWM content can influence psychophysical thresholds in a perceptual discrimination task (e.g., Teng & Kravitz, 2019), suggesting that VWM content biases perceptual experience. The latter evidence is provocative, as VWM is classically considered a form of visual memory that does not support conscious perception. Here, we examined the possibility that effects attributed to an interaction between VWM and perception are generated instead by interactions within VWM itself. We adapted the Teng and Kravitz (2019) method, in which participants completed a color discrimination task while holding a color in VWM. In the original study, the discrimination stimuli were spatially separated, and VWM may have been required to maintain color information across serial shifts of focal attention between stimulus locations, allowing for biases within VWM to influence discrimination performance. Here, we systematically varied the spatial separation between discrimination stimuli. When the spatial separation was relatively large, we successfully replicated the original effect indicating that the VWM color influenced discrimination thresholds. As spatial separation was reduced, reducing the demand to shift attention between locations, the effects of VWM on discrimination thresholds were likewise reduced. Crucially, when the discrimination stimuli abutted—so that discrimination could be achieved by attending to the color-contrast border, without the need to switch attention between locations—the effects of VWM on discrimination thresholds were eliminated. Together, our results indicate that effects of VWM on perceptual discrimination thresholds arise only when VWM is needed to support spatially separated comparisons and that, at least for the present task, VWM content does not directly alter perceptual experience.

23.341 EARLY VISUAL AREAS STORE INFORMATION ABOUT FEATURE BINDINGS IN WORKING MEMORY

Anna Zier[1,2] (zier@med.uni-frankfurt.de), Philipp Deutsch[1,2], Jochen Kaiser[1,2], Christoph Bledowski[1,2]; [1]Institute of Medical Psychology, Goethe University Frankfurt am Main, [2]Cooperative Brain Imaging Center (CoBIC), Frankfurt am Main

Previous studies on the neuronal basis of working memory have identified brain regions that temporarily store simple visual features like motion direction or color. In contrast, it remains unknown which brain areas temporarily store information about how those features are bound into an object. Each participant (N=20) performed a behavioral training and four separate functional magnetic resonance imaging (fMRI) sessions. Participants memorized two objects composed of two visual features each: color and motion direction. We manipulated the feature bindings within these objects across trials. To collect a sufficient number of binding repetitions per participant, each feature varied only across three exemplar categories (color: blue, yellow and pink; motion direction: leftwards, rightwards and downwards). Verbal strategies were minimized by slightly jittering color value and motion direction within each category. Together, our procedure allowed to isolate memorized binding information by comparing delay-related fMRI activity patterns between trials in which participants memorized pairs of objects with exactly the same features but the opposite binding. Additionally, the design enabled a comparison of activity patterns between trials that differed in a single feature, thereby isolating feature-specific information. Participants in our study were required to memorize two colors, two motion directions, and their binding on each trial. Despite this high memory load demand, we successfully decoded memory-related information about both motion direction and color from the visual cortex in an analysis of 10 of the 20 participants. Crucially, we were also able to decode the trial-specific binding information from the memory-related visual cortex activity. This novel, yet preliminary, finding indicates that early visual areas can retain specific information both about individual features and about the way these features are bound together.

23.342 STUCK TOGETHER: THE DIFFICULTY OF DIRECTED FORGETTING IN INTEGRATED WORKING MEMORY REPRESENTATIONS

Jiahan Yu[1,2] (psy_jyu@zju.edu.cn), Yingtao Fu[2], Mowei Shen[2], Hui Chen[2]; [1]Suzhou University of Science and Technology, [2]Zhejiang University

Integrated information is a crucial representational structure in working memory, where disparate elements are organized into cohesive units for efficient processing. Investigating how integrated information is selectively removed from working memory is crucial for understanding its storage mechanisms and operational principles. This study examined the directed forgetting of integrated information using retro cues presented during memory maintenance, instructing participants to forget specific parts of the memorized items. Experiments 1 and 2 focused on the directed forgetting of elements within a bicolor circle composed of two differently colored semicircles. In Experiment 1, the working memory-driven attentional capture effect was used to measure the degree of forgetting. Participants performed a visual search task after the retro cue. Results showed that the to-be-forgotten color caused greater attentional capture than the irrelevant color, indicating incomplete forgetting. In Experiment 2, the irrelevant change interference effect was used to measure forgetting. Participants judged whether a probe matched the to-be-memorized color. Longer response times for probes matching the to-be-forgotten color compared to irrelevant colors indicated incomplete forgetting. Experiments 3 to 5 extended these findings to multi-item sets organized by various cues. Experiment 3 investigated colored shapes that either conformed to or deviated from the Gestalt closure principle.

Experiment 4 investigated face pairs presented either facing each other or back-to-back to study the influence of social interaction cues. Experiment 5 explored real objects that were either semantically related or unrelated. Across all conditions, directed forgetting proved consistently more challenging for integrated information, regardless of the type of cue. This study highlights the inherent difficulty of forgetting integrated information in working memory, emphasizing the impact of integration on memory manipulation and control.

This work was supported by the Jiangsu Provincial Natural Science Foundation Youth Program (Grant No. BK20240992)

23.343 THE RELATIONSHIP BETWEEN PERCEPTUAL GROUPING AND VISUAL WORKING MEMORY'S POINTER SYSTEM

Shachar Lando[1,3] (shacharlando@mail.tau.ac.il), Roy Luria[1,2], Halely Balaban[3]; [1]Tel Aviv University, [2]Sagol school of neuroscience, [3]The Open University of Israel

Perceptual grouping plays a key role in enabling us to interact with a dynamic, details-rich visual world. Notably, while grouping cues such as Kanizsa figures demonstrate robust effects on various stages of visual processing, the cognitive processes responsible for grouping several objects into one unit are poorly understood. Here, we examined the interplay between perceptual grouping and the pointer-system, which is the mechanism implementing a steady correspondence between each object in the environment and its visual working memory (VWM) representation. Each unique pointer allows VWM to access the appropriate representation and modify it according to changes in the object's state in the world. Our main hypothesis was that perceptual grouping does not reflect pointer-system dynamics: each object in the group should still retain its independent pointer, while VWM maintains a compressed representation of the entire group. In our experiment, participants performed a change-detection task where three Pac-Man stimuli could move independently or in form of a Kanizsa triangle. On some trials, one Pac-Man abruptly changed its trajectory during the movement, breaking the Kanizsa in the grouped condition. We recorded EEG and monitored the contralateral delay activity (CDA), a neural marker of VWM. Our results indicated a lower CDA amplitude for the Kanizsa condition relative to independent movement, demonstrating a compressed grouped representation in VWM. Importantly, previous research revealed that breaking an object invalidates its VWM pointer, with a CDA drop indicating the temporary loss of the correspondence. In contrast, we found that breaking the Kanitza did not result in a CDA drop but a gradual increase, demonstrating an uninterrupted un-grouping process. This suggests that strong perceptual grouping is not the product of allocating a single pointer to the group, but rather of compressing several independent objects into one VWM representation, while still maintaining the original pointers for each group member.

23.344 LEARNING TO DECOMPOSE OBJECT-BASED ENCODING

Yingtao Fu[1], Longfei Ju[1], Mowei Shen[1], Hui Chen[1]; [1]Zhejiang University

Object has been considered as the basic unit for working memory encoding of external inputs. This is typically demonstrated by the involuntary encoding of task-irrelevant features on one object even when only one feature on that object is deemed task-relevant. The current study explored whether the object-based encoding could be decomposed through repetitive learning. Six experiments were conducted to test the working memory trace of task-irrelevant colors when they kept constant throughout the experiment. In Experiments 1 and 2, we found that the automatic encoding of the task-irrelevant color could be eliminated through repetition, providing initial evidence that object-based encoding can be modified through learning. Experiments 3 and 4 revealed that this learning effect diminished when multiple items needed to be learned, suggesting a strict capacity limitation for such learning. Additionally, Experiments 5 and 6 showed that the learning effect is constrained by the load of constant colors, rather than the load of objects containing those colors, indicating that the underlying unit of learning is a specific featural value. In summary, our findings demonstrate that object-based encoding can be further refined through repetitive learning of task-irrelevant features, with the effect being limited to a single featural value. These results highlight both the flexibility and capacity constraints inherent in cognitive selectivity.

Perceptual Organization: Segmentation, grouping

SATURDAY, MAY 17, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

23.345 CONNECTING THE LINES: CONTOUR INTEGRATION IS UNAFFECTED BY VISUAL INPUT SOURCE AND QUALITY

Angeline Yang[1] (angelineyang@berkeley.edu), Dongcheng He[2], Daniel R. Coates[3], Haluk Ogmen[4], Susana T.L. Chung[1]; [1]University of California, Berkeley, [2]Hefei Comprehensive National Science Center, [3]University of Houston, [4]University of Denver

Contour integration is a process whereby the visual system integrates local contour elements into a percept of a global shape. Previous electroencephalography (EEG) research has implicated the middle temporal N300 event-related potential (ERP) as a neural correlate of the contour integration process, along with the frontal P200, parietal contralateral and occipital N200, and parietal and occipital P400, which reflect attentional awareness, stimulus visibility, and task-related efforts, respectively. We aimed to investigate how these ERP signals associated with contour integration might be affected under binocular vs. monocular viewing, and in the presence of stimulus blur. Nine participants (6 females, 20.11±1.27 years) completed a contour integration task binocularly (no blur) and monocularly (no blur, 1, and 2 diopter of optical blur). The stimulus was a 27x15 array of 2x0.5˚ line segments that varied in orientation independently and randomly at 15 Hz, presented over 1,670 ms. At 735 ms, 12 line segments would form a 3x3 perfectly- or partially-aligned (by introducing orientation-jitter noise to the 12 segments) contour of a square for 152 ms at 6˚ to the right or left of a central fixation target. Participants indicated the contour location (right or left), under eye-position monitoring, while EEG activity was recorded. The amount of orientation-jitter noise was determined psychometrically for each viewing x blur condition before

EEG was performed. Across viewing conditions, participants' behavioral performance averaged 99.12% and 77.87% for the perfectly-aligned and partially-aligned conditions, respectively. There were no significant differences in the effects of viewing condition on behavioral performance or the signal strengths of the six associated ERPs. Additional Bayesian analysis favored the null effect over the effect of viewing condition on the ERP magnitudes. These results suggest that the neural correlates of contour integration are largely independent of the source (binocular vs monocular) and spatial-frequency content of local contour signals.

NIH Grant T32EY007043, UC Berkeley Undergraduate Research Apprenticeship Program

23.346 TEMPORAL DYNAMICS OF SPATIAL SEGMENTATION BASED ON TEMPORAL CORRELATION MEASURED BY CONTINUOUS TRACKING

Yen-Ju Chen[1] (ra51011daniel@gmail.com), Zitang Sun[1], Shin'ya Nishida[1]; [1]Graduate School of Informatics, Kyoto University

Temporally asynchronous changes in visual features induce visual spatial segmentation. We investigated the temporal dynamics of this phenomenon, particularly the temporal window for computing temporal correlation, using continuous tracking. Participants tracked by mouse a Brownian motion of the target. The temporal impulse response from perception to action was estimated from the cross-correlogram (CCG) of the trajectory between the target and mouse. The stimulus was a texture made of numerous Gaussian bulbs. Each bulb subtended 0.5 deg in diameter. Within the target area, the luminance or color (Red-Green or Blue-Yellow) of the bulb was temporally modulated with an interelement correlation of ~0.9. The background texture consisted of (1) static bulbs or (2) dynamic bulbs with an inter-element correlation of ~0.9 within the background, while the target-background correlation was ~0.0. Note that the target could be segregated from the background by temporal changes in the static condition, while by temporal asynchrony in the dynamic condition. We hypothesized that the difference in CCG between these two conditions must reflect additional temporal processing for correlation computation. The estimated CCGs (n=3) indicated slower and broader impulse responses for the dynamic condition than for the static condition. For luminance stimuli, the peak latency and half-height bandwidth were 333 ms and 209 ms in the static condition and 704 ms and 733 ms in the dynamic condition, respectively. Similar trends were observed for the color conditions. This temporal difference reflects an additional low-pass filtering in the dynamic CCG, providing an estimate of the temporal correlation window. Assuming a linear cascade processing system, the dynamic CCG is the static CCG convolved with the correlation processing time course. A mixed Gaussian kernel fit indicated that the time window for computing the correlation had a peak latency of 257 ms and a bandwidth of 753 ms.

This study is supported by SPRING, Kakenhi 24H00721, 20H05957

23.347 LUMINANCE EDGES INTERACT DIFFERENTLY WITH ORIENTATION- VS. CONTRAST-DEFINED EDGES

Christopher DiMattina[1] (cdimattina@fgcu.edu), Curtis L. Baker Jr.[2]; [1]Florida Gulf Coast University, [2]McGill University

Edges in natural images often give rise to spatially correlated changes in both first-order visual features (e.g., luminance), and second-order features (e.g., contrast or texture). Given the co-occurrence of first- and second-order cues at edges, it is important to understand how they are integrated for purposes of edge detection, and whether different kinds of second-order cues interact in a different manner with first-order luminance cues. Here, we compare the interactions between luminance modulation (LM) and contrast modulation (CM) with the interactions between LM and orientation modulation (OM) in an edge-detection task. Human observers detect which of two temporal intervals contain an edge between halves of a disc-shaped region with a filtered noise "texture" or carrier, whose contrast, orientation, or luminance differ across the boundary. We first characterize dependence on carrier spatial frequency, demonstrating different "tuning curves" for CM and OM, neither of which is explained by the carrier contrast sensitivity function. We then perform sub-threshold summation experiments in which observers detect either CM or OM in a 2IFC task with an LM pedestal varying from sub-threshold to supra-threshold levels. We find that an LM pedestal does not facilitate CM detection at sub-threshold levels, and actually masks CM at supra-threshold levels, consistent with previous studies. However we find that an LM pedestal has little or no effect on OM detection. When we consider detection of LM with CM and OM pedestals, CM pedestals impair the detection of LM, whereas OM pedestals do not. Therefore, we see that different forms of second-order structure interact in a different manner with first-order structure, with CM and LM exhibiting a mutually antagonistic relationship, and LM exhibiting little interaction with OM. These results suggest distinct mechanisms for detecting edges defined by differences in orientation or by changes in contrast.

C.D. is funded by NIH grant R15-EY032732-01, C.B. is funded by Canadian NSERC grant RGPIN-2023-03559

23.348 PSYCHOPHYSICAL REVERSE CORRELATION EXPERIMENTS ARE ACCELERATED VIA NOVEL TAILORED NOISE GENERATION

Ethan Duwell[1] (eduwell@mcw.edu), Adam S. Greenberg[1]; [1]Medical College of Wisconsin

Psychophysical reverse correlation (PRC) is a powerful data-driven method applicable in a wide array of sensory domains. In visual PRC experiments, subjects perform tasks on base images (BIs) with noise overlaid. After many trials, noise from correct and incorrect trials are quantitatively combined to generate classification images (CIs) which highlight BI features of a hypothetical internal template used in performing the task. Unfortunately, the number of trials required (5-10k) severely limits PRC's practicality. Our previous work suggests that specific noise characteristics can greatly improve PRC efficiency, and noises with feature profiles closer to the BIs perform best. But how might the noise be optimized for any given PRC experiment? Here, we address this via a novel noise-generation algorithm which samples the BIs with varying kernel window sizes and combines the output to form noise frames with spatial frequency (SF) spectra closely matching the BIs. Performance was tested in 3 different simulated tasks. In two tasks, a simulated observer detected right vs. left angled Gabors with 2 and 17 cycles/image, respectively. In the third task, the BIs were smiling vs. neutral human faces. Three different trial counts (1k, 5k, and 9k) and four different noise conditions were used: 'tailored-noise' (sampling BIs from the same task); 'tailored-noise' (sampling BIs from the other 2 tasks); and white noise. Each configuration was repeated 50 times and CI quality was assessed via structural similarity between resulting CIs and a veridical CI. In all tasks, tailored noise performed best and produced interpretable CIs within 1000 trials. Interestingly, noise SF content impacted which features were emphasized in resulting CIs. Thus, noise choice is critical for PRC optimization but can also act as a filter on the features resolvable in the CI. Overall, PRC may now be more feasible given the reduced trial count required to produce CIs.

This work was funded by NSF grant SBE 2122866 to A.G.

23.349 VISUALIZED AVERAGES PRODUCE POLARIZED SENTIMENT

Jeremy B Wilmer[1], Tugral Bek Awrang Zeb[1], Sarah H Kerns[1,2], Ken Nakayama[3]; [1]Wellesley College, [2]Dartmouth College, [3]University of California at Berkeley

A common graphing practice in science, health, education, and media is to focus on a single summary statistic, such as the average (mean) value. By hiding the variability in the raw data, a graphed average provides a single, salient reference point for comparison; a benchmark against which individual values, or other average values, may be compared. While known to be an over-simplification, such simplicity is deemed necessary for clear visual communication, especially to a general audience (Kerns & Wilmer, 2021). It is common practice, for example, in standardized test score reports, such as the one for the statewide K-12 Massachusetts Comprehensive Assessment System (MCAS), to provide the average score as a reference point. Yet could a single benchmark be over-interpreted? Potentially, a salient average value could polarize people's sentiments about scores. Rather than interpret the average correctly as merely the center of a spectrum, people might interpret it, without nuance, as a categorical boundary that artificially splits scores into two types: above average (good) and below average (bad). Here, we show that this is indeed the case. When asked to rate how happy one would be to receive various potential scores, shown relative to an average value, people's ratings are strongly polarized, with a disproportionate change when crossing the average score, relative to identical increments that do not cross the average. We next demonstrate a powerful, complete solution: show the data. When a full spectrum of individual values is graphed in place of the mean, polarization vanishes. We replicate both presence and removal of polarization across eight different practical domains, ranging from education to health to life skills. We conclude that graphed average values unnecessarily fuel polarized sentiment, and that an effective, actionable solution is to plot individual values.

Funded in part by NSF award #1624891 to JBW, a Brachman Hoffman grant to JBW, and a subaward from NSF grant #1837731 to JBW.

23.350 VISUAL IMAGINATION NETWORKS IN HUMANS AND LARGE LANGUAGE MODELS

Saurabh Ranjan[1], Brian Odegaard[1]; [1]University of Florida

Visual imagination is an aspect of consciousness that results in experiences of the sensory world without any external stimulation. How is the experience of imagination structured in human minds and machines? To explore this question, we probed how the vividness ratings of different imagined experiences were associated with one another in both humans and large language models (LLMs). We analyzed responses from the Vividness of Visual Imagery Questionnaire (VVIQ), which requires participants to imagine different visual aspects of a scene under given context, and report their vividness on a rating scale (1-5). To understand how the vividness of different experiences were associated with one another, we first constructed imagination networks using pairwise partial correlations (edge weights) of the items (nodes) from 1,776 human responses to the VVIQ. Next, we constructed the same networks based on VVIQ responses from LLMs to understand how their generative behavior differs from human experiences of visual imagination. We found that humans exhibited more positive edge associations between the vividness of different items, but LLM responses showed more negative edges. These differences in edges resulted in topological differences in the imagination networks as shown across multiple centrality measures like strength, closeness, betweenness, and expected influence of each node. Further, to understand how vividness of different experiences clustered in the networks, we investigated the community structure of nodes in networks from humans and LLMs. While all eight VVIQ contexts clustered in human imagination networks; clustering of items from LLM networks was extremely diffused. Overall, our study not only reveals how the vividness of different experiences are associated with each other in human visual imagination, but also shows how responses differ in the imagination of LLMs. Together, our results reflect differences in internal world-building across natural and artificial generative processes, resulting in different vividness responses to visual imagination.

No Funding Sources Available

23.351 COLOR/MOTION FEATURE BINDING ERRORS WITH 2D AND 3D MOTION

Sunny M Lee[1], Steven K Shevell[1]; [1]University of Chicago

The effortless perception of a flying red frisbee belies an active feature binding process of the visual system that integrates separate features, such as color and motion. In the periphery, sparser neural representations can cause the veridical feature pairing of color and motion to be overwritten by that in the center. Currently, the limits of the stimulus differences that still result in induced feature misbinding are unknown. Aim: We investigated how 2D and 3D motion interact across the central and peripheral visual field by measuring induced color/motion feature misbinding. Method: Participants reported the perceived motion direction of dots of one color (red or green) in the periphery, while fixating centrally and viewing two sets of colored dots (one set red, one set green) moving in opposite directions. Misbinding was induced by reversing the motion directions assigned to the two colors of dots between the center and the periphery. Dots moved radially from center to periphery, either contracting or expanding to give the percept of 3D radial optic flow motion or 2D "flat" radial motion. A control condition presented vertically moving dots at different speeds in center and periphery. Results: Significant color/motion feature misbinding in the periphery was found when the stimuli were entirely a single motion type, either 2D "flat" motion or 3D optic flow motion. Overlaying 2D and 3D radial motions, however, reduced misbinding by over 50% compared to both sets of colored dots moving as a single

motion type. Significant misbinding observed with vertically moving 2D dots of different speeds demonstrated that the different dot speeds in 2D versus 3D motion cannot explain these misbinding results. Conclusion: Reductions in peripheral feature misbinding with different motion types, but not speeds, is consistent with different visual-processing levels for color-motion feature binding for objects perceived in 2D and 3D motion.

Supported by NIH EY-026618

23.352 FIGURE SPACE IS DISTORTED RELATIVE TO GROUND SPACE

Timothy Vickery[1], Kallie Sweetman[1], Mary Peterson[2]; [1]University of Delaware, [2]University of Arizona

The presence of an object in a 2-D display distorts perceived space. For example, two dots placed on a region perceived to be an object appear farther apart than two equally spaced dots placed on region that appears to be its background ("object-based warping"). The extent to which perceived objecthood per se generates this distortion, as opposed to image factors, is still debated. We asked whether cues that lead to border assignment in classic figure-ground displays could generate distortions of space when other factors are held constant. We employed horizontally elongated rectangular images divided into eight alternating black and white convex and concave regions, presented on a gray background. (The black/white color of convex and concave regions was balanced and they were equal in area). In such displays, the convex regions are highly likely to be perceived as figures/objects and the concave regions appear to form a background. We placed two fixed-spacing "reference" dots on either a convex or concave region, and two "adjustment" dots on the other region type. Subjects adjusted dot spacing to match the reference. In a control condition, all dots were placed on a uniform rectangle of white or black. Subjects systematically estimated the dots as farther apart when the reference dots were on a figure region versus a ground region ($p < .001$). Furthermore, ground-region estimates were significantly underestimated compared to control estimates ($p < .001$) and figure-region estimates were significantly overestimated relative to control estimates ($p < .001$). Thus, even on displays matched for local boundedness of the dots, assignment of relative objecthood distorted space, such that perceived spatial distance was expanded in object regions, and shrunk in ground regions. The use of such figure-ground displays allows for the matching of many important low-level properties in the study of object-based distortion effects.

Supported by NIH NIGMS grant 5P30GM145765-03

23.353 **ARE YOU A VISUAL ‘SHADER’ OR A ‘BOLDER’?:** DIFFERENT VISUAL ROUTINES CREATE EVERYDAY **HALLUCINATIONS IN ‘SCAFFOLDED ATTENTION’**

Andrea Ying[1], Joan Danielle K. Ongchoco[1]; [1]The University of British Columbia

Visual processing of incoming sensory cues gives rise to the rich colours that fill contours and the contours that form objects. But people can also experience colours and contours in the *absence* of explicit sensory cues, albeit in more fleeting ways — as in the phenomenon of “scaffolded attention”. Consider a regular grid of squares. By definition, there is no structure there, but many people report seeing various shapes and patterns anyway (e.g., horizontal lines, block letters). But beyond *what* people see, perhaps more intriguing is *how* they experience it. Some describe the squares of perceived patterns to be brighter or differently shaded (i.e., “shaders”), while others note the squares as being ‘outlined’ or ‘traced’ (i.e., “bolders”). What determines when people experience one ‘type’ over another? Here observers reported which type they experienced, and reproduced the magnitude of bolding and/or shading through an interactive grid. We then explored the influence of *external* grid features (e.g., white squares with black outlines vs. black squares with white outlines), and *internal* factors (e.g., attentional breadth [via the ‘Functional Field of View task; FFOV], and sensitivity to figure-ground boundaries [via the Leuven Embedded Figures Test; LEFT]). First, the proportions of shaders and bolders overwhelmingly differed across grids, with reliably more bolders for black (93.2%) than for white grids (41.1%) — perhaps because the contrast differences change whether the squares or the lines are seen as figure or ground. Second, only the LEFT (and not FFOV) scores predicted whether people would be a ‘shader’ or ‘bolder,’ highlighting the role of segmentation processes in scaffolded attention. Thus, people’s everyday hallucinations can depend on what the mind selects — the squares on the white grids or the lines of black grids — which may be grouped together through different visual routines.

23.354 VISUAL CROWDING DIFFERENTIALLY AFFECTS BAR AND EDGE DISCRIMINATION

Emma Neto[1], Akilesh Sathyakumar[1], Isabella Crescenzi[1], Jessica Lim Tung Tseung[1], Allison B. Sekuler[1,2,3], Patrick J. Bennett[1]; [1]Department of Psychology, Neuroscience & Behaviour, McMaster University, [2]Rotman Research Institute, Baycrest, [3]Department of Psychology, University of Toronto

Relative phase discrimination is performed using even-symmetric and odd-symmetric visual mechanisms (bar and edge detectors, respectively; Field & Nachmias, 1984). In central vision, even- and odd-symmetric mechanisms are approximately equally sensitive, but it is unclear how sensitivity varies with increasing stimulus eccentricity. Studies that measured discrimination thresholds with compound gratings (Bennett & Banks, 1987) found that sensitivity of odd-symmetric mechanisms decreased dramatically in peripheral vision; however, studies that used isolated bars and edges (Morrone et al., 1989) found that sensitivity of even- and odd-symmetric mechanisms was nearly equal across the visual field. These different results could be reconciled if lateral masking in the periphery was much greater for edge discrimination than bar discrimination. We tested this hypothesis by measuring bar and edge discrimination thresholds for various inter-feature distances. Stimuli were modified versions of the even- and odd-symmetric stimuli used by Morrone et al. and consisted of 1, 2, 4, or 8 light/dark bars or right/left edges. Discrimination thresholds were measured by varying stimulus contrast in a 2-IFC task. In four observers, increasing the number of features (i.e., decreasing the inter-feature distance) had virtually no effect on edge and bar discrimination thresholds in central vision. However, when stimuli were presented at an eccentricity of 7.6 deg, increasing the number of features dramatically increased edge discrimination thresholds but had virtually no effect on bar discrimination thresholds. These results are consistent with the hypothesis that peripheral edge discrimination is especially sensitive to lateral masking.

NSERC

23.355 LATENT HIERARCHICAL PERCEPTUAL STRUCTURE BIAS COALITION FORMATION IN THE STRATEGY OF CONFLICTS

Zhen Li[1], Tao Gao[2]; [1]Zhejiang University, [2]UCLA

Vision plays a crucial role in non-visual processes by offering structured, universal representations that other cognitive systems can leverage without maintaining their own visual systems. This capability allows vision to influence downstream cognition, as seen in phenomena like the minimal group effect, where arbitrary color labels shape group identification, and focal point effect, where visual salience aids coordination among identical options. This study explores how latent hierarchical visual structures influence strategies of conflict and coalition formation in adversarial contexts. We developed a three-player territory expansion game where competition and cooperation coexist. Players competed for fixed territories while avoiding costly battles. Alliances, though beneficial, were challenging due to the absence of communication and mechanisms for enforcing alliances. Players navigated ambiguous relationships, where any individual could be an ally or adversary at any moment. In Experiment 1, participants played on one of two game boards: a uniformly colored Non-Structured board or a Structured board inspired by Mondrian art, featuring hierarchically organized colored blocks generated by a latent parse tree. Although irrelevant to gameplay rules, the visual structure significantly reduced battle frequency and intensity. On the Non-Structured board, players engaged in "a war of every man against every man," as described by Hobbes. Conversely, on the Structured board, players residing in sibling groups within the perceptual hierarchy were more likely to ally against the non-sibling group, which ended with the smallest territory. Crucially, these effects emerged only through gameplay. In Experiment 2, when participants were asked about their strategy after observing the structured board without gameplay, they showed no recognition of non-sibling group disadvantages or preference for allying with sibling groups. These results underscore that humans take advantage of structured visual representations in conflict resolution by transforming a structure of perception into a structure of bargaining, provided they share the same visual scene.

23.356 PATH INTEGRATION REFLECTS AN INTERMEDIATE REPRESENTATION IN CONTOUR PERCEPTION

Austin S. Phillips[1], Philip J. Kellman[1]; [1]University of California, Los Angeles

Path integration refers to perceptual grouping of oriented elements, often Gabor patches, embedded in a random array. Although the geometric relations that produce the pop-out of paths correspond to those that produce completion of illusory and occluded contours (Hess & Field, 1999; Kalar et al, 2010), path detection is not accompanied by perception of illusory contours, and the relation to contour interpolation has remained unclear. A two-process model of contour interpolation (Kellman & Fuchser, 2023) proposes that path detection depends on an intermediate representation of contours linked across gaps purely by their geometry, but that perception of continuous contours depends on additional constraints that act on this representation. The higher-level constraint appears to be that Gabors do not support perception of continuous contours because linked contours do not enclose a consistent surface color within and between elements. Altering Gabors so that their centers match the surround produces clear illusory contours connecting path elements (Kellman, Erlikhman & Carrigan, 2016). Here, we further tested path integration and illusory contour perception, using Gabors and modified elements with centers matching the surround. Participants searched for paths with 0-75 degree turning angles in a 2IFC paradigm (Field, Hayes & Hess, 1993). Consistent with the two-process model, we hypothesized that (1) elements supporting illusory contours would show no advantage in path detection, and (2) that illusory contours would be perceived only with elements satisfying both the geometry of relatability and the surface continuity constraint. The results confirmed both hypotheses. All element types showed similar path detection functions, and illusory contours were perceived only in displays supporting surface connections. For those displays, perceived illusory contours mirrored the path detection results. These results support the idea that path integration performance depends on an intermediate stage of contour-linking, necessary but not sufficient for perception of continuous contours across gaps.

We gratefully acknowledge support for this research from National Institutes of Health grant R01CA236791 to PK.

SATURDAY MORNING POSTERS IN PAVILION

Face and Body Perception: Experience, learning, expertise

Saturday, May 17, 8:30 am – 12:30 pm, Pavilion

23.401 AN OLDER BIAS OR AN ELDER BIAS?: DO MEMORY BIASES FOR OLD FACES REFLECT ASSIMILATION TOWARD AN OLD-AGE EXTREME OR TOWARD AN OLD-AGE CATEGORY CENTER?

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? In our previous work, we found that a briefly viewed, 60-year-old face is remembered as older than it really was: a 60-year-old face is more likely to be remembered as matching the age of a decoy which is ten years older, than it is to be remembered as matching the age of a decoy which is ten years younger. Does this effect reflect assimilation toward the extreme end of the age spectrum (around 90+ years old) or rather toward an 'old' prototype (perhaps around 75 years old)? To investigate this, here we tested whether memory for the age of an 80-year-old face is assimilated upward or downward. Participants saw a target 80-year-old face and subsequently saw two new decoy faces – one 70-years-old, and one

90-years-old — and selected the face that they thought matched the age of the target. We found a robust bias for participants to select the younger, 70-year-old decoy, suggesting that memory is assimilated towards a septuagenarian 'old' prototype.

23.402 TRAINING HORIZONTAL BIAS TO IMPROVE FACE IDENTIFICATION IN OLDER ADULTS

Allison B. Sekuler[123], Jamie G.E. Chocrane[3], Runzhi (Nina) Yue[12], Konka Paul[1], Yifei Wang[12], Ali Hashemi[34], Eugenie Roudaia[1], Patrick J. Bennett[3]; [1]Rotman Research Institute, Baycrest Academy for Research & Education, Toronto, ON, Canada, [2]University of Toronto, Toronto, ON, Canada, [3]McMaster University, Hamilton, ON, Canada, [4]Wilfred Laurier University, Milton, ON, Canada

The preferential use of horizontal facial structures, known as horizontal bias, positively correlates with face identification performance. Older observers, who have lower face identification abilities than younger observers, tend to show less horizontal bias; and, within each age group, individuals with greater horizontal bias tend to have better face identification. However, the extent of an individual's horizontal bias is not always fixed: Training younger adults to enhance horizontal bias improves face identification for both trained and novel faces, with the greatest degree of transfer observed when uninformative vertical context is included during testing. The current study investigates whether training horizontal bias, with or without uninformative vertical context, can enhance the horizontal bias and improve face identification in older adults, and whether these effects transfer to novel faces. Older adults completed a 10-alternative forced-choice task, identifying briefly presented faces filtered with horizontal or vertical orientation filters (alternating across blocks). Identification thresholds, which were estimated by varying the orientation filter bandwidth across trials, were measured pre- and post-training (days 1 and 4) in eight conditions that differed in face set (trained/novel), filter orientation (horizontal/vertical), and context (present/absent). Training occurred on days 2 and 3, with participants randomly assigned to one of two conditions: horizontal filters without context, or horizontal filters with uninformative vertical context. Both types of horizontal bias training improved face identification performance and reduced threshold bandwidths. Both training programs also showed partial transfer to novel faces with informative horizontal and vertical structure, with and without uninformative context. However, greater transfer was seen in the horizontally filtered conditions. These results suggest training horizontal bias can enhance face identification in older adults. Future research will examine whether this training improves face identification in other face-perception impaired groups, and determine how training affects the neural signatures of face perception.

Support from NSERC Discovery Grants (ABS and PJB), and Baycrest Foundation (ABS)

23.403 MODULATION OF FACE PROCESSING BY MOTOR DEVELOPMENT: INFANTS CAN DISCRIMINATE FACES IN A WIDER RANGE OF PICTURE-PLANE ROTATIONS BEFORE ROLLING OVER

Megumi Kobayashi[1] (mkobayashi@human.niigata-u.ac.jp); [1]Niigata University

It is well known that visual perception and motor ability are interrelated in infants' development; e.g., optical-flow perception drastically changed just before the emergence of voluntary locomotion (Shirai & Imura, 2013). We report evidence that motor development modulates infants' face perception. In infancy, face input to infants would shift with their motor development. For instance, infants after rolling over tend to observe more upright or nearly upright faces, whereas infants before rolling over observe from various angles. These changes in visual input with motor development should modulate the development of face processing. In adults, a non-linear relationship between performance and the angle of rotation of faces is shown, which is a significant decline in face recognition between 60° and 90° (e.g., Rossion & Boremanse, 2008). These results imply the difficulty in configural face processing at orientations of faces that are less frequently experienced. We hypothesized that the non-linear relationship between face discrimination and picture-plane rotation in infants may change with motor development. To test this hypothesis, we examined face discrimination ability at various rotation angles (0° to 180°, 4 angles) in 3- to 8-month-old infants. We confirmed that infants showed significant face discrimination when presented upright (0°, $p < .05$) but not upside-down (180°, $p > .05$). Also, as shown in adults, they discriminated between faces at 60° ($p < .05$). At 120°, however, infants before rolling over showed significant discrimination ($p < .05$), whereas infants after rolling over did not ($p = .19$). Furthermore, infants after rolling over showed a significant negative correlation between face discrimination at 120° and the number of days from the date of acquisition of rolling over ($p < 0.05$). These findings suggest that alterations in face input associated with motor development play a pivotal role in the development of face perception.

This research was supported by Grant-in-Aid for Scientific Research (B) from JSPS (23H01055/23K25752)

23.404 ELIMINATING INDIVIDUAL FACE SHAPE VARIABILITY IDIOSYNCRATIC TO SPECIFIC RACES REDUCES THE OTHER-RACE RECOGNITION DISADVANTAGE, BUT NOT THE OTHER-RACE CATEGORIZATION ADVANTAGE

Emily Martin[1] (emart459@fiu.edu), Fabian Soto[1]; [1]Florida International University

Other-race effects in face perception are well-documented and extensively studied in the literature. Specifically, the other-race recognition disadvantage (ORRD) involves a reduction in face recognition memory for races different from the perceiver's compared to the own-race, while the other-race categorization advantage (ORCA) involves faster and more accurate classification of other-race faces than own-race faces. These effects could stem from reliance on features that distinguish the two races (race-specific) or on features that distinguish individuals within each race (identity-specific) and that differ between the two races (race-idiosyncratic). Using three-dimensional face modeling, we generated faces in which identity-specific information was held constant across races (i.e., individuals differed from one another in the exact same ways in both races), but race-specific information was included and implemented as anthropometrically-validated three-dimensional race morphs. Controlling for self-reported other-race contact, we presented participants with our novel stimuli using both the ORRD and ORCA tasks. We find that our stimulus manipulation eliminated the ORRD but preserved the ORCA, suggesting that different mechanisms may underlie the two effects. Considering that identity-specific information was held constant across races while race-specific information was not, our results suggest the ORRD may result from a sensitivity to features that distinguish individuals in the own-race but not the other-race, while the ORCA may result from a sensitivity to race-specific features.

This work was supported by the National Science Foundation.

23.405 BEAUTY IS IN THE EYE(BROW) OF THE BEHOLDER: PERCEIVED ATTRACTIVENESS RELATED TO EYEBROWS

Jessie Peissig[1], Kareena Brennan[2], Rachel Salamone[1]; [1]California State University Fullerton, [2]California State University Long Beach

Many features of the face have been studied in relation to attractiveness. However, there are very few studies that look specifically at eyebrows. The worldwide market for eyebrow products was estimated to be $5.5 billion in 2023, suggesting eyebrows play an important role in the perception of facial attractiveness. This study explored specific features of eyebrows that can be manipulated with easy changes, like makeup and tweezing, to see how they may related to attractiveness. Different conditions were created to change the thickness, thinness, darkness, and lightness of the eyebrows. Both younger (n=171) and older adults (n=36) participated in this study, to test for any generational or age-related differences. Participants were shown 100 images of faces of younger men and women and then asked to rate how attractive they perceived each face. For younger participants, the main effect of eyebrow type was significant, $F(4, 170) = 5.10$, $p < .0005$. This indicates that participants responded differently to the different eyebrow conditions. Planned contrast tests showed that overall, the original eyebrows were rated as more attractive than the lighter eyebrows ($t = 4.45$, $p<.001$), the thicker eyebrows ($t = 3.96$, $p<.001$), the thinner eyebrows ($t = 3.96$, $p<.001$), and the darker eyebrows ($t = 2.65$, $p < .05$). Older participants showed a similar pattern, with a significant main effect of eyebrow type, $F(4, 35) = 6.46$, $p < .0005$. Planned contrasts showed that the original eyebrows were rated as more attractive than the lighter eyebrows ($t = 2.69$, $p<.01$), the thinner eyebrows ($t = 3.71$, $p<.001$), and the darker eyebrows ($t = 2.55$, $p < .05$), but not thicker eyebrows ($t = 1.61$, $p>.05$). Overall, participants rated the original eyebrow condition as more attractive than most of the other conditions, indicating a preference for natural eyebrows.

23.406 CONSTRUCTING THE CROSS-RACE TRIAD IDENTITY MATCHING (CRTIM) TEST

Geraldine Jeckeln[1] (geraldine.jeckeln@utdallas.edu), Alice J. O'Toole[1]; [1]The University of Texas at Dallas

Cross-race face identification is done routinely in applied settings (e.g., forensic, security). Despite the error-prone nature of cross-race identification, no publicly available tests exist to evaluate individuals' identification ability for own- versus other-race faces. Our goal was to develop a test of African American (AA) and Caucasian (CA) test items that would: 1) challenge individuals of varying abilities, 2) ensure comparable identification accuracy for both AA and CA (no item-race effect), and 3) provide measures of item difficulty. Item selection was guided by the performance of a face-identification algorithm (Szegedy et al., 2017) and AA and CA observers (n = 34 per race). The resulting test includes 25 AA and 25 CA face-image triads. Each triad contains two images of the same person and one image of a different person. The task is to select the image of the "different" person. We found a classic cross-over "other-race effect" [interaction of item and participant race: $F(1, 66) = 4.33$, $p = 0.041$, $\eta 2p = 0.06$; AA participants (AA faces: $M = 0.74$; CA faces $M = 0.71$); CA participants (CA faces, $M = 0.75$; AA faces $M = 0.72$)], with no main effect of item or participant race ($p > 0.05$). Next, item difficulty was assessed using Item Response Theory (Lord, 1980). This provides participant-ability and item-difficulty estimates on the same scale. Item responses for each item-race set and participant-race group were modeled separately. Results showed that test items covered a wide range of item difficulty levels, with average difficulty falling slightly below participant ability. Overall, the novel face-identification test allows for accurate assessment of participants' identification abilities for both AA and CA face identities. This work also provides a guideline for designing cross-race tests using the output of face-recognition systems and human observers.

Research funded by The National Institute of Standards and Technology, Grant 70NANB22H150 to A.OT.

23.408 HAVE WE MET BEFORE? FACE LEARNING OVER MULTIPLE INTERACTIONS

Alice Nevard[1], Soazig Casteau[1], Ulrik Beierholm[1], Holger Wiese[1]; [1]Durham University

Learning a new face can be described as forming a robust representation over time that allows for recognition in novel situations. Event-related potential research has linked this process with more negative amplitudes for learnt compared to novel faces at occipito-temporal channels from ~200 ms onwards. While this N250 effect represents robust learning, its amplitude is typically small - reflecting limited exposure to novel faces, which are typically learnt in a single session. The current research examines how new face representations are established over multiple sessions by comparing the N250 after repeatedly meeting a new person. In pre-training sessions, participants saw highly variable ambient images of the "to be learnt face", an "unmatched face" and a "matched" identity, the latter resembling the first with respect to age, gender, hair colour and ethnicity. In subsequent sessions across four days, participants first had a ten-minute in-person interaction with the same learnt identity, followed by a test phase in which participants saw ambient images of the three identities. While new images were used in all sessions, the learnt and unmatched identities were repeated. However, a new matched identity was presented in each session. There were no differences in the N250 in the pre-training session. However, after one interaction, N250 responses were larger for the learnt face compared to the unmatched identity. Critically, this effect was substantially enhanced after four interactions. Moreover, learning effects in comparison to the matched faces were generally small. Our findings show how representations of novel faces build up with repeated exposure. Importantly, while in-person interactions provide sufficient information for building such representations over multiple sessions, repeatedly viewing ambient images of the same person does not.

23.409 SERIAL DEPENDENCE OPERATES ON CATEGORICAL RATHER THAN STIMULUS REPRESENTATIONS: EVIDENCE FROM BEHAVIOR AND EEG

Pierre Costa[1] (pierre.costa@u-paris.fr), Thérèse Collins[1]; [1]Integrative Neuroscience and Cognition Center, Université Paris Cité & Centre National de la Recherche Scientifique (CNRS), Paris, France

Our environment is constantly changing but our perception is stable. This stability may result from the integration of successive visual images that smooths over spurious variations in input. Such a process manifests in a behavioral phenomenon called Serial Dependence (SD), an attractive bias between successive stimuli. Here we questioned whether SD operates on low-level stimulus representations or at a higher level. We constructed face morphs from three prototypes of the same face expressing anger, fear and sadness. We measured the perceived similarity between each pair of morphs with an odd-one-out experiment, to build a behavioral representational matrix. As expected, emotional expressions were perceived categorically, with greater perceived similarity between pairs within a category than between pairs across categories, for a fixed distance in stimulus space. Subjects then reproduced the perceived facial expression of successively presented morphs by adjusting a response cue while their cerebral activity was recorded with EEG. The categorical effect was mirrored in the EEG signal in a given time window and group of electrodes. Using representational similarity analysis, we computed an EEG-based correlation matrix based on the similarity between each pair of morphs. This neural representational matrix correlated more with the behavioral representational matrix than with the stimulus matrix. We also correlated the three matrices (stimulus, behavioral, neural) with the SD matrix, i.e. the extent to which perceptual reports were biased towards the facial expression seen in the past (SD). SD was stronger between two faces belonging to the same category than between two faces that did not belong to the same category, for a given stimulus difference. The SD matrix correlated better with the representational matrices than the stimulus matrix. These results reveal categorical SD, showing that the temporal integration serving visual stability operates on high-level visual representations.

ANR CNRS

23.410 HOW WE LEARN NEW FACES: GAZE PATTERNS AND INDIVIDUAL DIFFERENCES IN REAL-LIFE ENCOUNTERS

Soazig Casteau[1], Alice Nevard, Ulrik Beierholm, Holger Wiese; [1]Durham University

Face learning is an important cognitive ability that has been extensively studied, however, the precise factors contributing to this process remain unclear. Previous studies have shown that greater focus on specific facial regions (e.g., eyes) during face learning enhances recognition performance (e.g., Henderson et al., 2005). However, individual variations in optimal fixation locations suggest that effective strategies may differ across individuals (Royer et al., 2018; Peterson & Eckstein, 2013). The aim of the current study was to investigate the relationship between gaze patterns, individual differences, and face recognition performance in naturalistic settings. Participants engaged in a 10-minute real-life interaction with an unfamiliar person while their eye-movements were recorded using mobile eye-tracking. We then assessed individual personality traits as well as cognitive abilities related to face processing using personality and psychometric tests, and measured participants' ability to identify the learned face using a recognition test. Our results confirmed the importance of the eye region during face learning, as participants spent approximately half of their interaction time focused on faces, with around half of that time directed at the eyes. Interestingly, the proportion of time spent looking at the face or the eyes did not strongly predict participants' recognition performance. However, we found associations between recognition scores and personality traits. Specifically, there was a negative correlation between neuroticism and recognition accuracy. Conversely, extraversion and conscientiousness showed a positive correlation with recognition accuracy. Moreover, scores on the Cambridge Face Memory Test (CFMT+) and the Glasgow Face Matching Test (GFMT2) were significantly associated with accuracy in recognizing newly learnt faces. Our study confirms that while focusing on faces, particularly the eyes, is important for effective face learning, individual variability plays a significant role in this process. Particularly, personality traits and face recognition abilities are key factors influencing how well individuals recognize newly encountered face.

This research has been supported by an ESRC grant (ES/X014657/1)

23.411 HOLISTIC PROCESSING IN OCCIPITAL AND FUSIFORM CORTICAL COMPLEXES RELATED TO RADIOLOGICAL EXPERTISE

Haowen Guan[1], Min Guo[2], Yaou Liu[2], Sheng He[1]; [1]Institute of Biophysics, University of Chinese Academy of Sciences, [2]Tiantan Hospital, Capital Medical University

Radiologists' expert-level performance in detecting abnormalities in chest X-rays could be supported by holistic processing, a mechanism exemplified in face recognition. Visual areas such as the fusiform face area (FFA) and occipital face area (OFA), as indicated by their names, have been considered to be specialized in face processing, their role in holistic processing associated with visual expertise remains unclear. This study investigates the involvement of holistic processing when radiologists performed visual recognition tasks. Behavioural and 3T fMRI experiments were conducted with radiologists and laypersons. Participants viewed three types of images (faces, buildings, chest x-rays) with the images presented in two modes (upright vs. inverted and amodally completed vs. uncompleted), to modulate the degree of holistic processing. Behavioural results confirmed that radiologists rely on holistic strategies when interpreting chest X-rays. Multivariate pattern analysis (MVPA) of the fMRI data revealed that when holistic processing was disrupted, radiologists' decoding accuracy for X-rays versus buildings significantly decreased in the left FFA (lFFA) and left OFA (lOFA), suggesting that these regions encode expertise-specific holistic information. Furthermore, the right FFA (rFFA) showed improved decoding accuracy for inverted faces and stronger category selectivity for faces over buildings in radiologists, suggesting that radiological expertise modulates neural responses to stimuli that rely on holistic processing, regardless of whether the expertise for the images was professionally acquired or not. These results demonstrate that the lFFA and lOFA are central to the encoding of holistic information in radiological expertise, whereas radiological training also enhanced the rFFA's existing sensitivity to faces. This study advances our understanding of the neural basis of expert perceptual processing and its relationship to holistic mechanisms.

Face and Body Perception: Individual differences

Saturday, May 17, 8:30 am – 12:30 pm, Pavilion

23.412 DIFFERENCES BETWEEN THE PERCEPTION AND THE RECOGNITION OF CROSS-RACE FACES

Dilhan Töredi[1], Eric Y. Mah[2], Megan Lall[3], Haiyang Jin[4], Brett D. Roads[5], James W. Tanaka[2]; [1]John Jay College of Criminal Justice, [2]University of Victoria, [3]Adler University, [4]Zhejiang Sci-Tech University, [5]University College of London

The cross-race effect (CRE) refers to better recognition of same-race than cross-race faces (Lee & Penrod, 2022). The Face Space Model (Valentine, 1991) attributes the CRE to the fact that more frequently encountered same-race faces are represented with diverse features in perceptual space, making them more differentiated in memory. In contrast, less frequently encountered cross-race faces cluster around race-typical features, resulting in high perceptual similarity and reduced discriminability (Valentine, Lewis & Hills, 2016). To test the Face Space Model predictions, participants (N= 46 African, N = 45 Caucasian, N = 40 Chinese) judged the similarity of African, Caucasian, and Chinese faces by selecting the most similar faces to query faces from a 3x3 grid. Using these judgements, we generated multidimensional face spaces (i.e., psychological embedding) for same- and cross-race faces using the PsiZ platform (https://psiz.org). Contrary to the Face Space Model predictions, cross-race faces were judged as more differentiated from one another than same-race faces. However, results were less inconsistent when analyzed by participant race. Specifically, African participants showed a more differentiated face space for Caucasian faces, Caucasian participants for Chinese faces, and Chinese participants for African faces. Results were not explained by differential levels of racial contact. In a follow-up experiment, these results were replicated with Chinese (N = 35) and Caucasian (N = 25) participants and new stimuli. As a test of CRE, these stimuli were employed in a standard recognition memory paradigm with Caucasian participants (N = 37). We found a robust CRE; Caucasian participants showed better recognition of same-race faces than cross-race faces, notwithstanding these faces were not more differentiated in perceptual similarity judgements. The diverging results suggest that the Face Space Model cannot fully account for the CRE, suggesting that race-based differences in perceptual similarity cannot solely drive differences in recognizing same-race and cross-race faces.

This work was supported by an NSERC Discovery Grant awarded to James W. Tanaka (34665-58400)

23.413 DISTINCT MECHANISMS OF CONFIGURAL FACE PROCESSING BASED ON THE PREFERRED FIRST FIXATION LOCATION ON THE FACE

Puneeth N. Chakravarthula[1] (puneeth@wustl.edu), Ansh K. Soni, Miguel P. Eckstein; [1]University of California, Santa Barbara, [2]University of California, Santa Barbara, [3]University of California, Santa Barbara

Accurate face recognition relies on discerning subtle facial feature variations. The visual system uses specialized detectors that balance position invariance and specificity to encode facial shapes and positions. Recently, we found that individuals who prefer to initially fixate near the eyes (upper-lookers) have more position-invariant neural face codes than those who fixate lower (lower-lookers). How do these groups achieve efficient face recognition despite neural code differences? To investigate this, we tested 11 upper- and 11 lower-lookers' ability to recognize faces varying subtly in shape (featural faces) or position of facial features (configural faces). Shape and configural feature sets were difficulty-matched. During a 5-alternative forced-choice recognition task, observers maintained their gaze either at the eyes or near the mouth. Further, we also tested their sensitivity to variations in feature shape and position, comparing efficiency at integrating feature variations across the face under separate single-feature conditions (eyes-only, nose-only, mouth-only). Both groups recognized featural faces equally well across fixation locations. However, lower lookers could only recognize configural faces accurately when fixating the mouth, while upper lookers did equally well, irrespective of their gaze position (looker type × fixation location $F(1,21) = 19.9$, $p<0.001$). In single-feature conditions, upper-lookers were more efficient at extracting eye position (but not shape) information than lower-lookers ($t(18) = 4.3$, $p < 0.005$). No differences in featural or configural processing efficiency were found for the nose or mouth. Finally, using an integration index (a measure of the relative efficiency for processing whole faces vs. parts), we found that upper-lookers efficiently integrate shape and configural information for both fixation locations, while lower-lookers better integrated configural information when fixating the mouth compared to the eyes. Taken together our findings suggest that upper- and lower-lookers rely on distinct mechanisms for configural face processing.

23.414 INDIVIDUAL DIFFERENCES IN EYE MOVEMENTS TO FACES ARE STABLE BUT MALLEABLE

William G. Hayward[1], Nianzeng Zhong[1]; [1]Lingnan University

Although all faces have a highly similar spatial structure, previous studies have found individual differences in eye-movement patterns, showing that some people prefer to scan the upper region of faces (e.g. the eyes) whereas others prefer to look at the lower region of faces (e.g. the nose or mouth). An unresolved question is whether these different fixation patterns are sensitive to the information that is encoded about a face. In this study, we explored whether the facial information available for encoding would affect idiosyncratic eye movement patterns. Specifically, two groups of participants (an upper-focused group and a lower-focused group) performed two learning/recognition tasks. In one task they learned intact faces and in the other task they learned scrambled faces; in both tasks they were then given an old/new test for the studied faces but always in the intact format. We expected that participants would primarily fixate on the eyes of scrambled faces during the study phase, and we predicted that this would lead to a more upper-focused eye-movement strategy for intact faces at test. We found that following both study tasks, the upper-focused group continued to show a more upper-focused pattern than the lower-focused group, suggesting that individuals' looking preferences in the tasks are relatively consistent. In addition, both upper- and lower-focused groups used a more upper-focused pattern when they were learning scrambled faces then learning intact faces, suggesting that these stable idiosyncratic fixation patterns were nonetheless sensitive to encoded information about a face. Taken as a whole, these results show that an observer's fixation patterns when viewing a face are the product of an interplay between stable individual differences and context-specific task optimization.

This work was supported by a grant from the Hong Kong Research Grants Council (LU13605523) to William G. Hayward.

23.415 SOCIAL CORRELATES OF FAMILIAR AND UNFAMILIAR FACE RECOGNITION FOR CHINESE FACES IN AN EAST ASIAN CULTURE

Dana L. Walker[1], William G. Hayward[1]; [1]Lingnan University, Hong Kong SAR

Face recognition is presumed to be an important skill for everyday social interaction. Theoretically, it has been assumed that accurate and efficient recognition of the face may be a social asset for building and maintaining social relationships. However, there is minimal empirical support for the social correlates of individual differences in face recognition. Recent research has found a focused link between face recognition ability and the quality of social connections, rather than the sheer quantity (Engfors et al., 2024). However, this conclusion was based on a Caucasian sample from a highly individualistic country. An interesting question remains whether these findings generalise to other ethnicities, and especially a primarily collectivist culture. Social connections are central in collectivist societies with more rigid social hierarchies when compared to more individualistic societies. Therefore, we aimed to replicate and extend the research of Engfors and colleagues by investigating multiple social correlates of both unfamiliar and familiar face recognition in an East Asian sample. A newly developed naturalistic face familiarity task and the Chinese version of the Cambridge Face Memory Test were administered alongside measures of social network size, social bonds, personality traits, social interaction indices, and social anxiety traits. Consistent with previous findings, quality of social connections was a correlate of familiar face recognition, but not sheer quantity nor extraversion-related personality traits. Interestingly, and in contrast to Engfors and colleagues' findings, familiar face recognition was associated with autistic-like personality traits (social and non-verbal communication dimensions) as well as social anxiety traits. Furthermore, there was minimal evidence for the social correlates of unfamiliar face recognition. Therefore, the key link between the ability to recognise familiar faces and social bonds, rather than may be particularly important to more collectivist cultures.

This work was supported by a grant from the Hong Kong Research Grants Council (LU13605523) to William G. Hayward.

23.416 ENHANCING FACE PERCEPTION RESEARCH WITH THE CHINESE FACE AND BODY DATASET (CFBD)

Ying Hu[1,2] (huy@psych.ac.cn), Ruyu Pan[1,2], Yaqi Xiao[1,2], Zihan Zhu[1,2], Geraldine Jeckeln[3], Xiaolan Fu[4]; [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, [4]School of Psychology, Shanghai Jiaotong University, Shanghai 200030, China

Face perception research typically focuses on identity differences, primarily using White subjects. This approach often neglects how appearances can change across images and the diversity of other racial groups. We aim to develop a dataset to address these gaps and to assess the effectiveness of the dataset in facilitating the exploration of within-model variability under a range of conditions. The Chinese Face and Body Dataset (CFBD) offers a publicly accessible collection of 2,195 images from 117 Asian models. This dataset includes both standardized lab photos and personal photos taken in natural settings at various times. Each image was manually annotated for eight objective attributes: facial expressions, face views, body views, postures, photo types, environments, image quality, and accessories. Participants rated cropped face images from the CFBD for two social traits (attractiveness, trustworthiness) and one identity-relevant trait (distinctiveness). Attractiveness and distinctiveness showed higher within-model variability than between-model variability, as indicated by Wilcoxon rank-sum tests ($W = 5733$, $p = .022$ for attractiveness; $W = 7137$, $p < .001$ for distinctiveness), whereas trustworthiness did not ($W = 11700$, $p = .547$). Personal photos received higher ratings than lab photos for attractiveness ($t(88) = 9.676$, $p < .001$), distinctiveness ($t(88) = 9.618$, $p < .001$), and trustworthiness ($t(88) = 6.993$, $p < .001$), underscoring the impact of type of photos on face perceptions. These findings suggest that perceived attractiveness and distinctiveness are more dependent on cues that vary across images, whereas trustworthiness is relatively dependent on cues that remain consistent. By enabling the exploration of within-model variability under diverse conditions, the CFBD significantly enhances face perception research, particularly for a previously underrepresented group in face perception.

National Natural Science Foundation of China (NSFC) grant No. 32200866 (YH). National Natural Science Foundation of China (NSFC) grant No. 62276259 (YH).

23.417 EXPLORING FACIAL DISTINCTIVENESS THROUGH DEEP LEARNING: INSIGHTS ACROSS IMAGE FORMATS AND GENDER

Artem Pilzak[1] (apilz042@uottawa.ca), Arda Erbayav[2], Alice O'Toole[3], Isabelle Boutet[4]; [1]University of Ottawa, [2]School of Psychology, [3]SCOPE Lab

Introduction: Facial recognition relies heavily on distinctiveness, with atypical faces being more recognizable than typical faces (e.g., Light et al., 1979). Objectively measurable face spaces are available from deep convolutional neural networks (DCNNs), which are highly accurate at face recognition. At VSS2024, we reported that FaceNet (Schroff et al., 2015, 2018), a pre-trained DCNN, can provide a model of human ratings of distinctiveness. Here, we asked (i) whether FaceNet can effectively quantify facial distinctiveness across variations in illumination and viewing angle, and (ii) whether we can replicate findings reported at VSS2024 with another face database and across image formats. Methods: We used 64 male and 64 female identities from the FIE database (de Oliveira Junior & Thomaz, 2006). For each identity, three images were tested: half-profile; and front-profiles lighter illumination, and front-profile darker illumination. We computed distinctiveness scores for each identity and for each image format based on the total cosine distance between embeddings. Results: DCNN-derived distinctiveness scores correlated strongly across image formats (r=0.96 between half-profile and front-profile lighter illumination; r=0.88 between half-profile and front-profile darker illumination; and r=0.91 between the two front-profiles lighter vs darker illumination; all p < 0.001), indicating that DCNN embeddings robustly capture invariant aspects of face typicality. DCNN-derived distinctiveness scores averaged across formats correlated with human ratings of commonality (r=0.31, p=0.01) and typicality (r=0.27, p = 0.03) for male faces, and with human ratings of memorability (r=0.34, p<0.01) and sociability (r=0.29, p=0.02) for female faces. Conclusion: FaceNet effectively models facial distinctiveness across variable image conditions and correlates meaningfully with human-rated aspects of distinctiveness. This underscores the utility of DCNN embeddings for quantifying distinctiveness in studies of face recognition. Further analyses will explore associations between DCNN-derived distinctiveness scores and human ratings for each format tested.

Supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (2022-03998) to IB.

23.418 NO EVIDENCE OF A DUNNING-KRUGER EFFECT FOR FACE RECOGNITION ABILITY: CONVERGING EVIDENCE FROM GLOBAL AND LOCAL METACOGNITION

Kayla Kusel[1] (kjk6km@virginia.edu), Hanxinyi He[1,2], Leah Kirsch[1], Jorge Morales[3], Alison Campbell[1,2,4], Jeremy Wilmer[5], Laura Germine[6,], Sarah Bate[7], Joseph DeGutis[1,4,6]; [1]Boston Attention Learning Lab, VA Boston Healthcare System, [2]Department of Psychological and Brain Sciences, Boston University, Boston MA, [3]Department of Psychology, Northeastern University, Boston MA, [4]Department of Psychiatry, Boston University Chobanian and Avedisian School of Medicine, Boston MA, [5]Department of Psychology, Wellesley College, Wellesley MA, [6]Department of Psychiatry, Harvard Medical School, Boston MA, [7]Institute for Technology and Psychiatry, McLean Hospital, Belmont MA, [8]Department of Psychology, Bournemouth University, Poole UK

One important aspect of metacognition is awareness of one's own cognitive abilities or performance. Lower-performers often overestimate their ability while higher-performers underestimate their ability, but it is unclear whether this 'Dunning-Kruger effect' is domain-specific or universal. To date, no studies have thoroughly examined whether face recognition ability, a specific ability where individuals have moderate insight, shows a Dunning-Kruger effect. Using a web-based sample (N=3,590), we first examined global metacognition, awareness of overall ability, by computing metacognitive sensitivity (correlation) and bias (difference score) using self-reported (Cambridge Face Memory Questionnaire-CFMQ) and objective (Cambridge Face Memory Test-CFMT) face recognition assessments. We found no evidence for differences in global metacognitive sensitivity (r's from lowest-to-highest ability=.28/.32/.32/.33/.30/.25) or bias across abilities (p's>.54). We next examined whether local, trial-to-trial, metacognition (sensitivity: m-ratio, bias: type-2 criterion) differed across face recognition ability by comparing 94 developmental prosopagnosics, 138 controls, and 18 super-recognizers performing an Old/New face recognition task with confidence ratings. For overall performance, super-recognizers outperformed controls, who outperformed prosopagnosics. Compared to controls, prosopagnosics were also more biased to say "new," while super-recognizers were more biased to say "old." Notably, when examining m-ratio we found no significant sensitivity differences between prosopagnosics (M=.98, SD=.72), controls (M=.91, SD=.74), or super-recognizers (M=.93, SD=.48). When examining response-specific type-2 criteria, no group differences were observed when subjects classified a face as "new." However, significant differences emerged in confidence-rating patterns when subjects responded "old" (DPs: M=-1.34, SD=.41, controls: M=-1.0, SD=.44, super-recognizers: M=-.63, SD=.58, p's<.002). In other words, prosopagnosics were more conservative when making high-confidence judgments whereas super-recognizers were more liberal. Together, these findings are inconsistent with a Dunning-Kruger effect, but rather reveal that face recognition ability affects confidence biases when encountering previously studied faces.

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

23.419 ITRAK: USING A NOVEL APPROACH TO SHOW THAT PREFERRED FIXATION LOCATION REMAINS STABLE IN ADULTS

Siobhan McCourt[1], Matt Peterson, Miguel Eckstein[2], Brad Duchaine[1]; [1]Dartmouth, [2]University of California, Santa Barbara

Individuals' first eye movements to faces are consistently directed to a preferred fixation location (PFL). Observers show substantial individual differences in the vertical position of their PFL. Each individual also has an optimal fixation location (OFL) on the face at which performance is best and then rapidly declines when fixations deviate from it (retinotopic tuning). Most people's PFLs and OFLs are tightly linked (Peterson & Eckstein, 2013), and these locations might be expected to remain stable across a lifetime, with participants consistently preferring to fixate on their optimal position. Two findings, however, suggest these positions may migrate down the face. PFLs were higher in younger adults than older adults in a small sample (Peterson et al., 2019), and discrimination of the eyes declines as adults age, whereas discrimination of the mouth remains stable (Fry et al., 2023). Here, we investigate PFLs' shift down the face with age, and the relationship between PFLs and retinotopic face tuning. We utilized iTrak, a novel method for determining single fixation locations in internet-based studies. iTrak briefly presents a target stimulus that is immediately followed by a 27x27 grid of bigrams for 200ms. Participants must respond to the target stimulus and report the bigram they saw, with the bigram indicating fixation location. Participants completed two sex discrimination tasks: a free-fixation task to determine PFLs and a forced-fixation task to estimate retinotopic tuning curves. Results from the free-fixation task (N=43, age range=18-64) showed no correlation between age and vertical PFL r=-.05, p=.73). The forced-fixation task (N=19) revealed the expected negative correlation between distance from PFL and sex discrimination performance (r=-.52, p=.02). These findings indicate that PFLs remain stable with age and the match between PFLs and OFLs demonstrates that iTrak effectively collects eye position data.

Face and Body Perception: Neural

Saturday, May 17, 8:30 am – 12:30 pm, Pavilion

23.420 NEURAL MECHANISMS OF FACE ENCODING AND RECOGNITION IN DEVELOPMENTAL PROSOPAGNOSIA AND HEALTHY CONTROLS

Alison Campbell[1,2], Michael Esterman[1,2,3], Joseph DeGutis[1,4]; [1]Boston Attention and Learning Laboratory, VA Boston Healthcare System, [2]Boston University Chobanian and Avedisian School of Medicine, [3]National Center for PTSD, VA Boston Healthcare System, [4]Harvard Medical School

Developmental prosopagnosia (DP) has the potential to reveal neural mechanisms critical for face recognition, yet its neural basis remains unresolved. Although DP is characterized by an impaired ability to learn and recognize faces, previous research has relied almost exclusively on category localizers to examine activation to faces during passive viewing. By contrast, there is a near complete lack of research on neural processing when there is an explicit demand to encode and recognize faces. One possibility is that the neural regions that normally support face encoding and recognition are dysfunctional in DPs, leaving them to engage regions less or to rely on fewer or alternative regions. To investigate this question, we used event-related fMRI to compare DPs and neurotypical controls while they completed an old/new face recognition task. Notably, performance was equated between groups by reducing the number of study faces for DPs (8 blocks of 10) vs. controls (2 blocks of 40). Prior to the task of interest, participants performed a face/eye region/object localizer and we individually identified regions selective to faces (Faces>Objects) and eyes (Eyes>Objects) in the occipito-temporal (OFA, FFA), prefrontal (IFG), and medial temporal lobe (hippocampus, amygdala). In the face learning paradigm, participants alternated between blocks of study and test in the scanner and a subsequent memory analysis was used to compare successful encoding (subsequent hits > misses during study) and successful recognition (hits > misses during test) in the localized ROIs. Preliminary results show commonly reported group differences in face-selectivity, as well as a novel finding of reduced selectivity for the eye region in DPs. We also found that neural responses during successful encoding and recognition engaged similar medial temporal lobe and face-selective regions. This suggests that, during face learning and recognition, neural differences in DP are more quantitative rather than qualitative.

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

23.421 FACE-TRAINED DEEP NEURAL NETWORK SHOWS HUMAN-LIKE SHAPE BIAS IN FACE SIMILARITY JUDGMENTS

Virginia E. Strehle[1], Frank Tong[1]; [1]Vanderbilt University

Deep neural networks (DNNs) trained for face recognition have surpassed human accuracy (Parde et al., 2023; Phillips et al., 2018). However, there is mixed evidence regarding how sensitive these models are to face shape (Abudarham et al., 2019; Strehle et al., 2024). Here, we asked if a face-trained DNN is truly sensitive to face shape and whether it resembles human perception by leveraging a 3D morphable model of face appearance (Paysan et al., 2009) to directly manipulate face shape and texture (Jozwik et al., 2022; Yildirim et al., 2020). We generated multiple trios of faces to be compared, which consisted of one target face and two alternate faces that differed in shape or texture to various degrees. Half of the trials presented the target and alternate faces from the front view, while the other half presented the alternates offset by +/-22.5° from the front view target. Human observers were asked to report which of the alternate faces was most similar to the target. For comparison, we extracted responses from the final fully connected layer of a face-trained Inception-ResNet-V1 network (Szegedy et al., 2017) and measured the Pearson correlational similarity between responses to each target face and its two alternates. Both humans and the DNN showed an overall shape bias in their face similarity judgments in the front view condition (Human: 70.1% of trials; DNN: 60.4% of trials) and also in the view-offset condition (Human: 60.07% of trials; DNN: 58.3% of trials). Moreover, human and DNN shape biases were positively related for both front-view and offset-view conditions (Pearson r = 0.51and r = 0.46, respectively). Our results indicate that humans exhibit an overall shape bias in their face similarity judgments, DNNs exhibit a similar bias, and more important, the face-trained DNN showed a significant correspondence with human judgments of face similarity.

This research was supported by NEI grants R01EY035157 to FT and P30EY008126 to the Vanderbilt Vision Research Center.

23.422 MODELING ACTION-PERCEPTION COUPLING WITH RECIPROCALLY CONNECTED NEURAL FIELDS

Xinrui Jiang[1], Martin A. Giese[1]; [1]CIN & HIH, University Clinics Tübingen

Visual perception is modulated by various contextual factors, including self-motion. The existence of the mirror neuron system highlights a strong connection between action observation and execution. Yet the mechanisms underlying action-perception coupling remain incompletely understood. METHODS: We developed a hierarchical neural model that represents perceived and executed actions by recurrent neural networks (neural fields) which are reciprocally connected to investigate interactions between visual and motor regions. Our model includes four interconnected parts: a visual pathway, neural fields representing perceived actions and motor plans, and a highly simplified motor pathway. The visual pathway recognizes body shapes based on a pre-trained deep neural network. The neural fields representing visually perceived shape sequences and motor programs include recurrent interactions that result in sequence selectivity or support autonomous traveling pulse solutions. The interaction between these fields and fields representing different types of actions is designed in a way that results in inhibition between different actions, and in mutual excitation between temporally coherent perceived and executed actions. The motor pathway generates image sequences of executed actions from the activation patterns in the motor neural field. We validated the model by exploiting hand

movement image sequences, varying the temporal or movement-type congruence between executed and visually perceived actions. RESULTS: In congruent conditions, synchronized visual inputs and action execution enhanced visual and motor neural field activation, consistent with prior psychophysical findings. Temporal incongruent conditions revealed visual neural field inhibition when delays between visual cues and motion exceeded a critical threshold. Notably, our model aligned with recent fMRI findings that presenting congruent visual stimuli during movement suppressed the activation of units encoding incongruent movement patterns. CONCLUSIONS: Our model successfully captured action-perception interaction across varying congruence conditions. The model makes specific predictions regarding excitatory and inhibitory interactions between visual and motor representations of actions dependent on task conditions.

The work was funded by ERC 2019-SyG-RELEVANCE-856495. The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Xinrui Jiang.

23.423 EXPLORING SEMANTIC AND VISUAL INFORMATION IN FACE PERCEPTION AND SELF-PERCEPTION WITH DEEP NEURAL NETWORKS

Arijit De[1], Shao Feng Liu[1], Kinkini Monaragala[1], Adrian Nestor[1]; [1]University of Toronto

Extensive work has evaluated the contributions of visual and semantic information to face perception, with recent efforts leveraging deep neural networks. Building on this body of work, the present study evaluates the robustness and effectiveness of various neural network models in capturing these contributions. To this end, female White participants (n = 40) rated the pairwise similarity of unfamiliar and familiar (i.e., famous) faces, including their own faces. The stimuli comprised female White young adult faces with neutral expressions. In addition, participants rated all faces for attractiveness and familiarity. Regarding semantic information, a sentence generative pre-trained transformer (SGPT) (Muennighoff, 2022) reliably accounted for relevant variance in the behavioral data. Its explanatory power, as expected, was modulated by face familiarity and depended on the source of information (e.g., celebrity descriptions provided by AI conversational agents were more effective than Wikipedia entries). Regarding visual information, discriminative models (e.g., ArcFace; Deng et al., 2019) and generative models (e.g., StyleGAN; Karras et al., 2020) trained with face images provided complementary and overlapping contributions to explaining the data. Further, we found that explanatory power varied as a function of training set and architecture (e.g., StyleGAN2 outperformed StyleGAN3 in this respect). Last, StyleGAN2's explanatory power was harnessed to map behavioral data into its latent space. Then, we used its generator to synthesize hyper-realistic approximations of unfamiliar and familiar face percepts, including the participants' own faces. These findings demonstrate the utility of combining semantic and visual models to study face perception and highlight the potential of generative networks to recover visual representations. Further, this approach provides a novel framework for exploring the cognitive basis of self-perception.

Natural Sciences and Engineering Research Council of Canada

23.424 HYPERFAMILIARITY FOR FACES ENHANCES FUNCTIONAL CONNECTIVITY BETWEEN VISUAL AND NON-VISUAL REGIONS OF THE BRAIN DURING NATURAL VIEWING

Kira N. Noad[1], Marie-Luise Kieseler[2], David M. Watson[1], Brad Duchaine[2], Timothy J. Andrews[1]; [1]University of York, [2]Dartmouth College

Hyperfamiliarity for faces is a rare condition characterised by an abnormal sense of familiarity for unknown faces. Disorders of face recognition can provide insights into the brain regions involved in the neural representation of familiar faces. Here, we describe a 49-year-old woman (JP) who began experiencing hyperfamiliarity after a severe migraine. She reports having a strong sense of familiarity for faces of people that she does not know, which was not evident before the onset of the disorder. To determine the neural correlates of hyperfamiliarity, we compared the neural responses of JP with neurotypical participants. Structural MRI scans failed to reveal any discernible abnormalities. We then measured fMRI responses in the core face-selective regions to images of unfamiliar faces. JP showed a normal pattern of face-selective responses in the core face regions of the visual brain. To explore the neural response to faces in a more naturalistic setting, we recorded fMRI responses during movie watching. JP watched a movie composed of clips from the TV series Game of Thrones (GoT), which she had not seen before. We measured the response across different regions of the brain in JP and neurotypical participants who were either familiar or unfamiliar with GoT. We found neural responses in the medial temporal lobe of JP were more like familiar participants than unfamiliar participants. We also found that functional connectivity between core face regions and the medial temporal lobe in JP was more like familiar participants than unfamiliar participants. Together, these results demonstrate that hyperfamiliarity for faces manifests in both visual and non-visual parts of the brain, and that connectivity between these regions may play a critical role in the recognition of familiar faces during natural viewing.

23.425 FACE-SELECTIVE BRAIN REGIONS SHARE THE VISUAL-FIELD ANISOTROPIES OF EARLY VISUAL CORTEX

John A. Greenwood[1] (john.greenwood@ucl.ac.uk), Anisa Y. Morsi[1], Hugo T. Chow-Wing-Bom[1], D. Samuel Schwarzkopf[1,2], Valérie Goffaux[3], Tessa M. Dekker[1]; [1]University College London, [2]University of Auckland, [3]Université Catholique de Louvain

Face recognition has long been considered "special", involving dedicated brain regions with selectivity for faces invariant of their visual-field location. Recent findings show that the location of faces in our visual field does however alter their perception – face recognition has higher resolution along the horizontal vs. vertical meridian and in the lower vs. upper visual field (Morsi et al, 2024), matching the variations of low-level vision. We sought to determine whether these performance anisotropies are driven solely by variations in early visual cortex, or by similar variations in the spatial properties of face-selective brain regions. Using fMRI, retinotopic mapping and population receptive field (pRF) analyses were conducted with large-field bars (±21° eccentricity) of either upright or inverted faces in both early (V1-V3) and face-selective cortical regions (OFA, pFus, mFus). The size and number of pRFs, and associated visual field coverage (the sum of all pRFs), was estimated in wedges around the visual field. Though pRF sizes increased with eccentricity in all areas, they did not vary reliably with polar angle in either early visual cortex or face-selective regions. Both early cortex and face-selective areas nonetheless showed both a greater number of pRFs and a concomitant increase in visual-field coverage along the horizontal vs. vertical meridian and in the lower vs. upper field. These variations in visual-field sampling could therefore drive the variations observed in face-recognition abilities. We further show that pRF numbers (but not pRF size or coverage) were higher for upright than inverted faces in mFus, providing a link between spatial variations and the perceptual advantage for upright faces. The shared pattern of these variations in visual-field sampling supports recent proposals for “visuospatial coding”, whereby the spatial selectivity of higher-level areas is built upon the selectivity of early visual cortex, even for specialised processes like face recognition.

Supported by the UK Biotechnology and Biological Sciences Research Council [BB/J014567/1].

23.426 DECODING FACES UNDER DISCONTINUOUS FLASH SUPPRESSION

Shao-Min (Sean) Hung[1], Yen-Ju Feng[2], Po-Jang (Brown) Hsieh[2]; [1]Waseda Institute for Advanced Study, Waseda University, [2]Department of Psychology, National Taiwan University

Although psychophysical experiments have revealed a wide range of behavioral effects driven by invisible stimuli, the underlying neural machinery remains largely unknown. Previous attempts at localizing subliminal, especially high-level, signals suffered from low signal-to-noise ratios due to weak raw signals and/or active suppression. In this study, we implemented dynamic faces, a novel variant of interocular suppression (discontinuous flash suppression), as well as regions-of-interest (ROI) analyses to enhance the stimulus and statistical power. A total of forty-three participants was included based on a pilot study with n = 10. At the whole-brain level, our univariate analysis first replicated a typical widespread activation by conscious faces, which was absent by unconscious faces. However, with multivariate analysis, the classification was successful in the occipital-temporal regions between static/dynamic faces and their counterpart scenes. ROI analysis further showed that bilateral FFA and OFA successfully differentiated unconscious dynamic faces and scenes. Together, these results indicate in-depth processing of unconscious faces in the ventral visual stream. Our successful localization of unconscious high-level signals highlights the importance of the selection of stimulus, paradigm, and analysis.

The Ministry of Education (MOE), Yushan Young Scholar Program (NTU-110V0202), Ministry of Science and Technology (109-2410-H-002-004-MY3), National Taiwan University (111L9A00701), and Waseda Grant Program for Promotion of International Joint Research.

23.427 WHEN THE EYES ARE NO LONGER DIAGNOSTIC: A CLEAR DEMONSTRATION OF THE N170 AS AUTOMATIC AND GOAL-ORIENTED INFORMATION PROCESSING

Chanelle Demeule[1] (chandem09@gmail.com), Anthony Proulx[1], Justin Duncan[2], Caroline Blais[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais, [2]Ottawa University

Previous studies suggest that early perceptual brain activity in face recognition (N170) is linked both to the processing of the eye region and to the processing of diagnostic information (i.e., task-relevant; Schyns et al., 2007; Smith et al., 2004). However, in such previous studies, eye information and diagnostic information were confounded - that is, the eyes conveyed diagnostic information. The present study aimed to disentangle the two by creating facial stimuli where the eye region was completely identical between the two versions of the same identity (smiling or neutral). This will allow us to precisely measure the role of the eye region in the occurrence of the N170. We recorded brain activity using EEG for 29 participants (12 men, 17 women) while they categorized the expressivity of faces. On each trial, Faces were revealed only partially through randomly located Gaussian apertures (“Bubbles”; Gosselin & Schyns, 2001). Using time-resolved classification image techniques, we were then able to reveal image pixels that were associated with brain activity at PO8, from -200ms to 800ms post-stimulus, on a subject basis. A pixel test (p<.05; Stat4Ci Toolbox; Chauvin et al., 2005) revealed a significant association between N170 amplitude and the contralateral eye, the ipsilateral eye and the mouth. These results add credibility to the theory that the N170 reflects both a sensitivity to the eye region, independent of task demands, and the integration of goal-oriented information.

The present study is supported by the Natural Sciences and Engineering Research Council of Canada to Daniel Fiset (#RGPIN-2022-04350)

23.428 FUNCTIONAL ORGANIZATION OF VISUAL RESPONSES IN THE ANTERIOR TEMPORAL LOBE

Ben Deen[1], Chencheng Shen[1]; [1]Tulane University

Object perception in primates is thought to be supported by a collection of brain areas within the ventral visual stream, with a hierarchical organization along a posterior-to-anterior axis. Decades of neuroimaging research has described the organization of category-sensitive visual areas within human posterior ventral temporal cortex. However, category-sensitive responses within with anterior temporal lobe (ATL) remain poorly understood due to signal dropout in fMRI data resulting from magnetic susceptibility artifacts. Here, we investigate the functional organization of visual responses in the ATL in healthy young adult data from the Human Connectome Project, leveraging a large sample size to overcome limits in data quality. We selected a subset of N = 864 participants based on quality control measures and a head motion cutoff (mean framewise displacement of .3mm). Cytoarchitectonic subregions of the ATL, including the temporal pole (TP) and perirhinal cortex (PR), were defined using a probabilistic atlas of hand-drawn regions from a separate set of human participants. Category-sensitive functional regions-of-interest were defined in individual participants by comparing responses to face, body, tool, and scene conditions, and response magnitudes and functional connectivity were assessed in independent data. We identify three distinct functional responses within the ATL: a socially-selective response within TP, a face-selective response within PR, and a non-face-object-selective response within PR. These three regions showed distinct patterns of functional connectivity with areas outside the ATL, each showing preferential connections parts of cortex with similar category preferences. These results demonstrate the presence of multiple distinct visual areas within the ATL, and show that the principle of category-sensitive organization extends to the top of the ventral visual processing hierarchy.

23.429 CONTRIBUTIONS OF PRIMATE SUPERIOR COLLICULUS TO FACE PROCESSING IN TEMPORAL CORTEX

Gongchen Yu[1], Leor Katz[1], Richard Krauzlis[1]; [1]Laboratory of Sensorimotor Research, National Eye Institute, NIH

We recently reported that neurons in the macaque superior colliculus (SC) exhibit a preference for face stimuli at extremely short latencies. This finding raises the question of whether and how SC contributes to the well-established selectivity for faces in temporal cortex. To address this question, we used multichannel probes to record activity from face-selective neurons in the middle fundus face patch (MF) of the temporal cortex, both before and during SC inactivation. Visual object stimuli—including faces, bodies, and human-made objects—were presented either within the visual field location affected by SC inactivation (5° away from fixation, contralateral to the recorded hemisphere) or outside it (5° ipsilateral). Our results show that SC inactivation significantly reduced the selectivity of face-selective neurons in MF. This effect was predominantly observed for faces presented in the visual field affected by SC inactivation, and was primarily centered on the initial visual response period following stimulus onset (~70–140ms). In addition to face-selective neurons, our dataset included a population of neurons selective for bodies. Preliminary analyses indicate that SC inactivation reduced the selectivity of body-selective neurons as well, albeit not as strongly as observed for face-selective neurons. Our findings demonstrate that the SC causally contributes to face processing in the temporal cortex, particularly at short visual latencies. Thus, SC activity might be critical for prioritizing the processing of objects—especially faces—and for facilitating subsequent higher-order processing in cortex.

23.430 NEURODYNAMICS OF ACTIVE FACE PERCEPTION DURING FREE VIEWING WITH EYE TRACKING AND INTRACRANIAL EEG

Casey Becker[1], Witold Lipski[1], Mary-Kate Richey[1], Arish Alreja[1], Jorge Gonzalez-Martinez[1], Taylor Abel[1], Avniel Ghuman; [1]University of Pittsburgh

Since at least the seminal studies of Yarbus, we have known that people look at faces in stereotyped ways. Furthermore, disruptions of these gaze patterns are a hallmark of face processing deficits. Nonetheless, with a few notable exceptions, nearly all we know about the neurodynamics of face perception relies on paradigms where faces are presented abruptly at central fixation, overlooking the dynamic role of eye movements in natural vision. Here, we characterize the cortical and subcortical neurodynamics of face perception during unrestrained sequential fixations around a face using eye tracking and intracranial EEG in human participants. Participants free viewed static images of happy, angry, fearful, and neutral facial expressions for 1500 milliseconds. Behaviorally we found an initial freezing of eye movements upon stimulus onset, followed by a first saccade around 200ms consistently across participants. This timing aligns with the N170 timeframe and give rise to the question of how much of early face-related potentials are associated with this oculomotor freezing and the planning of the first saccade. Comparing responses across multiple fixations (fixation-related potentials) to traditional event-related potentials (ERPs) elicited by sudden onset faces uncovered large differences in the neurodynamics between the onset response and neural responses to subsequent fixations. Using Multi-Temporal Pattern Analysis (MTPA) we characterize distinct cortical and subcortical responses to different facial features, including presaccadic activity related to the facial feature participants saccade to next. These neurodynamics highlight the neural foundations of the dynamic active sensing process in face perception during free viewing.

23.431 MOTOR MOVEMENTS AND REACTION TIME AS WINDOWS INTO CORTICAL SELECTIVITY

Timothée Maniquet[1], Hans Op de Beeck[1]; [1]KU Leuven, [2]Leuven Brain Institute

One of the hallmarks of the human visual system is the functional specialisation of areas in the occipito-temporal cortex (OTC). Category-selective patches of cortex can be found where neurons preferentially respond to a given semantic category. Examples include the fusiform face area (FFA), extrastriate body area (EBA), and parahippocampal place areas (PPA), with selective responses to faces, bodies, and places, respectively. Across clusters of selectivity, super-clusters can be observed, in particular for the distinction between animate and inanimate categories. This animacy representation seems to be continuous, which has been interpreted as preference for categories closer or further away on an intuitive taxonomy scale, or alternatively as a reflection of how bodies and faces are coded as being more or less human-like. Here, we test the latter hypothesis through the combination of behavioural and neuroimaging methods. We recorded participants behaviour as they classify images either as animate or inanimate, or as faces, bodies and scenes, and tried predicting reaction times and gradual movement trajectories from univariate activity from FFA, EBA and PPA in response to the same images. Our results show a nuanced pattern whereby FFA activity is strongly predictive of reaction times and mouse positions, more so than EBA and PPA. While EBA and PPA can both predict increases in reaction times for non body and non scene stimuli, respectively, only FFA activity can accurately predict decreases in reaction times for stimuli rated as more face-like. Additionally, FFA surpasses others in predicting patterns of the movement trajectories during animacy classification. Overall, we find partial support for the face & body account of animacy selectivity in the cortex. We interpret our results in light of theories of OTC organisation.

23.432 ALLOCATION OF ATTENTIONAL RESOURCES TO FACES IS DOMAIN-SENSITIVE AND INDEPENDENT OF FAMILIARITY

Linda H. Lidborg[1], Anna Yue Gao[1], A. Mike Burton[2,3], Holger Wiese[1]; [1]Durham University, [2]University of York, [3]Bond University

Previous research has suggested that attentional capacity limits permit the processing of just one face at a time. However, we have previously shown that the brain is capable of processing two different facial identities, if faces are simultaneously presented in equally relevant spatial locations. Here, we used event-related potentials (ERP) and immediate repetition priming to test whether two facial identities are still processed simultaneously when one face is made more relevant than the other. In Experiment 1, stimuli consisted of both famous and unfamiliar faces. Two facial primes were presented simultaneously:

one prime appeared in a central location in every trial, while an additional, peripheral prime unpredictably appeared to the left or to the right of the central prime. Here, we observed an ERP priming effect with more negative amplitudes at occipito-temporal electrodes for central but not peripheral familiar faces from 300 ms onwards. Similarly, an unfamiliar central prime also did not allow for the processing of a peripheral familiar prime. In Experiment 2, primes consisted of faces or written names of familiar celebrities. The prime locations were identical to Experiment 1; however, while the peripheral prime was always a face, the central prime could be either a face or a name. Here, we observed ERP priming effects for peripheral faces when the central prime showed a name, but not when the central prime showed a face. Together, these results suggest that while it is possible for the brain to process peripheral (less relevant) faces, this only occurs if the central stimulus is not also a face; a centrally presented face captures domain-sensitive attentional resources independent of familiarity and blocks the processing of an additional face.

Funded by Economic and Social Research Council (ESRC, UK) grant ES/X014657/1 to HW and AMB.

Face and Body Perception: Emotion

Saturday, May 17, 8:30 am – 12:30 pm, Pavilion

23.433 LINKING CONCEPTUAL KNOWLEDGE OF EMOTIONS TO VISUAL REPRESENTATIONS OF FACIAL EXPRESSIONS USING A REVERSE CORRELATION APPROACH WITH GENETIC ALGORITHMS

Émilie St-Pierre[1] (stpe06@uqo.ca), Jeanine Ohene-Agyei[2], Arianne Richer[1], Alexis Bellerose[1], Francis Gingras[1,3], Zohair Mharchat[1], Camille Saumure[4], Daniel Fiset[1], Roberto Caldara[4], Caroline Blais[1]; [1]University of Quebec in Outaouais, Canada, [2]University of Toronto, Canada, [3]University of Quebec in Montreal, Canada, [4]University of Fribourg, Switzerland

Facial expressions of emotion are crucial for effective social communication, yet remain subject to perceptual confusion. Previous research using a reverse correlation technique demonstrated that individual conceptual knowledge shapes and predicts visual representations of emotions (Brook & Freeman, 2018). However, reverse correlation methods require many trials and have been used to explore conceptual and visual relationships between pairs of affective states. Combining a genetic algorithm with reverse correlation to manipulate facial features on photorealistic avatars should offer a faster, more efficient, and naturalistic method for generating facial representations (Binetti et al., 2022). To probe this hypothesis, we used this approach to investigate the relationship between participants' conceptual knowledge of six basic emotions and pain and their visual representations of facial expressions. Each participant completed a conceptual task, rating associations between these seven affective states and 45 word or phrase stimuli. This enabled the construction of individual-level similarity matrices, reflecting each participant's unique patterns of conceptual overlap among these affective states. Participants also completed a reverse correlation task to capture their visual representations of facial expressions for the affective states in male and female faces. Perceptual overlap among visual representations was measured using two approaches. First, individual-level similarity matrices were generated based on the degree to which the representation of each expression shared similar activation patterns across 59 facial features. Then, a second set of similarity matrices was generated from ratings by 20 independent observers assessing the perceived intensity of each affective state. Inter-matrix correlation analyses examined the relationship between conceptual and perceptual similarity matrices. Permutation analyses confirmed that these correlations significantly exceeded chance level, providing robust evidence of a link between conceptual and perceptual similarity in facial expression representations. Crucially, our data show that this relationship was characterized by distinct individual differences, revealing a complex interplay between these two factors.

This study is supported by the Canada Research Chair in Cognitive and Social Vision (Caroline Blais, #CRC-2023-00019), NSERC Discovery Grant (Caroline Blais, #RGPIN-2019-06201), and the NSERC Graduate Scholarship Doctoral Program (Émilie St-Pierre, #ES D-590029).

23.434 MAPPING THE IDIOSYNCRATIC RECOGNITION OF FACIAL EXPRESSIONS OF EMOTION

Anita Paparelli[1] (anita.paparelli@unifr.ch), Lisa Stacchi[1], Inês Mares[2], Louise Ewing[3], Marie L. Smith[4], Roberto Caldara[1]; [1]Eye and Brain Mapping Lab, Department of Psychology, Univerisity of Fribourg, Fribourg, Switzerland, [2]William James Center for Research, Ispa – Instituto Universitário, Lisboa, Portugal, [3]School of Psychology, University of East Anglia, Norwich, UK, [4]School of Psychological Sciences, Birkbeck College, University of London, UK

The recognition of facial expressions of emotion (FER) is a critical biological skill of human social cognition. Recently, data-driven work from our laboratory revealed that Western adults sample facial information through idiosyncratic fixation patterns during FER, while maintaining comparable categorization accuracy (Paparelli et al., 2024). Importantly, we also reported that the same observer consistently adopts the same sampling strategy independently of the facial expressions of emotion (FEE). However, the understanding of the underlying factors behind idiosyncratic fixation patterns are yet poorly understood. The current study seeks to provide further insights to this important issue by assessing whether the individual differences observed in sampling strategies during FER relate to differences in information use. To probe this hypothesis, we tested healthy adult Western observers on a FEE-categorization task using first an eye-tracking (ET) and subsequently a Bubbles reverse correlation (RC) paradigm. Both experiments investigated the 6 basic FEEs, in addition to the neutral condition. In the ET paradigm, observers freely viewed images while their eye-movements were recorded. On the other hand, during the Bubbles paradigm, stimuli were presented through masks that revealed different facial features across trials. Our results replicate previous findings and show robust idiosyncratic fixation patterns with minimal FEE-dependency. In contrast, the results from the Bubbles RC paradigm highlight that information use is mostly FEE-specific, while showing less variation across observers. By investigating for the first time information fixation and use at the single subject level, our data show that while individuals sample facial information during FER

differently, they use the same facial features to categorize FEEs. This indicates that FER is achieved through a complex interplay between foveal and parafoveal visual signals integration across fixations. The recognition of FEE can be efficiently achieved by distinct biological idiosyncratic tunings.

AP was supported by the Swiss National Foundation grant n° P000PS_227296 / 1. RC was supported by theSwiss National Foundation grant n° 10001C_201145 / 1.

23.435 THE CONTEXTUAL UNDERSTANDING EFFECT: A SPECIALIZED VISUAL SKILL BASED ON THE IMPLICIT USE OF VISUAL CONTEXT INFORMATION

Jefferson Ortega[1], David Whitney[1]; [1]University of California, Berkeley

Context shapes everything in vision. To recognize an object's lightness, shape, color, or motion, we rely on background information. This reliance is even more critical when interpreting emotions. Facial expressions are inherently ambiguous, often requiring contextual cues for accurate understanding (Barrett et al. 2011; Chen and Whitney 2019). But does using context involve a distinct visual skill? Do some individuals intuitively seek and use contextual clues more effectively than others? We explored this question through an eye-tracking study, examining how individuals use background context to recognize emotions in real time. Participants continuously reported the valence and arousal of an actor in a film under two conditions: a context-only condition (where the actor was blurred out, leaving only the background scene visible) and a ground-truth condition (where the entire scene was visible, including the actor). When observers are prevented from seeing the actor's face (context-only condition), they need to search the background scene context for valuable information. We find that some observers are better at this than others and these skilled observers reveal where, in the background scene context, the informative information is found that facilitates emotion recognition. Surprisingly, we found that the observers who were best able to continuously report emotion in the ground truth condition (where everything was visible) were those who gazed at the same (non-face) locations in the background as the most accurate observers in the context-only condition. The results reveal that specific spatiotemporal moments in the background scene context are highly informative for emotion recognition, but not all observers understand or are able to use this information. This suggests that certain individuals possess a unique, implicit visual skill: the ability to recognize and extract critical contextual information for emotion recognition. Our results highlight significant individual differences in how we integrate context to decode emotions.

Supported in part by the National Institute of Health (grant no. R01CA236793) to D.W and (grant no. 1F99NS141343) to J.O.

23.436 IDENTIFYING DISTRUST FROM FACIAL EXPRESSIONS BY FACIAL ACTION CODING SYSTEM

Gabriel RongYang Lau[1] (gabriel.laury@ntu.edu.sg), Zihao Zhao[1], Shuyi Sun[1], Nicole Zhi Ee Ng[1,2], Bee Chin Ng[2], Hong Xu[1]; [1]School of Social Sciences, Nanyang Technological University, [2]School of Humanities, Nanyang Technological University

Previous research showed that social-behavioural cues can indicate trust. This raises the question of whether we can determine people's trust from their facial expressions. In this study, we used the Facial Action Coding System (FACS) to identify facial Action Units (AUs) associated with trust toward social media news. Thirty-two participants viewed 192 randomized videos of speakers delivering news that varied by display type (monitor, hologram), emotion (neutral, happy, angry), attire (casual, doctor, nurse, police), gender (female, male), and content veracity (ambiguous, fake, real, no news). Participants rated the trustworthiness of the speaker and news while their facial expressions were recorded using a Logitech BRIO Webcam (except one participant for incomplete data). We coded the AUs displayed in each trial and conducted a Wilcoxon rank-sum test (for non-normal distribution and unequal sample size) to compare trust ratings between trials with and without displayed AUs. Participants were more likely to display the following AUs in brows and lip movements when they reported lower trust in both the speaker ($W = 928{,}058$, $p < .001$) and the news content ($W = 625{,}781$, $p < .001$): AU4 (Brow Lowerer), AU23 (Lip Tightener), AU28 (Lip Suck), AU20 (Lip Stretcher), AU1 (Inner Brow Raiser), AU12 (Lip Corner Puller), and AU15 (Lip Corner Depressor). Logistic regression analysis showed that the likelihood of AU display increased when participants viewed angry speakers ($p < 0.001$) sharing ambiguous ($p = 0.022$) or fake news ($p < 0.001$) on monitor display ($p < 0.001$). Our results highlight brow and lip movements as consistent indicators of distrust, influenced by news veracity, display mode, and speaker emotion. These distrust-related AUs may reflect heightened emotional arousal and increased cognitive effort in information processing. Our findings provide insights into the physiological expression of distrust and suggest potential applications for detecting trust using facial AUs.

23.437 DO EMOTIONAL ENSEMBLES SHAPE BEHAVIOR? INVESTIGATING THE ROLE OF AVERAGE EMOTIONAL EXPRESSION IN APPROACH-AVOIDANCE DECISIONS

Eliz Shimshek[1], Marco A. Sama[1], Jonathan S. Cant[1]; [1]University of Toronto Scarborough

The visual system efficiently processes multiple sources of information by leveraging ensemble encoding, the ability to extract statistical summaries (e.g., the average size of circles) from sets of similar objects. This also occurs for high-level stimuli such as faces, and summary statistics such as the average expression in a crowd of faces provides critical cues about the intentions of others. Emotional expressions are pervasive drivers of decision-making, and similarly, affective images, such as the valence of photographs, have been shown to significantly influence approach-avoidance decisions. While ensemble processing has been extensively studied in visual perceptual tasks, its influence on approach-avoidance decisions, a fundamental aspect of human behavior, remains insufficiently explored. To address this, we examined how ensemble processing influences approach-avoidance decisions in real-world scenarios, namely, when deciding whether or not to watch a movie. To investigate this, we examined the relationship between implicit ensemble processing (i.e., passive viewing of ensemble stimuli) and approach-avoidance behavior. Participants viewed ensembles of six faces expressing a positive, negative, or neutral average emotion. Following the presentation of each face ensemble, participants completed an approach-avoidance task using a social decision-making paradigm.

Specifically, participants were presented with a positive, negative or neutral movie poster and then quickly decided whether or not they would prefer to watch the depicted film. We found that average expression did not influence approach-avoidance behaviors towards affective movie posters. Instead, viewing decisions depended only on the affective content of the movie poster, independent of the face ensemble it was paired with. Critically, this research deepens our understanding of the intersection between ensemble processing and decision-making. As a next step, we will test whether explicit ensemble processing influences approach-avoidance decisions, as explicit judgments of an average feature are known to create more precise ensemble representations compared with implicit processing.

This research was funded by an NSERC Discovery Grant to J.S.C.

23.438 FACIAL EMOTION PROCESSING IN NATURALISTIC CONTEXTS

Xiaoxu Fan[1], Abhishek Tripathi[2], Lily Chamakura[1], Kelly Bijanki[1]; [1]Baylor College of Medicine, [2]Rice University

Facial emotion processing has been widely investigated using highly artificial controlled paradigms, yet it remains unclear how we interpret facial emotions in real life where multi-modal and contextual cues are all available and naturally interact. In this study, we analyzed neural data collected from human participants watching an audiovisual film, using AI models to extract 48 dynamic facial emotions and encoding models to evaluate neural representations of these emotional signals. We found that activity in dorsal lateral prefrontal cortex (DLPFC), anterior superior temporal cortex (aSTC) and posterior superior temporal cortex (pSTC) represent facial emotion in naturalistic contexts, however, in our data, children's DLPFC does not encode these emotional features, suggesting a developmental difference in affective processing. Additionally, human voice appears to enhance the representation of facial emotion in pSTC, and alters the encoding of specific emotion categories in pSTC and aSTC. Taken together, our results reveal the coding of 48 facial emotions in ecologically valid settings, leveraging AI-based emotion analysis to advance our understanding of affective processing in the human brain.

23.439 VISUAL EMOTION RECOGNITION OF FACES AND SCENES IN ALEXITHYMIA

Ilya Nudnou[1] (ilya.nudnou@ndsu.edu), Benjamin Balas[1]; [1]North Dakota State University

Alexithymia is a personality trait defined as a difficulty of identifying, describing, and distinguishing between one's feelings. Individuals with high alexithymia symptoms also exhibit deficits in categorizing facial expressions, but retain the ability to discriminate emotional expressions. In three experiments, we presented emotional faces and scenes to participants with varying alexithymia severity to further explore the impact of alexithymia on the visual recognition of emotional content. In experiment 1 (N=98), we presented upright and inverted emotional faces and scenes depicting the emotions of anger, fear, and sadness, and asked participants to categorize these images. In experiment 2 (N=81), the same faces and scenes were presented during an odd-one-out discrimination task. Finally, in experiment 3 (N=73), the same faces and scenes were used in a Stroop-like task to examine interactions between face and scene emotional content. Neither accuracy nor reaction time data correlated with alexithymia severity in experiments 1 and 2. In experiment 3, participants with more severe alexithymia were slower at categorizing fear and sad scenes, and sad faces. Scene categorization accuracy was significantly lower during incongruent trials in experiment 3 compared to experiment 1, implying that emotional faces interfered with emotional scenes during categorization, but not vice-versa. Further, the differences in reaction times between experiment 1 and 3 demonstrate that individuals with alexithymia are still sensitive to the congruence of emotion categories perceived from different stimulus types. Altogether, our results demonstrate that people with alexithymia display difficulties with emotion categorization of face and scene stimuli, suggesting a general deficit in the recognition of emotional content.

23.440 THE INFLUENCE OF OBSERVER AND ACTOR AGE ON THE COMPLEXITY OF LABELS OFFERED IN AN EMOTION PERCEPTION TASK

Andrew Mienaltowski[1], Tashaunda Grimmett[1], Lynnsey Cole[1], Natalie Spiva[1]; [1]Western Kentucky University, Bowling Green, Kentucky USA

Advancing age is often associated with poorer emotion recognition performance in images of static facial expressions. However, weak to no age differences are observed when participants characterize emotion in dynamic facial stimuli. The current study investigated whether age differences in emotion perception for dynamic stimuli may be more prevalent in the complexity of the labels freely offered to describe emotional expressions than in measures of label selection accuracy common to emotion perception tasks. Younger and older adults observed younger, middle-aged, and older adult actors expressing anger, disgust, fear, sadness, and happiness via dynamic video stimuli. Labels offered by participants were coded for measures of lexical complexity, as well as for the relevance of each label to the objective emotion category from which the stimulus was drawn. Participant age had minimal impact on the complexity of the emotion labels offered or on their relevance to the stimulus category. However, the complexity of the participants' labels did vary by the age of the actor expressing emotion. More complex labels were offered for older actors than for younger actors, especially for expressions of anger and sadness. Although this suggests nuance in the perception of anger and sadness in older actors' facial cues, the labels offered by participants were less relevant to the stimuli's objective emotion categories. In addition, participants were less accurate at selecting the correct emotion label for older actors expressing anger and sadness. Taken together, these findings suggest that participants, regardless of their age, struggle more to decode the facial cues of older adults expressing negative emotions than they do those of younger adults, possibly due to a reduction in the emotion signal found in the expressive cues of older adults.

23.441 AGE-RELATED DIFFERENCES IN THE PERCEPTION OF NEGATIVE EMOTIONS INCLUDING PAIN

Alana Medina Disponett[1], Sheila To[1], Andrew Mienaltowski[1]; [1]Western Kentucky University, Bowling Green, Kentucky, USA

Age-related deficits in negative facial emotion recognition are generally weaker in tasks using dynamic emotion expressions relative to those using static expressions. This study extended prior aging research on emotion perception by asking adults ranging from 20 to 79 years of age to observe dynamic negative emotion expressions from numerous emotion categories, including pain, in the context of an emotion identification task. Stimuli were presented in a randomized manner in blocks such that participants were asked to consider every stimulus relative to a single emotion category in a given block. Signal detection measures of sensitivity and response bias were examined for each emotion identification block. Overall, chronological age was only negatively associated with participant sensitivity to sadness. Age did not predict signal sensitivity for the other negative emotions. Additionally, bias in perceiving pain in non-pain stimuli increased with age. This bias was not associated with participant self-reported physical or mental health, nor with the participants' tendencies to interact with people experiencing pain. These findings replicate prior research in which age differences in emotion perception were minimal when dynamic expressions were employed, but also highlight that, as we age, we may display a pain empathy bias where we perceive that others are in pain when they are expressing negative emotion.

23.442 THE ORIGIN OF AFFECT-SPECIFIC NEURAL REPRESENTATIONS OF EMOTIONAL SCENES IN EARLY VISUAL CORTEX

Yujun Chen[1] (yujunchen@ufl.edu), Andreas Keil[1], Mingzhou Ding[1]; [1]UNIVERSITY OF FLORIDA

Recent fMRI studies in human subjects have found affect-specific neural representations of emotional scenes in early visual cortex. The origin of these representations is debated. One group of hypotheses suggests that these representations result from reentrant feedback from anterior emotion-modulating structures (e.g., the amygdala), whereas another group of hypotheses states that sensory cortex, including retinotopic visual cortex, may itself code for the emotional qualities of visual stimuli, without the necessity for feedback processing. We examined this problem by employing a neural encoding model that can generate synthetic fMRI responses to natural images in early visual cortex. The model works by linearly mapping features extracted by convolutional neural networks onto voxel-wise BOLD responses in different visual areas and is trained on the Natural Scenes Dataset. Dividing the images in the International Affective Picture System into three broad categories: pleasant, neutral and unpleasant, we found that in early visual cortex, the neural patterns evoked by the emotional images cannot be decoded from that evoked by the neutral images, in contrast with the findings from recent fMRI studies in human subjects. Because the model-generated responses are free from emotion-modulated reentrant feedback, this finding can be seen as lending support to the reentry hypothesis. Interestingly, when face stimuli from the AffectNet were shown to the model, the neural patterns evoked by emotional faces in early visual cortex can be significantly decoded from that evoked by neutral faces, suggesting that the early visual cortex may contribute differently to the emotional processing of faces versus scenes.

23.443 THE EFFECT OF GENDER ON PAIN EXPRESSIONS AND THEIR RECOGNITION: EXPLORING FACIAL MOVEMENT PATTERNS IN POSED AND SPONTANEOUS DYNAMICS

Arianne Richer[1] (rica30@uqo.ca), Camille Saumure[2], Daniel Fiset[1], Caroline Blais[1]; [1]University of Quebec in Outaouais, [2]University of Fribourg

Although individuals recognize pain facial expressions (PFE) above chance, they often confuse them with expressions of other negative affective states (Kappesser & De Williams, 2002; Roy et al., 2015; Simon et al., 2008). Moreover, pain is less easily recognized in women's faces compared to men's (Riva et al., 2011). At last year's VSS, we presented findings based on stimuli from the Delaware Pain Database (DPD; Mende-Siedlecki et al., 2020), showing that automatic recognition algorithms identified men's posed PFE as more frequently containing features typically associated with pain. Furthermore, when presented to a group of human observers, women's posed PFE were more often perceived as conveying emotions such as fear, sadness, and surprise, despite no greater signal of these emotions being present in the stimuli. This project examines whether similar effects emerge when using dynamic stimuli of both genuine and posed PFE. We used the Pain E-motion Faces Database (PEMF; Fernandes-Magalhaes et al., 2023) containing 55 videos of White young adults (38 women, 17 men) posing PFE and 55 videos of genuine PFE. OpenFace was used to quantify the activation levels of 17 Action Units (AUs). A Principal Component Analysis revealed four significant components. The first component includes AUs typically associated with pain and is significantly more associated with posed than genuine PFE. Additionally, the fourth component is significantly more associated with men's PFE than women's. Ongoing analyses investigate how these differences in terms of expressive signal predict the way they are perceived by external observers. This approach enables us to explore the interplay between encoding and decoding processes in PFE, which may help us understand why women's pain expressions are not recognized as easily as men's.

The present study is supported by the Canada Research Chair in cognitive and social vision to Caroline Blais (#CRC-2023-00019) and by the Canada Graduate Scholarship - Doctoral program of the Natural Sciences and Engineering Research Council of Canada to Arianne Richer (#CGS D - 589787).

23.444 DEVELOPMENT AND VALIDATION OF A NEW DYNAMIC FACIAL EXPRESSION DATABASE

Sophia Lipetzky[1], Susan G. Wardle[1], J. Brendan Ritchie[1], Chris I. Baker[1], Shruti Japee[1]; [1]National Institute of Mental Health

Facial expressions are essential for effective interpersonal communication and social interaction. Most prior research aimed at understanding how humans process the expressions of others has used static images of racially homogenous, highly posed facial expressions as stimuli, which have low real-world relevance. To improve the ecological validity of stimuli used in facial expression research, we developed the Facial and Body Movement Database (FaBMoD) consisting of video recordings of twenty-three racially diverse professional actors making a variety of facial expressions. Each actor made nine different facial expressions oriented in seven different directions. To validate these expression videos, participants (n=10) completed a computer task where they were shown each of the 204 front-facing videos and asked to identify the facial expression. Participants were also asked to rate each video on a scale of 1-9 for intensity, genuineness, and valence. Results indicate good overall correspondence for most videos between the intended expression and the expression perceived by participants. This was especially the case for videos depicting happiness and sadness, while videos intended to depict fear were sometimes perceived as surprise. Intensity ratings for most emotional expressions were reasonably high, with participants predominantly rating videos between 6 and 8 on a scale from 1 (not intense at all) to 9 (very intense). By contrast, videos depicting sadness or neutral expressions were generally rated as less intense. Genuineness ratings showed a similar trend with emotional expressions being rated as quite genuine, while neutral expressions were typically rated as less genuine. This new dynamic facial expression database along with the validation data will be used in our future studies of facial expression processing, as well as shared with other researchers world-wide.

23.445 ANGRY CROWD BIAS IN THE NEWS

Delaney McDonagh[1], Darla Bonagura[2], Timothy Sweeny[1], Gorkem Er[1], Sarah Lamer[2]; [1]University of Denver, [2]University of Tennessee Knoxville

Crowds hold a unique place in perception; people are adept at making rapid and precise judgments about clear, prototypical emotions. Facial expressions, however, are often nuanced, and the judgements people make about these expressions are thus made with an unavoidable degree of uncertainty. When making low-confidence evaluations, perceivers are typically biased to interpret facial expressions as negative, which potentially serves a protective function. When people make judgments about crowds—which are associated with increased potential for threat—negativity bias is even further amplified. While this bias has been observed with tightly controlled stimuli (e.g., faces portraying posed expressions with no visual context), we hypothesized that this bias would extend to naturally occurring scenes where visual complexity increases ambiguity (e.g., context, demographic composition). To test this, we selected images (N=1972) featuring crowds and individuals from 7 news sources (e.g., NPR, National Geographic). All images were normed by coders (N=248) who rated the anger or happiness of people visible in vertical slices of each crowd or individual image. In a within-subjects design, participants (N=265) viewed 250 angry and happy crowds and individuals for 1s and rated the emotional intensity of the image from 0-9. Participants were randomly assigned to rate the images on anger or happiness. Relative to the norm ratings, participants overestimated the emotional intensity of angry crowds more than angry individuals, but there was no amplification effect for happy crowds. Thus, crowds – even ones containing visual complexity – potentiate perceivers' tendency to overestimate the anger of others.

Spatial Vision: Clinical

Saturday, May 17, 8:30 am – 12:30 pm, Pavilion

23.446 HIGHER MACULAR PIGMENT DENSITY IS RELATED TO IMPROVED CONTRAST SENSITIVITY WHEN MEASURED UNDER BLUE HAZE CONDITIONS

Cameron Wysocky[1], Jacob Harth[1], Lisa Renzi-Hammond[1], Billy Hammond[1]; [1]University of Georgia

Introduction: A long-standing question in Vision Science is why, under certain conditions, and all things equal (e.g., acuity), some individuals can see farther than others. A common explanation is the limiting effects of short-wave dominated light scatter on objects near the horizon (so-called blue haze). It has been argued for nearly a century that this may be partly why so many animals evolved intraocular blue-absorbing filters (i.e., to cut through this haze and enhance the contrast of distant targets). Humans also possess yellow foveal macular pigments (MP), and past data suggest that these pigments also serve this purpose. In this study, we provide additional data on the effect of MP and blue haze, on contrast sensitivity at multiple spatial frequencies. Methods: Fifty-eight healthy participants (N = 58; 18-28 years; M = 20.45 ± 2.90, 73% female, 33% non-White) with 20:40 uncorrected visual acuity or better, were assessed. MP optical density (MPOD) was measured using heterochromatic flicker photometry. Visual range (i.e., how far an individual can see in the distance) was measured by quantifying the amount of simulated blue haze (broadband xenon light paired with a blue-sky interference filter) necessary to completely obscure sinusoidal target gratings (3.2, 8, and 16-cpd). Results: MPOD was significantly related to the amount of blue haze required for participants to lose sight of the 3.2 ($r(56) = .54$, $p < .001$) and 8-cpd ($r(56) = .36$, $p < .01$) gratings. The relationship between MPOD and the 16-cpd grating was not significant ($r(47) = -0.03$, $p = .86$). Conclusions: MPOD is associated with increases in the visibility of low and medium spatial frequency targets when measured through simulated blue haze. This relationship does not extend to high spatial frequencies, indicating that individual differences in resolution acuity and target size may influence MP's effect on visual range.

23.447 SEEING IN DIM LIGHT: AGING FURTHER EXACERBATES DECLINES IN READING VISION AND FUNCTIONAL FIELD OF VIEW

Boris Penaloza[1] (b.penalozarojas@northeastern.edu), MiYoung Kwon[1]; [1]Northeastern University

While most research evaluates visual function under photopic (bright daylight) conditions, mesopic environments, such as dimly lit rooms or nighttime settings, pose real-world visual challenges. Older adults, despite having normal photopic visual acuity, often report significant difficulty with daily visual activities like reading or navigation in low-light conditions. Mesopic vision relies on both rod and cone photoreceptors, making it more susceptible to age-related retinal and optical changes, such as reduced pupil size, rod-photoreceptor loss, and optical opacity. However, age-related changes in visual function under mesopic conditions remain poorly understood. Here we investigate how aging affects reading vision and the functional field of view (FFV - the spatial extent of the visual field relevant to everyday tasks) under photopic and mesopic conditions. Using the MNREAD iPad app, we evaluated mesopic (2 cd/m²) and photopic (220 cd/m²) reading vision in 157 participants aged 18–84. Reading vision, indicated by maximum reading speed, critical print size, and reading acuity, declined linearly across adulthood, from the 20s to the 80s, with significantly steeper declines observed under mesopic conditions ($p<0.001$). A similar pattern was observed when FFV was assessed under divided attention: FFV was significantly reduced in older adults compared to young adults under both photopic and mesopic conditions ($p<0.01$). However, older adults experienced even greater reductions in FFV ($p<0.01$) under mesopic conditions. Our findings demonstrate the added burden of dim light on the aging population, thereby highlighting the critical need to assess the interplay between aging and mesopic conditions on functional vision. Understanding these interactions can inform the design of lighting solutions, accessibility measures, and targeted interventions to mitigate age-related visual decline and improve quality of life for older adults.

23.448 EXCESSIVE NOISE EXPLAINS THE DORSAL STREAM DEFICIT IN WILLIAMS SYNDROME: A COMPUTATIONAL APPROACH

Zvi Shapiro[1] (zrshapi@emory.edu), Alexander Weigard[2], Daniel Dilks[1]; [1]Emory University, [2]University of Michigan

Williams syndrome (WS) – a genetic disorder – is often characterized by a dorsal stream deficit. However, beyond impairment on many dorsal stream tasks, yet sparing on many ventral stream tasks, what does "dorsal stream deficit" mean? What is the mechanism driving this impairment? A recent computational modeling study found excessive noise explained the impairment on one dorsal stream task (a visually-guided navigation task) relative to one ventral stream task (a scene categorization task) in adults with WS. This finding leads to the intriguing hypothesis that excessive noise explains impairment on all dorsal stream tasks in WS. To directly test this hypothesis, we modeled the performance of adults with WS and mental-age (MA) matched controls on 3 pairs of dorsal/ventral tasks: i) intuitive physics/intuitive psychology, ii) configural/featural face processing, iii) configural/featural scene processing. 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 "threshold." Consistent with our hypothesis, we found disproportionately more noise in each of the three dorsal tasks in WS adults compared to MA controls, relative to the three ventral tasks. No such difference between groups was found for signal or threshold. Taken together, our results reveal excessive noise explains the deficit across dorsal stream tasks in WS, and suggests a reformulation of WS from simply being characterized by a dorsal stream deficit to being characterized by excessive noise in the dorsal stream.

NIH NEI R01 EY297242DDD

23.449 PREDICTING AMBLYOPIA AND CROWDING FROM APPEARANCE CAPTURES

Kaneiya Desai[1] (kaneiyad@berkeley.edu), Bilge Sayim[2], Dennis M. Levi[1], Ângela Gomes Tomaz[1]; [1]University of California, Berkeley, [2]University of Lille, France

Amblyopia is a neurodevelopmental disorder characterized by reduced visual acuity due to uncorrelated visual input early in life. Observers with amblyopia perceive stimuli as distorted (e.g., straight lines perceived as jagged) and experience stronger foveal crowding (i.e., worse performance with flanked targets) than controls. In a previous study (Gomes Tomaz et al., 2023), observers recreated the appearance of letters presented isolated or flanked to the fovea of the dominant/fellow or non-dominant/amblyopic eye. Here, we trained a convolutional neural network (CNN) with the appearance captured in that study to predict whether a response was made by an observer with or without amblyopia and under which condition, isolated or flanked. The architecture consisted of four convolutional, activation, and maximum pooling layers each, followed by two fully connected layers. Batch normalization followed the first, third, and fourth activation layers. The training dataset consisted of all appearance responses (captures), along with their classification label (e.g., "control" for a response made by a control observer (when classifying for group)). The validation dataset, consisting of 30% of all appearance captures, was randomized and balanced within and between all variables. The captures included in the validation dataset were not part of the training dataset. The best-performing models for classifying group and condition showed training accuracies of 88.0% and 95.0% and validation accuracies of 69.2% and 90.4%, respectively. The lower accuracy of the group predicting model is likely due to the high variability of features in the appearance space that characterizes each group. These results highlight the potential of CNNs in the analysis and classification of target appearance. Accurate models that classify and predict stimulus appearance in amblyopia will be lesioned to further study amblyopic visual perception in the future.

Supported by a grant from the National Eye Institute awarded to Dennis Levi (R21EY030609)

23.450 COULD SPATIAL SCRAMBLING LIMIT LETTER ACUITY IN AMBLYOPIA? INSIGHTS FROM EXAMINING LETTER MISTAKE PATTERNS

Raffles Xingqi Zhu[1] (raffles.zhu@mail.mcgill.ca), Robert F. Hess[1], Alex S. Baldwin[1]; [1]McGill University

Behavioural research in amblyopia has shown evidence of a "spatial scrambling". These effects would be consistent with a topographical disorganization affecting neural projections in the visual system. We investigated where in the processing hierarchy this scrambling might occur, and how it might contribute to the letter acuity deficit that is used clinically to characterize the severity of amblyopia. We constructed our stimuli using a physiologically-inspired algorithm. We consider two possible sites for scrambling to occur, either affecting: i) the outputs of log-Gabor filters (which mimic the oriented simple cell receptive field in V1), or ii) the inputs from isotropic subunits (i.e. LGN afferents) that form those oriented receptive fields. We refer to these image space manipulations as cortical and subcortical scrambling (CS and SCS), respectively. These distortions differ from typical additive noise such as bandpass noise, which we also tested for the purpose of comparison. The experiment involved a monocular task where the participant identified one of four possible noisy letters (bandpass and centered at 3 cycles/letter). Their sizes were rendered at half of the eye's acuity limit, obtained through a separate monocular letter acuity experiment. We hypothesized that amblyopic participants should exhibit different patterns of mistakes (obtained from confusion matrices) based on viewing condition as a result of spatial scrambling for amblyopic eye viewing. In total 18 amblyopic (13 strabismic, 5 anisometropic) participated in this study. We found moderate correlations between the ratio of interocular acuity difference and our metric of the mistake pattern divergence in the two eyes when stimuli were affected by scrambling noise (SCS: $r=0.65$, $P=0.0038$; CS: $r=0.47$, $P=0.047$) but not for letters shown in bandpass noise ($P=0.35$). These results suggest neural scrambling, as early as the monocular inputs to the cortex, may play an important role in limiting letter acuity in amblyopia.

23.451 IMPROVING MOBILITY FOR PEOPLE WITH LOW VISION USING AUGMENTED REALITY

Sarit F.A Szpiro[1], Lior Maman[2]; [1]Department of Special Education, University of Haifa, [2]Department of Information Systems, University of Haifa

Goal People with low vision struggle with various daily tasks, including walking between hazards. Augmented reality (AR) has been demonstrated to improve the accessibility of various items in daily tasks by modifying, magnifying, and changing the visibility of items. Here, we examined the utility of modifying the visibility of hazards using AR in a walking task. Methods Two groups participated in the study - typically sighted and low vision participants. Both groups had to walk an obstacle course with various items along the route in two conditions: (1) No AR augmentation; (2) With AR augmentation. In both condition, physical items were present on the route, but in the AR condition each item's outline was highlighted to attract attention and increase its visibility. Results We found a difference between groups. For the low vision group, we found that walking time was shorter and less collisions in the AR condition. In contrast, in the typically sighted group time and collisions did not significantly change. These results illustrate the potential of AR glasses as a vision enhancement aid in navigation and highlight areas for future work in computer vision, design, and vision science.

Research grant ISF 1198/22 to SFAS

23.452 HOW TO SIMULATE VISUAL IMPAIRMENT FOR RESEARCH IN VISUAL PERCEPTION?

Carla Lanca[1] (crl9600@nyu.edu), Robert Volcic[1,2,3]; [1]New York University Abu Dhabi, [2]Center for Artificial Intelligence and Robotics, New York University Abu Dhabi, [3]Center for Brain and Health, New York University Abu Dhabi

The use of simulation to degrade visual function is an essential tool in visual perception research. However, there are no standardised protocols for simulating different diseases and levels of impairment. Here, we explore various methods used to degrade visual function, inclusion criteria for selection based on participant's baseline visual functions, and discuss potential advantages and limitations of each method. Methods such as spherical and toric lenses, Bangerter filters, fog filters, Cambridge simulation glasses, cataract glasses, prisms, and goggles for simulating visual field lesions are described. Their effects on visual functions, including visual acuity, contrast sensitivity, stereoacuity and visual field, are also discussed. Simulations may not always fully replicate true visual impairment, mainly due to age of onset of the disease, lack of adaptation to the specific visual condition, and the presence of multiple deficits in patients. Nevertheless, simulations allow the study of sudden and acute visual loss and short-term adaptation processes, and may contribute to developing improved rehabilitation strategies for patients with ophthalmic diseases.

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.

23.453 USING SSVEPS TO CHARACTERISE WIDE-RANGING RETINOPATHY: IMPLICATIONS FOR CLINICAL TRIALS

Kim Eliane Staeubli[1,2,3] (kim.staubli.21@ucl.ac.uk), Marc Pabst[1,2], Roni Maimon-Mor[1,2], Mariya Moosajee[1,3,4], H. Steven Scholte[5], Tessa Dekker[1,2,3]; [1]Institute of Ophthalmology, University College London, London, UK, [2]Experimental Psychology, University College London, London, UK, [3]Moorfields Eye Hospital NHS Foundation Trust, London, UK, [4]The Francis Crick Institute, London, UK, [5]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Inherited retinal diseases (IRDs) are a leading cause of blindness worldwide. CRB1-related retinopathies are an IRD in which mutations to a single gene cause a wide range of sight loss profiles for reasons not well understood. With emerging gene therapies aiming to restore retinal gene function, we need to characterise the impacts of the disease on retino-cortical development and function, and establish objective clinical trial endpoints that can be used across disease phenotypes. So far, this has been challenging due to patients' widely varying abilities and the associated variability in test feasibility and sensitivity. We used a child-friendly electroencephalography (EEG) paradigm to quantify retinocortical visual system function in a cohort of patients with CRB1-related retinopathy. This involved recording steady-state visual evoked potentials (ssVEPs) in response to flickering sinusoidal gratings and full-field flicker embedded in age-appropriate movies. Our results reveal significant ssVEP attenuation in patients with CRB1-related retinopathy compared to sighted controls, with the most pronounced reductions in those with generalised retinal involvement and milder reductions in those with isolated macular dystrophy. While ssVEP curve amplitude differed across CRB1 phenotypes, we found no significant differences in curve shape at the group level. Intra-session reliability was high across the sample, but lower in patients with severe impairment, likely due to fixation instability. Notably, full-field flicker stimuli elicited reliable responses in these patients, strongly correlating with visual acuity, suggesting that this stimulus is an important addition to phase-reversing patterns when assessing visual function in severe vision loss using VEPs. These findings highlight the utility of VEPs as a non-invasive, objective measure of visual function that can be applied across a broad spectrum of IRD phenotypes and visual abilities. VEPs thus provide a valuable complement to traditional assessments of visual function for characterising disease characteristics and assessing treatment efficacy in IRDs.

This research was supported by the Santen SenSyT studentship scheme.

23.454 CROWDING AFFECTS VISUAL ACUITY MORE FOR VERTICALLY THAN HORIZONTALLY ALIGNED LETTER CHARTS

John Siderov[1] (j.siderov@hud.ac.uk), Frantisek Pluhacek[2], Andrea Sláviková[2], Sarah Waugh[1]; [1]University of Huddersfield, [2]Palacky University Olomouc

The impact of horizontal/vertical asymmetry in crowding on visual acuity is not clear as reports differ. To assess crowding asymmetry on measurements of visual acuity, we measured visual acuity using custom letter charts arranged in vertical and horizontal alignment with varying inter-letter and inter-row/inter-column spacings. Specially constructed logMAR letter charts based on the formats of the ETDRS (Ferris et al, 1982) and Waterloo (Strong and Woo, 1985) charts were used. Charts comprised rows or columns of dark Sloan letters, ranging from 1.0 to -0.3 logMAR at the 6 m test distance, printed onto white cardboard backgrounds. Edge-to-edge spacings between adjacent letters of 100%, 75%, or 50% of the optotype size were used for each alignment condition (vertical and horizontal). Row and column spacing was fixed at the inter-optotype spacing of the row below or column to the right, in each chart, respectively. An additional set of charts was created to assess test-retest repeatability. Monocular best-corrected acuity was measured in 30 adult observers. Charts were presented in random order (i.e. vertical or horizontal alignment and spacing) and observers asked to read the letters from largest to smallest. Measurements continued until 3 or more errors were made on a single row/column. Final visual acuity was determined using letter-by-letter scoring. Measurements were repeated on a separate day. LogMAR acuity was significantly better for the horizontally aligned formats ($F_{1,29}$ $p<0.001$), and worsened with decreasing letter spacing ($F_{2,58}$ $p<0.001$), indicating more crowding for vertically aligned charts and closer inter-optotype separations, respectively. However, differences were small, about 0.04 logMAR (i.e. 2 letter difference). Test-retest results resulted in a significant interaction between letter spacing and alignment ($F_{2,58}$ $p<0.05$), indicating a learning effect for the vertical aligned charts at the largest inter-optotype separation, suggesting that practice may overcome some of the differences due to crowding.

Supported, in part, by Grants URN020-01 from the University of Huddersfield and IGA_PrF_2017_003 and IGA_PrF_2018_007 from the Faculty of Science, Palacky University Olomouc, Olomouc, Czech Republic.

23.455 CLASSIFICATION OF GLAUCOMA USING ON-AND OFF-PATHWAY BIASING SSVEPS

Martin T W Scott[1], Hui Xu[1], Alexandra Yakovleva[1], Robert Tibshirani[1], Jeffrey L Goldberg[1], Anthony M Norcia[1]; [1]Stanford University

The ON and OFF retino-cortical pathways encode increments and decrements (respectively) in perceived luminance. Recent evidence from small animal models and human electrophysiology suggests that the OFF-pathway is more vulnerable to damage from Glaucoma than the ON-pathway. Thus, OFF-pathway based measurements of visual function may be useful in the diagnosis of Glaucoma. The steady-state visually evoked potential (SSVEP) measured with high density electroencephalography (EEG) can be used to non-invasively make these functional measurements. Here, in a cohort of 98 Glaucoma patients and 71 age-matched controls, we examine the extent to which SSVEP measurements obtained using ON- and OFF-pathway biasing visual stimuli are useful for case-control classification. Full-field SSVEPs were obtained using a spatial array of low-contrast, flicker-frequency tagged hexagonal probes. The luminance of these probes was modulated at 2.7Hz with a saw-tooth function, the fast phase of which biased SSVEPs towards either the ON- or OFF-pathways. Using both a logistic regression and a random forest classifier, we report moderate overall classification performance, with SSVEP features evoked by OFF-pathway biasing stimuli producing a small improvement in predictive accuracy over ON-pathway biased SSVEP features. Perhaps more importantly, classification based on signal phase features significantly outperformed classification based on signal amplitude features, particularly when OFF-biasing stimuli were used. While the classification performance we obtain is not yet sufficient for direct clinical use, it highlights the high information value of phase-related features, a metric that is often omitted in existing related work using the isolated-check VEP.

Research to Prevent Blindness, NIH (R01-EY030361-01, P30-EY026877)

23.456 INCOMPLETE LETTER RECOGNITION IS LIMITED BY CORTICAL AND NOT OPTICAL FACTORS: SIMULATING THE DEFICITS OF VISUAL-LED DEMENTIA IN HEALTHY ADULTS

Zien Huang[1] (zien.huang.23@ucl.ac.uk), Tessa M. Dekker[1], Sebastian J. Crutch[1], Keir X. X. Yong[1], John A. Greenwood[1]; [1]University College London

The Graded Incomplete Letters Test (GILT) was recently developed to detect visual symptoms in dementia, including Posterior Cortical Atrophy (PCA; a neurodegenerative syndrome predominantly affecting visual cortex). The GILT measures the threshold for recognising letters degraded by removing pixelated letter sections (decreasing 'completeness'). Thresholds are strongly elevated in PCA patients relative to typical adults (Yong et al, 2024), though it is unclear why PCA patients struggle on the task. We aimed to distinguish the role of cortical factors associated with PCA (elevated crowding and impaired feature integration) from disruptions by age-related optical factors (e.g. blur or low contrast through glaucoma) that could be mistaken for PCA-related visual loss by simulating these effects in typical adults (n=6). To examine optical factors, we applied blur and lowered contrast separately to incomplete letter stimuli, with participants required to identify one of 12 uppercase letters presented foveally on each trial. Stimuli were degraded at different completeness levels using QUEST. Without optical factors applied, thresholds averaged ~5% completeness. With blur or low contrast, there was very little effect until blur/contrast approached detection/visibility thresholds, where small elevations to 8% completeness were evident. These deficits do not reach the levels seen in PCA (median: 47% completeness). To examine cortical factors, we simulated elevated crowding (impaired object recognition in clutter) by moving stimuli into peripheral vision, where crowding is high in typical adults. We simulated issues in feature integration by varying pixel size to alter the distribution of letter degradation (limiting the spatial integration of letter features), or by degrading letters dynamically with limited-lifetime pixels (limiting temporal integration). The combination of elevated crowding and feature integration impairments increased thresholds as high as 40% (with dynamic presentation in peripheral vision), comparable with average PCA deficits. Poor incomplete letter recognition is therefore more strongly associated with cortical than optical factors.

23.457 MAPPING VISUAL FIELD LOSS IN LEBER HEREDITARY OPTIC NEUROPATHY USING FMRI

Hugo Chow-Wing-Bom[1,2] (hugo.chow-wing-bom.15@ucl.ac.uk), Matteo Lisi[3], Noah Benson[4], Patrick Yu-Wai-Man[1,5,6,7], Frederic Dick[1,2], Roni Maimon-Mor[1], Tessa Dekker[1,2]; [1]University College London, UK, [2]Birkbeck-UCL Centre for NeuroImaging, London, UK, [3]Royal Holloway University of London, UK, [4]University of Washington, USA, [5]University of Cambridge, UK, [6]Cambridge University Hospitals NHS Foundation Trust, UK, [7]Moorfields Eye Hospital NHS Foundation Trust, London, UK

Assessing visual function across the visual field is very challenging when vision is severely impaired because dense central scotomas limit the ability to keep the eyes comfortably fixed on a central target, and therefore the precise stimulation of retinal loci. Functional magnetic resonance imaging (fMRI), combined with structure-based retinotopic atlases could be valuable tools for measuring visual function alongside standard behavioral perimetry approaches in such cases, because the retina can be passively stimulated without the need for accurate fixation. We investigate the use of this approach for assessing severe visual field loss in Leber Hereditary Optic Neuropathy (LHON). LHON is a mitochondrial disease causing acute loss of retinal ganglion cells, leading to severe central vision loss, often extending up to 10-15 degrees of eccentricity. Building on our previous fMRI work showing variations in cortical contrast sensitivity across a large visual field (40deg) in normal-sighted individuals, we used our brain-based approach to map visual field loss in 10 LHON patients. Our initial approach combined population receptive field (pRF) modelling with the modelling of BOLD modulations to large-field contrast-varying sinusoidal gratings, to quantify cortical contrast sensitivity across eccentricities and visual field quadrants. To apply this to severely sight impaired patients, we estimated retinotopic tuning of visual cortex using a fixation-free, structure-based atlas linking cortical activity to visual space locations (i.e., Benson retinotopic template), with additional model-based eccentricity scaling. Our results show that this approach accurately maps V1 contrast sensitivity to visual field loss, showing alignment with behavioral perimetry maps. This objective cortical mapping could offer new insights into how disease progression and recovery in severe sight loss impact visual information processing, including potential responses to gene therapy, which are emerging as potential treatment options for LHON.

Supported by Moorfields Eye Charity (GR001315; London, UK) and the Birkbeck-UCL Centre for NeuroImaging (BUCNI; London, UK).

Visual Memory: Models

Saturday, May 17, 8:30 am – 12:30 pm, Pavilion

23.458 A POPULATION VECTOR MODEL OF VISUAL WORKING MEMORY FOR NATURALISTIC SCENES

Steven J. Luck[1], John E. Kiat[1]; [1]University of California, Davis

Visual working memory is essential for navigating and interacting with the visual environment, so it is important to understand how complex natural scenes are represented in working memory. Natural scenes are characterized by complex contours, continuously varying feature gradients, and spatial relationships, but most recent research on visual working memory has focused on simplified arrays of discrete artificial objects, favoring experimental control over ecological relevance. Similarly, quantitative models of visual working memory have focused on pre-parsed discrete objects that vary along a small number of simple dimensions such as color and orientation, and it is not clear how these models could be updated to represent complex, photograph-like scenes. To address this limitation of current models, we have developed a "population vector model" of visual working memory that represents a complex scene as a noisy vector of activation values across a population of neurons in visual cortex. It uses CORnet—a convolutional neural network designed to mimic the properties of the ventral object recognition pathway—to estimate the population representation of a given scene at different levels of abstraction, from V1 through inferotemporal cortex (IT). We tested this model in neurotypical young adults using behavioral change detection experiments and using a modified 1-back design to collect EEG data. We found that the model predicted behavioral accuracy and response times extremely well, accounting for over 75% of the variance across scenes. The model also predicted differences in the scalp patterns of EEG activity across scenes during the delay period. In addition, we found that the abstract IT-like representations of the model accounted for much more unique variance in behavior and brain activity than the spatially detailed V1-like representations. Although far from complete, this model provides a path toward understanding how complex naturalistic scenes are stored in working memory.

This work was made possible by NEI grant R01EY033329

23.459 MEMORY STRENGTH AS ACCUMULATED ACTIVATION OVER TIME: UNIFYING TEMPORAL ORDER AND REPETITION EFFECTS

Maria Robinson[1] (maria.robinson@warwick.ac.uk), Timothy Brady[2]; [1]University of Warwick, [2]University of California, San Diego

Can time-based memory effects (like recency) and strength-based effects (like repetition) be unified under a single explanatory framework? We addressed this question by testing whether the Target Competition Confusability (TCC) framework (e.g., Robinson & Brady, 2023) could predict performance in repetition conditions using model parameters estimated solely from temporal order effect conditions. We developed a summative model that conceptualizes repetition-based effects as the cumulative activation strengths elicited by individual items over time. The model treats repeated items as the sum of their individual presentation activations, using parameters that are independently estimated from a sequential task condition where items were probed in specific orders. For instance, in a six-item sequence with the first two items repeated, the model predicted performance for the repeated item by summing the activations from single-presentation trials that probed the first and second items in the sequence. Our results showed that this summative model successfully predicted the entire distribution of memory errors in repetition conditions with no free parameters used to fit these distributions. Furthermore, it did so across different numbers of repetitions and for different combinations of temporal orders. While prediction accuracy did not match test-retest reliability, the model captured both the shapes of error distributions and the ordinal effects of repetition-based manipulations. These findings demonstrate that a single modeling framework can unify seemingly distinct temporal order and repetition effects on memory,

and we discuss how this elucidates the underlying principles that govern memory performance more broadly.

23.460 VARIABLE STIMULI AND TESTS GENERATE VARIABLE CONFIDENCE: CHALLENGING THE 'RECOLLECTION' SIGNAL IN WORKING MEMORY

Timothy Brady[1] (timothy.brady@gmail.com), Yong Hoon Chung[2], Jamal Williams[3], Viola Stoermer[2]; [1]University of California, San Diego, [2]Dartmouth College, [3]Yale University

The dual-process signal detection (DPSD) model has been influential in long-term memory research. This model makes the strong claim that when a large number of high-confidence responses occurs in a memory task, this provides direct evidence for an all-or-none recollection process that is distinct from familiarity signals. This model has recently been extended to visual working memory; for example, suggesting that stimuli like real-world objects and colors engage different amounts of 'recollection'. Here, we demonstrate that the DPSD model incorrectly considers trial-by-trial variability in memory to be evidence of recollection. We collected confidence ratings in two color change detection conditions matched for overall performance but differing in variability. In Experiment 1, we compared a fixed condition (set size 3) to a variable condition that mixed set sizes (1-6). Despite matched overall performance, the DPSD model attributed the excess high-confidence responses in the variable condition to 'recollection' ($Ro=0.31$ vs. $Ro<0.01$ in fixed, $p<0.01$). Experiment 2 used a fixed set size (four) but different test probes to induce different amounts of variability. In the fixed condition, we used consistent 40° color wheel separations between targets and foils, while the variable condition used varying foil distances. Again, the DPSD model incorrectly interpreted the excess high-confidence responses in the variable condition as greater 'recollection' ($p<0.001$). These results challenge the DPSD model's claim to measure 'recollection' through confidence ratings alone, without independent measures like context memory. While multiple processes may contribute to memory match signals, it is likely that people collapse these onto a single evidence axis when making change detection decisions, making them inseparable through confidence reports and ROCs derived from those reports. By demonstrating how task variability shapes confidence patterns, this work provides a more parsimonious framework for interpreting memory decisions across different stimulus types and experimental conditions.

NSF BCS 2146988

23.461 COMPUTATIONAL MODEL SELECTION FAILS TO RECOVER THE CAPACITY-LIMIT IN SHORT-TERM MEMORY RECALL

Jessica Golding[1], Tianye Ma[2], Weiwei Zhang[3]; [1]University of California Riverside

Schurgin et al. (2020) developed the Target Confusability Competition (TCC) model, a uni-dimensional signal detection model for short-term memory (STM) recall. Formal model comparisons, using goodness-of-fit indices like Bayesian Information Criterion (BIC), favored the TCC model over capacity-limited models. such as those incorporating a pure-guess component. Nevertheless, the validity of this computational model selection method in inferring the capacity limit in STM is unclear. To address this, the current study evaluated the validity of model selection by replicating Schurgin et al.'s analysis on simulated data from a process model explicitly incorporating a capacity limit. Simulations were conducted using a revised version of the Binding Pool model (Swan & Wyble, 2013), with population code decoding as the response mechanism. The TCC model was compared to two capacity-limited models: the Slot model and the TCC with capacity (TCCk) model, both of which were included in the original analysis by Schurgin et al. (2020). Results indicated that parameters from the TCC and Slot models, but not the TCCk model, successfully tracked parameter changes in the process model. However, the TCC model consistently outperformed the other models across a wide range of parameters in the Binding Pool simulations. Importantly, model selection failed to detect the capacity limit embedded in the process model, suggesting that computational model selection alone is insufficient for inferring underlying memory mechanisms. These findings highlight the risk of over-interpreting good model fits without careful theoretical validation.

23.462 MODELS OF LONG-TERM MEMORY CAN ALREADY PRODUCE SEVERAL OF THE KEY SIGNATURES OF VISUAL WORKING MEMORY STORAGE

Geoffrey Woodman[1], Sean Polyn[1]; [1]Vanderbilt University

The study of visual working memory and visual long-term memory diverged decades ago, resulting in views about the diagnostic nature of certain measures that might not be valid. For example, in the study of visual working memory, it is viewed as critical to show a capacity limit in storage as a way of demonstrating that we are measuring the limited capacity visual working memory store and not the unlimited capacity visual long-term memory store. Although this is sound logic, it remains to be seen whether several of the hallmarks of working memory might already fall out of existing models of human long-term memory. Here we show that the precision and set size effects that we visual working memory researchers often use to validate our assumption that we are studying visual working memory are also observed as a natural result of the dynamics of contextual models of long-term memory storage and retrieval. We discuss how the situation motivates re-examining unified models of human memory, paired with multi-modal empirical studies that target key questions about the nature of visual working memory and long-term memory.

This work was supported by grants from the National Science Foundation (BCS-2147064) and the National Eye Institute (T32-EY007135).

Theory

SATURDAY, MAY 17, 8:30 AM – 12:30 PM, PAVILION

23.463 BAYESIAN COMPARISONS BETWEEN REPRESENTATIONS

Heiko H. Schütt[1]; [1]University of Luxembourg

A fundamental methodological question for vision science is how we test whether the representations in different models and brain areas are similar to each other or not. Due to the high dimensionality of visual representations comparing them is not easy and it remains unclear which methods are most useful with variations of encoding models and kernel or distance based methods being the main contenders. Here, I propose novel methods to compare representations based on Bayesian statistics about read out models from the representations. When we apply a prior to the read out weights, we can compute the predictive distribution for read outs from our representation. The predictive distribution is a full description of the inductive bias and generalization of a model in Bayesian statistics, making it a great basis for comparisons. To compare representations, one can assume that one representation is read out from the other one and use the likelihood of getting that representation. Alternatively, one can use distances for probability distributions like the total variation distance or Jensen-Shannon distance to compare the predictive distributions. For a linear readout with Gaussian priors, we can analytically solve all computations without dimensionality reductions of the representations and our metrics just depend on the linear kernel matrix of the representations. Thus, the new methods connect linear read-out based comparisons to kernel based metrics like centered kernel alignment and representational similarity analysis. I apply these new methods to compare deep neural networks trained on ImageNet to each other and to fMRI data from the natural scenes dataset. The new methods broadly agree with existing metrics, but consider smaller sets of representations to be equivalent. They vary less across different random image samples, and have some theoretical advantages. Thus, these new metrics nicely extend our toolkit for comparing representations.

23.464 NEURAL NETWORKS RELY ON TOP-DOWN REPRESENTATIONS UNDER CONDITIONS OF UNCERTAINTY

Holly Kular[1], Robert Kim[2], John Serences[1], Nuttida Rungratsameetaweemana[3]; [1]University of California San Diego, [2]Cedars-Sinai Medical Center, [3]Columbia University

Natural visual environments are highly structured and statistical regularities can be used to form expectations that guide behavior. However it is not well understood how top-down information such as prior expectation modulates information processing along the cortical hierarchy. Here, we test the hypothesis that expectations can shape processing at all levels of information processing, particularly when stimulus uncertainty is high. We isolated effects of expectation using a hierarchical continuous-time recurrent neural network (RNN) trained on a perceptual decision task while simultaneously manipulating stimulus uncertainty and set size. To establish top-down expectations in the network, the RNN was trained with the expected stimulus presented more frequently than others. We show that the RNN extracts sensory statistics and develops expectations that improve processing when stimuli are uncertain. Linear SVM classifiers were trained on RNN firing rates to decode stimulus identity, comparing conditions of low/high uncertainty and small/large stimulus sets. Decoding accuracy was highest for expected stimuli across all conditions. Additionally, the expected stimulus could be more readily decoded from activity in the final layer under the high uncertainty condition, suggesting a strong top-down influence from higher layers. This effect is more pronounced when the stimulus set size is large, indicating a greater role for top-down modulation from the final layer as the amount of incoming sensory information decreases. These findings show that the RNN developed expectations and relied on top-down signals under uncertainty and with larger stimulus spaces, supporting the hypothesis that expectations facilitate visual processing particularly when bottom-up sensory information is limited.

23.465 USING TRANSFER LEARNING TO IDENTIFY A **NEURAL NETWORK'S ALGORITHM**

John Morrison[1,2], Nikolaus Kriegeskorte2[2,3,4], Benjamin Peters[5,6]; [1]Philosophy Department, Barnard College, Columbia University, [2]Zuckerman Mind Brain Behavior Institute, Columbia University, [3]Department of Psychology, Columbia University, [4]Department of Neuroscience, Columbia University, [5]School of Informatics, University of Edinburgh, [6]School of Psychology & Neuroscience, University of Glasgow

Algorithms generate input-output mappings through step-by-step operations on representations. In vision science, algorithms explain biological and artificial processes. For example, feature weighting explains image categorization, sequential sampling explains visual search, and Bayesian inference explains cue combination. The standard parts-based approach is to look for parts in the underlying network corresponding to the parts of the algorithm. But we have not been able to find many such parts, perhaps because they are too entangled. We propose the alternative approach of identifying a system's algorithm by assessing how quickly it learns alternative input-output mappings, that is, its transfer learning profile. We use artificial networks to demonstrate that this approach is promising. In our first toy experiment, we used transfer learning to show that networks can use different algorithms to compute the same arithmetic function. Importantly, the distinction between these two types of network was not evident in traditional parts-based analyses such as encoding and decoding analysis, weight similarity analysis, and representational similarity analysis. In our second experiment, we used transfer learning to show that convolutional networks use different algorithms to classify images of ellipses. The ellipses were generated from three latent variables: area, color, and circularity. The networks were trained to classify ellipses into two classes defined by a linear function of their latent variables. We could identify algorithms of two types. The first used independent feature detectors (one for area and one for circularity) and the second used joint feature detectors (one for both area and circularity). Once again, the distinct algorithms were not evident in traditional parts-based analyses. In an ongoing experiment, we are using transfer learning to show that convolutional networks use different algorithms to classify objects in naturalistic images with more complex latent features. Combined, our results suggest that transfer learning is a promising alternative to parts-based approaches.

23.466 VISUAL COGNITION IN VISION-LANGUAGE MODELS

Krista A. Ehinger[1]; [1]The University of Melbourne

Large language models (LLMs) show human-like performance in a range of language tasks such as question answering, text editing, and text composition. These models are trained on massive text datasets and demonstrate an impressive ability to flexibly recombine what they have learned in novel ways to perform arbitrary tasks (Brown, et al., 2020). The latest generation of LLMs are multimodal and able to process images as well as text. Do these vision-language models (VLMs) learn similarly flexible representations for visual tasks? A state-of the art VLM (GPT-4) was tested on a battery of visual cognition tasks. In addition to scene and object recognition tasks, the model was tested on fundamental abstract visual tasks like numerosity judgement (counting), contour completion, determining whether two points lie on the same contour, and determining whether a point is inside or outside a contour. The model performed well on recognition tasks and showed some human-like patterns in performance in other tasks; for example, more errors in numerosity judgements for larger set sizes. However, the model also showed striking failures, such as an inability to perform contour completion even in the simplest cases. In general, the model seemed to perform poorly in tasks which involve "visual routines" (Ullman, 1984) or flexible reasoning about the elements of images. These results suggest that some high-level visual tasks, like recognizing scenes and objects, don't actually require more fundamental visual processes like contour completion. Models can learn to do recognition tasks well while apparently lacking basic visual operations. The results also suggest that basic, general-purpose visual mechanisms don't necessarily emerge automatically when models are trained on image recognition tasks.

23.467 VISUALIZING THE UNSEEN: PERCEPTOGRAPHER, AN AI ENGINE FOR VISUALIZING BRAIN-STIMULATION-INDUCED PERCEPTUAL EVENTS

Elia Shahbazi[1] (elia.shahbazi@nih.gov), Drew Nguyen[1], Rasel Ahmed Bhuiyan[2], Adam Czajka[2], Arash Afraz[1]; [1]National Institutes of Health, [2]University of Notre Dame

We recently developed a novel paradigm called perceptography1 to visualize complex perceptual distortions induced by local stimulation of the inferotemporal (IT) cortex. Perceptography uses machine learning to create and optimize specific complex image distortions that were hard for the animal to distinguish from the state of being cortically stimulated. This paradigm opens the door to scientific measurement of subjective perceptual events but comes with a serious image generation challenge. In the absence of a theory linking neuronal activity to visual perception, the perceived visual distortions following brain stimulation may be of any nature. Thus, to avoid biases, our image generation engine that aims to mimic stimulation-induced visual distortions should be able to create any possible distortion in the image. State-of-the-art AI provides two fundamentally different approaches for image generation (for example, face generation): Generative adversarial networks (GAN), which are suitable to create any possible natural face but unable to develop off-manifold distortion to the face, and diffusion models (DM), which could make any image from a given text prompt but have difficulty fine-tuning image/face identities in a continuous space using prompts2,3. We introduce Perceptographer, a novel structure designed to solve this problem. It combines a GAN(StyleGANEX), an autoencoder, and a DM (pix2pix-instructor&LLM) to create a novel, customizable engine to navigate this dense multidimensional space. We invert each GAN-DM output image into a perturbable latent space, enabling the Perceptographer to generate off-manifold distortions, apply various distortion levels, and reconstruct any random point in a continuous feature-space. Perceptographer offers a novel, customizable framework for visualizing brain-stimulation-induced perceptual events in different parts of the visual cortex. This framework overcomes the limitations of current image generation models in handling complex, off-manifold image distortions, providing new opportunities for visualizing and understanding brain-stimulation-induced perceptual phenomena across multiple cortical regions.

This research was supported by the Intramural Research Program of the NIMH ZIAMH002958

23.468 PRIVILEGED REPRESENTATIONAL AXES IN BIOLOGICAL AND ARTIFICIAL NEURAL NETWORKS

Meenakshi Khosla[1] (meenakshik1993@gmail.com), Sudhanshu Srivastava[1], Alex Williams[3], Josh McDermott[2], Nancy Kanwisher[2]; [1]University of California San Diego, [2]Massachusetts Institute of Technology, [3]New York University

How is information coded in the brain? To answer this question, neuroscientists have increasingly adopted methods such as representational similarity analysis, linear encoding and decoding, canonical correlation analysis and centered kernel alignment that analyze the geometry of the population code while disregarding the actual tuning of neurons. But are neural tunings in fact arbitrary and irrelevant, or might they matter, privileging some representational axes over others? We developed methods to probe for privileged representational axes in biological and artificial neural networks, and applied them to multiple types of neural data from diverse brain systems, and to DCNNs trained on natural sensory stimuli. We found that representational axes were consistent between individuals, between artificial neural networks varying in architecture and learning objectives, and between brain systems and artificial neural networks trained on the same modality. These results indicate that representational axes in neural systems are not arbitrary, and can arise in artificial systems with none of the priors and meanings assigned to these axes by humans. We further found that the privileged axes used in the brain and DCNNs confer important computational advantages, including economy in the number of active neurons, minimization of downstream wiring costs and improved generalization under biologically realistic constraints. Finally, our metrics of axis alignment also distinguish the fit of models to the brain that are not well discriminated by standard metrics. These findings underscore the importance of representational axes in both biological and artificial neural systems, offering new insights into their origins and functional relevance.

23.469 SPATIAL-SCALE INVARIANT PROPERTIES OF VISUAL CORTEX IN MAMMALS

Raj Magesh Gauthaman[1], Brice Ménard[1], Michael Bonner[1]; [1]Johns Hopkins University

How does visual cortex encode information about the visual world in the coordinated activity of hundreds of millions of neurons? Our recent work analyzing a large-scale fMRI dataset containing neural responses to natural scene images has demonstrated that cortical representations of natural images are high-dimensional and exhibit scale-free covariance structure. Surprisingly, this characteristic statistical signature is not only universal across low- and high-level regions of human visual cortex, but is also observed in the population activity of single neurons in mouse primary visual cortex (V1). What properties of cortical population codes allow us to observe the same statistical structure at the level of single neurons in mice and voxels in humans? To investigate this question, we analyze two datasets: the Natural Scenes human fMRI dataset and a large-scale mouse calcium imaging dataset, both containing V1 responses to natural images, but measured at dramatically different resolutions ranging from single neurons (~20 μm) to voxels containing ~10^5 neurons (1.8 mm). Using a cross-decomposition estimator, we confirm that stimulus-related variance is distributed as a power law along all available latent dimensions (>10^3) in both humans and mice. In fact, these latent dimensions are patterned on the cortical surface with characteristic spatial scaling: high-variance dimensions vary on coarse scales while low-variance dimensions vary on fine scales. Crucially, we discover a stable power-law relationship between variance and spatial scale that is identical across both mammalian species. Together, this remarkable universality in the covariance statistics of human and mouse V1 population activity suggests a generic encoding principle of visual cortex. More broadly, this result explains why studying cortical responses using neuroimaging at spatial scales far removed from single neurons reveals interesting principles of visual encoding — the statistics of visual responses are self-similar across many orders of magnitude of spatial scale.

This research was supported in part by a Johns Hopkins Catalyst Award to MFB, Institute for Data Intensive Engineering and Science Seed Funding to MFB and BM, and grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP).

23.470 TRAINING CONVOLUTIONAL NEURAL NETWORKS WITH BLURRY IMAGES ENABLES THE LEARNING OF MORE HUMAN-ALIGNED VISUAL REPRESENTATIONS

Ikhwan Jeon[1], Connor Parde[1], Frank Tong[1]; [1]Vanderbilt University

Although convolutional neural networks (CNNs) can achieve human-level recognition accuracy on natural images, research has revealed systematic deviations between CNNs and human vision, including susceptibility to visual noise (Jang et al., 2021; Geirhos et al., 2018) and insensitivity to shape information (Geirhos et al., 2019). However, recent work has shown that these deviations can be reduced by providing CNNs with auxiliary training on blurry images (Jang and Tong, 2024). In this work, we further demonstrate how blur training can improve the alignment between CNNs and human vision by evaluating the quality of metameric stimuli generated from CNNs (Feather et al., 2023). The metamers of CNNs are defined as image pairs that produce nearly identical responses; such metamers can be generated by modifying an initially random noise image until it produces CNN responses that closely mimic the responses to a reference object image. To investigate the potential benefits of CNN blur training, we generated metamers from both standard and blur-trained CNNs. We also considered metamers generated from both RGB- and grayscale-trained models to test the hypothesis that color information may allow CNNs to learn “shortcut” strategies that are less aligned with human vision. All metamers were generated using the responses from the final convolutional layer of a given CNN. Human observers and cross-validated CNN models then classified the metamers generated from each CNN. Across all conditions, the metamers of blur-trained models were recognized more accurately than those generated from clear-trained CNNs. This general benefit of blur training for creating recognizable CNN metamers indicates that blur training improves alignment between the internal representations of CNNs and the human visual system. Furthermore, this effect was more pronounced in grayscale than RGB, suggesting that color-based short-cut learning may have been mitigated, facilitating the learning of more canonical visual representations.

This research was supported by NEI grants R01EY035157 to FT and P30EY008126 to the Vanderbilt Vision Research Center.

23.471 A SIGNAL DETECTION MODEL FOR THE ANALYSIS OF CONTINUOUS RESPONSE GRADIENTS AND AN APPLICATION TO OTHER-RACE EFFECTS

Fabian A. Soto[1], Emily R. Martin[1]; [1]Florida International University

Many perceptual tasks result in behavioral gradients depicting a continuous response as a function of stimulus value. Examples are gradients of confidence in perceptual decisions or response times. In many cases, researchers are interested in linking the mechanisms underlying continuous behavioral measures and perceptual choices. For example, while the other-race effect (ORE) in face recognition is usually measured using proportion of correct responses, the other-race categorization advantage (ORCA) is usually measured using response times. Understanding the mechanisms behind these two other-race effects could benefit from a way of measuring both of them in the same scale, such as sensitivity measures and thresholds obtained from detection theory. Here, we propose a generalization of the signal detection model for the psychometric curve that deals with continuous responses such as response times. As in the traditional model, we assume normally-distributed decision variables with means and variances that change depending on the presented stimulus. We also assume that a monotonic link function transforms such variables into the measured responses, which are perturbed by random normal noise. The model is a generalization of traditional signal detection models, which are obtained by assuming a staircase link function. We propose an algorithm that uses a combination of quantile functions and monotone spline regression to estimate the parameters of this model from data, and show that the inclusion of a flexible link function allows the model to fit continuous data better than ROC analyses previously proposed for continuous data. We show through simulations that our estimation procedure produces accurate parameter recovery. One can directly compare parameters estimated from the generalized SDT model across tasks that share the same stimulus space or behavioral response. We show an example of this by analyzing choice data obtained from an ORE task and response time data obtained from an ORCA task.

This research was funded by NSF grant 2319234 awarded to Fabian Soto

23.472 PRINCIPAL DISTORTIONS FOR DISCRIMINATION OF IMAGE REPRESENTATIONS

Jenelle Feather[1,2,3], David Lipshutz[1,2,4], Sarah E. Harvey[2], Alex H. Williams[2,3], Eero P. Simoncelli[2,3]; [1]Equal Contribution, [2]Flatiron Institute, [3]New York University, [4]Baylor College of Medicine

Similarity between image representations is often quantified by measuring their alignment over a set of natural images that span many object categories, viewpoints, or environments. However, systems with comparable representational similarity measures on these sets of natural images can have strikingly different sensitivities to small stimulus distortions. We propose a framework for comparing a set of image representations in terms of their sensitivities to small distortions. We quantify the local geometry of a representation using the Fisher Information matrix (FIM), a standard statistical measure of sensitivity to stimulus perturbations, and use this to define a metric on the local geometry of representations in the vicinity of a base image. This metric may then be used to optimally discriminate a set of representations by defining a pair of "principal distortions" that maximize the variance of the representations under this metric. We use this framework to compare a set of simple models of the early visual system, identifying a pair of image perturbations that allow immediate comparison of the models by visual inspection, and naturally extend to psychophysical experiments measuring the discrimination thresholds for the perturbations. In a second example, we apply our method to a set of deep neural network models and reveal differences in the local geometry that arise due to architecture or training protocol. These examples demonstrate the use of our framework to elucidate informative differences in local sensitivities between complex computational models, and lay the groundwork for comparison of model representations with human perception.

23.473 SUBOPTIMAL VISUAL CUES INTEGRATION IN LOCATION ESTIMATION PARADIGM

Waragon Phusuwan[1,2] (waragonphoosuwan@gmail.com), Arp-Arpa Kasemsantitham[2,3], Chaipat Chunharas[2,4]; [1]Medical Science, Faculty of Medicine, Chulalongkorn University, [2]Cognitive Clinical & Computational Neuroscience Lab, King Chulalongkorn Memorial Hospital, [3]Faculty of Medicine, Chulalongkorn University, [4]Chula Neuroscience Center, King Chulalongkorn Memorial Hospital

Humans rely on multiple information to guide their actions and decisions in complex environments. Multisensory cue integration (e.g., using visual and proprioceptive inputs to guide action) has been extensively studied, mostly within the Bayesian information integration framework. However, little attention was given to cue integration within a single modality. This type of information integration is commonly deployed in navigation, where humans estimate goal locations and directions using visual cues such as faraway landmarks, signs, or even celestial objects like the brightest star in the sky. In this study, 9 participants played a simple computerized treasure-hunting game. For each trial, they had to estimate the treasure's location from 2 distinct visual shapes within 1 second. After the response was made or the time limit expired, the actual treasure location appeared as feedback. The locations of the two visual cues were generated by applying a horizontal shift from the randomized treasure location and adding jitter sampling from a normal distribution, with different standard deviations assigned to each shape. In this study, the two shapes (triangle and circle) were assigned with a standard deviation of 2 degrees and 8 degrees, respectively. This experimental setup allowed participants to make decisions based on the learned statistics of cues. We performed multiple linear regression with two visual cue locations as the predicate and the estimated location as the dependent variables to quantify prediction dependence on each cue location. The results showed a strong dependency of predicted locations on the reliable cue (2-degree deviation), with a coefficient of 0.88, std 0.06. However, they were weaker than the predictions from maximum likelihood estimation (MLE), with a predicted coefficient of 0.95. These findings highlight suboptimal visual cue integration and continue to raise questions about the factors influencing optimal performance in single-modality and multisensory integration.

I want to express my gratitude to Assoc. Prof. Dr. Xuexin Wei for guidance in data analyses, and Prof. Konrad Kording for the experimental design knowledge in Bayesian framework research. Lastly, I appreciate help from other CCCN lab members' contributions during the time of this study.

23.474 SUPERVISED QUADRATIC FEATURE ANALYSIS: TASK-SPECIFIC FEATURES THAT MAXIMIZE QUADRATIC DECODING

Daniel Herrera-Esposito[1] (dherrera1911@gmail.com), Johannes Burge; [1]University of Pennsylvania

Vision begins with the extraction of features in the retinal images that carry information useful for estimating and classifying properties of the environment. Here, we present a novel feature-learning technique called Supervised Quadratic Feature Analysis (SQFA) that maps class-conditional data (e.g. labeled image set) to a low-dimensional feature space that maximally preserves second-order class-differences, and thus second-order class discriminability. Differences in class-conditional second-order statistics can aid task performance, but few dimensionality reduction methods focus on second-order differences. SQFA exploits a theoretical relation between class discriminability (i.e. Fisher Information) and the Information geometry of second-moment (or covariance) matrices as points on the symmetric positive definite (SPD) manifold. The technique learns the features that maximize the distances between points (i.e. second-moment matrices) in this manifold. We demonstrate the usefulness of SQFA on a set of vision tasks where second-order statistics are known to be crucial. In many cases, SQFA finds features that are similar to the optimal features learned with Accuracy Maximization Analysis, a more computationally intensive approach. SQFA is distinct from other well-known methods in several respects. Unlike unsupervised techniques (e.g. ICA, PCA), SQFA learns task-specific features that are useful for classification and/or estimation. Unlike many supervised techniques (e.g. LDA), SQFA learns features that are sensitive to information other than first-order class differences. And unlike many non-linear and machine-learning methods (e.g. kernel methods, DNN's), SQFA results are easily interpretable. Thus, besides representing an encouraging first step in the use of Information-geometry-based methods for feature learning, SQFA should be a useful addition to the vision scientist's toolkit. Further, because SQFA can also be used to find features that exploit correlations in neural population activity, it should find broad application in neuroscience as well.

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.

Multisensory Processing: Visual-haptic and visual-vestibular integration

23.475 BEHAVIOURAL SIGNIFICANCE OF FOVEAL CORTICAL PROCESSING FOR HAPTIC SIZE ESTIMATION

Samantha Sartin[1], Domenico Dal Monte[2], Fabio Del Giudice[1], Laura Caleca[2], Elena Prosperi[2], Federica Carini[2], Federica Danaj[3], Irene Sperandio[2], Simona Monaco[1]; [1]Center for Mind/Brain Sciences, University of Trento, [2]Department of Psychology and Cognitive Science, University of Trento, [3]University of Regensburg

Human neuroimaging work shows that the foveal cortex is recruited during haptic exploration of objects, in the absence of visual input. Here we investigate whether haptic processing of object properties in the foveal cortex is relevant for behaviour. To this aim, we tested 30 individuals in a behavioral paradigm (Exp.1). Participants used their right-dominant hand to haptically explore one of three differently sized cylinders placed behind a monitor, while fixating a central cross. During the haptic exploration, dynamic visual noise (DVN) was presented in central vision (Noise condition) to disrupt haptic processing in the foveal cortex, if present. In a control condition, no DVN was present. Subsequently, participants were instructed to manually estimate the size of the explored stimulus. We then tested 34 volunteers (Exp.2) to investigate whether DVN effects on haptic size estimation may be explained by: 1) foveal processing, 2) the contribution of bimodal cells with overlapping visual and somatosensory receptive fields, 3) attentional distraction. To test these hypotheses, we adapted the previous behavioural paradigm such that in Noise trials the DVN randomly appeared in the center, on the right or left side of the screen. Results of Exp.1 show that size estimation is larger in the Noise as compared to the control condition, suggesting uncertainty when DVN is present. Further, standard deviations of size estimates increase with stimulus size, suggesting that haptic size processing, like visual processing, follows Weber's law. Results of Exp.2 reveal no significant effect of DVN location on performance, suggesting that none of the tested hypotheses can explain the results of Exp.1. While the results of Exp.1 suggest that the foveal cortex supports haptic size estimation, further research is needed to understand if the simultaneous interplay between multiple mechanisms could explain these results.

We thank Next Generation EU, MUR, and the University of Trento for their financial support.

23.476 THE INFLUENCE OF ADDITIONAL HAPTIC INFORMATION IN SPEEDED REACHING MOVEMENTS

Logan McIntosh[1], Robert Volcic[1,2,3]; [1]New York University Abu Dhabi, [2]Center for Artificial Intelligence and Robotics, New York University Abu Dhabi, [3]Center for Brain and Health, New York University Abu Dhabi

Previous research has shown that when visual information about the target location is complemented by haptic information, such as grasping a nail while striking it, movements tend to be faster and more precise. However, it is unclear whether the benefits of haptic information are also seen in speeded reaching movements when the target changes location unpredictably during the actual movement, or when the actual target among multiple target locations is revealed after movement onset. In this study, participants made speeded reaching movements in various conditions of potentially predictable or unpredictable target locations to one of two physical targets, placed 30 cm away and 6.25 cm to the left and right from the midline. The target location was cued at the start at of each trial through illumination of the monitor area underneath the target. However, on trials in which target location was unpredictable, the target location changed with 50% probability to the other location shortly after the movement was initiated, triggering an on-line movement correction. Movements were completed either purely by vision or with additional haptic information provided by holding the physical target in the left location, counterbalanced across different blocks. Thus, on trials with additional haptic information, sensory information about the left target location was provided by both vision and haptics. In contrast, sensory information about the right target location was derived exclusively from vision. We found that the addition of haptic information influences movement trajectories, with further differences between movements in predictable and unpredictable target location conditions. We also found idiosyncrasies in movement trajectories, with some participants showing a repulsion effect away from the haptic target location and others showing an attraction effect. Our study shows that the availability of additional sensory information affects the planning and execution of speeded reaching movements.

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.

23.477 TOUCHING SOUNDS: EXAMINING THE IMPACT OF NON-VISUAL EXPOSURE ON THE DEVELOPMENT OF AUDIO-TACTILE SOUND-SHAPE CORRESPONDENCES

Shibo Cao[1] (shibo.cao001@umb.edu), Rong Tan[1], Julia Kelly[1], Cuong Nguyen[1], Vivian M Ciaramitaro[1]; [1]University of Massachusetts Boston

In one type of crossmodal correspondence, nonsense words, such as "bouba", are associated with rounded abstract shapes, and "kiki" with angular shapes. Such associations are found between auditory and visual objects (AV), and between auditory and tactile objects, touched but not seen (AT). Visual experience can influence AT associations: AT associations are weak in early-blind adults (Fryer et.al, 2014) and fully-sighted 6-8-year-olds, with naïve visual experience (Chow et.al, 2021), and can be enhanced if fully-sighted children see the visual shapes first, prior AV exposure. Here, we examine how the type and amount of prior non-visual exposure influences AT associations. Children (N=122, 6-8-year-olds) completed 4 or 8 trials of a tactile attention (TT), or visual imagery (TI) task. For TT exposure, children felt a round and spiky shape inside a box, hidden from view, and then felt a smaller shape and judged which larger shape best matched the smaller shape. For TI exposure, children felt either a round or spiky shape inside a box and judged if they imagined the shape best resembled a flower or star. After TT or TI exposure, children completed 16 AT test trials: they felt a round and spiky shape inside a box and judged which shape best matched a nonsense sound. No feedback was provided during exposure or test trials. We found that 8 trials of prior TT, but not TI, exposure enhanced AT associations. However, neither 4 trials of prior TT or TI exposure enhanced AT associations, complementing our related findings that 8, but not 4, trials of prior AV exposure enhanced AT associations. Our results suggest that the type and amount of prior exposure is important in helping children understand how to represent objects experienced only via touch when forming sound-shape correspondences and that direct visual exposure may be sufficient but not necessary.

23.478 FUNCTIONAL NEURAL DYNAMICS OF VISUOMOTOR LEARNING: FUNCTIONAL CONNECTIONS BETWEEN SENSORY, MOTOR, MULTISENSORY, AND FRONTAL REGIONS REVEAL NETWORK FOR VISUO-PROPRIOCEPTIVE RECALIBRATION

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, [3]Ruhr University Bochum

Visual and proprioceptive (position sense) cues are used to estimate hand position. These cues can be spatially mismatched experimentally, e.g. by shifting a cursor away from the unseen hand, or naturally, e.g. viewing the hand underwater. This mismatch causes visuo-proprioceptive recalibration, a process that often appears in motor learning paradigms (i.e. visuomotor adaptation), but is rarely itself the focus of empirical study. This is problematic under the distributed processing perspective, where both sensory (e.g. primary somatosensory cortex S1, lateral occipital cortex LOC) and motor regions (e.g. primary motor cortex M1, ventral premotor cortex PMv, cerebellum) contribute to both sensory and motor behaviors, suggesting that to understand one it may be necessary to consider the other. Proprioception has been related to neural excitability in both M1 and S1; while multisensory (visual and proprioceptive) behaviors have been related to PMv and CB, as well as the multisensory anterior superior parietal lobule, aSPL. This evidence suggests the importance of investigating sensory processes in the context of motor behavior (and vice versa). Participants completed two blocks of visuo-proprioceptive pointing trials with resting state functional magnetic resonance imaging (fMRI) scans after each. Visual and proprioceptive cues were correctly aligned in the veridical block but gradually offset in the misaligned block to induce recalibration. Comparison between the resting state scans revealed the functional connections related to visuo-proprioceptive recalibration. Behavioral measures of visuo-proprioceptive recalibration were used in a seed-based analysis to identify a functional network with increased positive synchrony between CB, S1, striatum, LOC, and multisensory regions; and increased negative synchrony of M1, PMv, frontal areas, and multisensory regions related to recalibration. The cerebellum is also identified as a key region involved in visuo-proprioceptive recalibration, evidenced by its high global efficiency in our network, implicating the involvement of a traditionally motor region with visual recalibration.

23.479 INVESTIGATING VISUAL WEIGHTING DURING POSTURAL CONTROL USING VIRTUAL REALITY

Yingying Bei[1], Jeffrey Allen Saunders[1]; [1]The University of Hong Kong

Multiple sensory cues, including vision, vestibular, and proprioception, are available to maintain balance when standing. We investigated the contribution of visual cues to posture control by measuring physical responses to continuous pseudorandom perturbations in virtual reality (VR). We varied the spectral composition of the perturbations to manipulate visual uncertainty. Optimal integration predicts reduced reliance on vision with higher uncertainty. On a trial, participants maintained a standing posture for 4.5 minutes in a virtual room that continuously rotated in the roll direction. In the baseline condition, perturbations were superimposed waves with low frequencies: .085 Hz, .115 Hz, .155 Hz. In the added-oscillation conditions, perturbations included an additional wave with frequency of .35Hz or .70 Hz. Spectral analysis of head movements in the baseline condition revealed detectable responses at the perturbation frequencies. In the added oscillation conditions, there was little or no response to the higher frequencies but the response to low frequencies was reduced compared to baseline. This suggests a reduced reliance on visual information when the visual uncertainty was increased by the additional oscillations. None of the participants reported motion sickness in any of the conditions. Our results demonstrate that a continuous psychophysics approach can measure visual contributions to posture control with short exposures that are not sickness-inducing and that the presence of higher-frequency oscillations can reduce responses to lower-frequency oscillations.

23.480 WHAT OR WHERE/HOW DOES THE BODY-TILT ILLUSION OCCUR?

Sophia R. Baia[1] (srbaia@asu.edu), Michael K. McBeath[1,2]; [1]Arizona State University, [2]Max Planck Institute for Empirical Aesthetics

Introduction: Previous research found participants exhibit a systematic bias to overestimate the angle that their body is tilted, and that this bias increases under conditions where it is more costly or difficult to prevent falling. The current study investigates if this illusion is consistent with the “What” vs “Where/How” system by comparing body angle estimates of 45° to the maximum body angle at which individuals cannot prevent a fall by taking a step forward or backward. Method: 25 Participants leaned forward and backward to determine the maximum body-tilt angle before needing to take a step. They then had their body tilted in a body-sized gimbal gyroscope in which they estimated both when they felt tilted 45° and when they felt tilted at the angle needed to take a step. We also measured individual differences in balance ability as a potential mediating covariate. Results: We found that generally participants accurately estimated their body-tilt angle at which they would need to take a step (within 1° forward and 4° backward), but significantly underestimated when they were tilted 45° (by an average of 14°). We also found that individual balance metrics were not significant mediators of any body-tilt estimation effects. Discussion: Participants were accurate in estimating the angle of body-tilt needed to take a step, consistent with usage of the Where/How System (analogous to accurately estimating hill slant with a palm board). They also verbally underestimated their angle of 45° body-tilt, consistent with usage of the What System (analogous to verbally overestimating hill slant angle). This pattern of findings along with the lack of individual differences due to balance ability and the previous pattern of findings support that the body-tilt illusion is a functional distortion that likely helps people better keep their balance with little cost due to feeling more tilted than actual.

23.481 MOTION PERCEPTION BY A MOVING OBSERVER

David Yu[1], Guang Yang[1], David Alais[1], Reuben Rideaux[1], Frans Verstraten[1]; [1]The University of Sydney

Typically, motion-sensitivity studies are performed with observers sitting in darkened rooms, often constrained by forehead support and a chinrest. This is fine if one is interested in the sensitivity of the substrates that underlie motion detection and perception. However, it has become increasingly clear that vestibular input when observers are moving plays an important role in the way we perceive the world (see Davidson et al, Nature Communications, 2024). In a series of experiments, we investigated motion sensitivity under conditions where the observer is being moved. We used a CKAS W25R motion simulation platform and had observers perform a left-right motion direction discrimination task (2AFC). The participants were either stationary or were being moved by the motion platform. To make sure the vestibular system was active, the motion platform was accelerating when the task was performed. The stimulus was a random pixel array moving at different speeds. We used a signal-to-noise-ratio paradigm (see Fredericksen et al. Vision Research, 1993), where random pixel noise is added when the observer indicated the correct motion direction 3 times in a row. A single mistake decreased the noise level, resulting in a threshold level of 79% correct. There were 4 main conditions, where the motion platform is either stationary, moving in the same direction as the motion on the display, moving in the opposite direction (either translate or rotate), or moving vertically perpendicular to the stimulus direction. Surprisingly the results show that vestibular input has only little effect, if any, on the motion sensitivity thresholds under these conditions.

23.482 STATIONARITY PERCEPTION DEPENDS ON VESTIBULAR-OCULOMOTOR ADAPTATION

Paul MacNeilage[1] (pmacneilage@unr.edu), Robert S. Allison[2], Val Rodriguez[1]; [1]University of Nevada, Reno, [2]York University

As we move, the images on our retinas move due to head and eye movement, but under normal circumstances we tend to perceive the world as stationary. However, perception of world stationarity will be compromised if there is a significant discrepancy between observed and expected optic flow. This occurs, for example, when getting used to new spectacles, but we normally adapt to these changes. Here we studied the relationship between oculomotor adaptation and perceived stationarity. We adapted observers' vestibular ocular reflex (VOR) by exposing them to a virtual environment where a fixation point in an otherwise empty world moved with them when they turned their head ~15 deg to the left or right; movement direction on each trial was cued and randomized. The gain of dot motion decreased from 1 to 0.6 over the course of 150 trials. Trials with no fixation point (dark environment) were interspersed to measure VOR adaptation state. Joint analysis of eye and head movements on these trials indicated that many observer's VOR adapted toward the targeted gain of 0.6. Before and after adaptation we measured each observer's point of subjective stationarity (PSS), that is the scene motion gain perceived as stationary. On each trial, participants made similar cued head movements in a virtual world composed of a 3D cloud of spheres with a central fixation point and then judged whether the scene appeared to move with or against their head motion. Gain was varied trial to trial according to adaptive staircases to estimate the PSS. Before VOR adaptation, average PSS was not significantly different from a gain of 1, but afterwards, average PSS was significantly less than 1 ($p=0.03$), indicating that VOR adaptation state influences stationarity perception. This suggests that it may be useful to monitor VOR adaptation state in users of extended reality systems.

Supported by the Canada First Research Excellence Fund (CFREF) under the Vision: Science to Application (VISTA) program.

SATURDAY AFTERNOON POSTERS IN BANYAN BREEZEWAY

Eye Movements: Models, clinical, context

Saturday, May 17, 2:45 – 6:45 pm, Banyan Breezeway

26.301 DECODING EYE CLOSED GAZE POSITION USING DEEPMREYE

Sina M Kling[1], Uriel Lascombes[1], Matthias Nau[2], Guillaume S Masson[1], Martin Szinte[1]; [1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Université, Marseille, France, [2]Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Eye movements, even with eyes closed, provide valuable insights into human cognition and are a critical variable in numerous functional magnetic resonance imaging (fMRI) studies. Here, we track eye movements while the eyes are closed using DeepMReye, a deep learning framework for camera-less eye tracking in fMRI. We designed an experiment where participants moved their gaze towards a sequence of known positions, with both eyes open and closed, under conditions with and without visual input, while fMRI data were being acquired. While DeepMReye was initially trained on fMRI data from classical eyes-open tasks, the network could successfully generalize to decode gaze position during eyes-closed periods. Furthermore, model performance improved when the network was specifically trained to generalize across conditions, including eyes-closed and varying visual input. These findings demonstrate that reliable eye movement monitoring during eyes-closed periods is feasible in fMRI, enabling a more effective integration of eye tracking in fMRI research and therefore advancing our understanding of human cognition.

This research was supported by an ANR JCJC and a Fyssen Foundation grant to MS.

26.302 A NEW TECHNIQUE FOR UNDERSTANDING HOW THEORY-BASED FACTORS COMBINE TO CONTROL EYE MOVEMENTS

Abe Leite[1] (abrahamjleite@gmail.com), Gregory J. Zelinsky[1]; [1]Stony Brook University

In recent years, the performance of models using deep learning architectures to predict human eye movements during free-viewing and search has approached the ceiling imposed by inter-observer agreement. In parallel, a more theory-driven set of studies have investigated how much of gaze behavior can be explained by theoretically-relevant factors like visual salience or object recognition uncertainty. With some exceptions, this work has been limited in two ways: (i) it typically compares models incorporating one factor each, and (ii) it typically does not yield probabilistic models that make concrete predictions about behavior. We introduce a novel statistical method to understand how multiple theoretically important factors are integrated to produce free-viewing behavior. Based on Kümmerer, Wallis, and Bethge (PNAS 2015)'s technique of "phrasing saliency maps probabilistically", in which a probability map is obtained by applying a piecewise-linear monotonically increasing function to values of a single factor, our novel "probabilistic signal integration" combines multiple factors using a piecewise-linear function monotonically increasing in all of its inputs. The standard softmax function is then applied to obtain a predictive probability distribution. Our approach allows multiple factors influencing fixation to be integrated, under the strong but reasonable assumption that the probability of fixating a point must increase to some extent whenever any predictive factor increases. No parcellation of fixations is required. Applying this approach to many factors allows us to see how much a model's likelihood drops when each factor is omitted (like nested GLM) in order to assess how much information each factor contributes uniquely to the model. Unlike GLM, we do not assume a linear relationship (only a monotonic one) for either main effects or interactions. We present an initial test of our multi-factor integration approach by applying it to the free-viewing data analyzed by Chakraborty, Samaras, and Zelinsky (JoV 2022).

This project is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2234683.

26.303 EYE-TRACKING-BIDS: THE BRAIN IMAGING DATA STRUCTURE EXTENDED TO GAZE POSITION AND PUPIL DATA

Martin Szinte[1], Dominik R. Bach[2], Dejan Draschkow[3], Oscar Esteban[4], Benjamin Gagl[5], Rémi Gau[6], Klara Gregorova[7], Sina Kling[1], Sourav Kulkarni[2], Julia-Katharina Pfarr[6,8]; [1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Université, Marseille, France, [2]University of Bonn, Transdisciplinary Research Area "Life & Health", Hertz Chair for Artificial Intelligence and Neuroscience, Bonn, Germany, [3]Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom, [4]Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland, [5]Self-Learning Systems Lab, University of Cologne, Cologne, Germany, [6]Neuro Data Science ORIGAMI Laboratory, McConnell Brain Imaging Centre, Faculty of Medicine, McGill University, Montréal, Canada, [7]Department of Psychology and Sports Sciences, Goethe University, Frankfurt am Main, Germany, [8]Philipps-University Marburg, Department of Psychology, General and Biological Psychology, Marburg, Germany

The Brain Imaging Data Structure (BIDS) is an extensive community effort to establish data and metadata management standards. Due to its enormous success in human brain imaging research, numerous extensions have been developed. Here, we present Eye-Tracking-BIDS, an extension to BIDS for gaze position and pupil data. This community-driven extension aims to enhance practices and transparency within the eye-tracking community, as well as to facilitate the effective integration of eye-tracking data within fMRI, EEG, MEG, and behavioral standardized datasets. It includes standardization of raw synchronous gaze position and pupil data, asynchronous model parameters and trigger messages, as well as detailed metadata information. Compliance with the Eye-Tracking-BIDS by the concerned research community will foster the development of efficient, robust, open, and automatized BIDS-Apps for pre- and post-processing of eye-tracking data, thus enhancing reproducibility and replicability in the field.

26.304 REDUCED INHIBITORY CONTROL IN PARKINSON DISEASE AS REVEALED BY SACCADIC INHIBITION

Celeste Cafaro[1] (celestecafaro97@gmail.com), Alessio Fracasso[2], Ciro Rosario Ilardi[3], Marco Cerrone[4], Giovanna Vermiglio[5], Carlo Cavaliere[4], Giovanni Cirillo[6], Antimo Buonocore[1]; [1]Suor Orsola Benincasa University, Italy, [2]University of Glasgow, [3]University of Naples Federico II, Italy, [4]IRCCS Synlab SDN, Italy, [5]University of Messina, Italy, [6]University of Campania Luigi Vanvitelli, Italy

Oculomotor deficits are a hallmark of Parkinson's disease (PD) and provide important insights into its associated visual, motor, and cognitive dysfunctions. For instance, antisaccade tasks—which require redirecting gaze to the opposite location of a cued target—have been widely used to describe reduced inhibitory control. However, the impact of PD on reflexive oculomotor inhibition, independent of cognitive control, remains unclear. We hypothesized that saccadic inhibition (SI), which relies on oculomotor structures known to be impaired in PD patients, may more accurately reflect disease progression. These measures are less influenced by cognitive abilities, making them promising candidates as biomarkers for the condition. To test this hypothesis, we assessed eye movements in eleven PD patients (including three de novo cases) and six healthy controls (HCs) across classic oculomotor tasks, including visually guided saccades, prosaccades, antisaccades, smooth pursuit, and visual search. The latter two tasks were modified by introducing visual transients to evaluate SI. As expected, PD patients exhibited hypometric saccades and slower reaction times in visually guided saccade tasks compared to HCs. In the antisaccade task, PD patients failed to inhibit responses to cued targets and displayed high inter-participant variability. Smooth pursuit performance showed reduced gain and an increased intrusion of catch-up saccades. Interestingly, both groups exhibited fast SI in response to visual transients, suggesting preserved reflexive inhibitory abilities in PD. However, SI in PD patients was weaker and lasted longer, followed by a diminished rebound phase, suggesting deficits in oculomotor reprogramming. A similar SI pattern was observed for catch-up saccades during smooth pursuit. Unlike the antisaccade task, reflexive inhibitory measures were consistent across PD patients and distinguishable from HCs, including the de novo cases. These findings suggest that SI measures in visual search tasks may be effective for the early detection of deficits in PD and potentially other neurodegenerative disorders.

C.C, G.V., G.C, A.B. were supported by a grant from the Italian Minister of Research and University (PRIN PNRR 2022 P2022ST78T).

26.305 LIMBUS SEGMENTS ARE A POTENTIAL ALTERNATIVE TO P-CR SIGNALS FOR GAZE ESTIMATION

Kevin Willeford[1] (kwillefo@nova.edu); [1]NOVA Southeastern University College of Optometry

Purpose. The limbus remains visible during all ocular rotations. Thus, the goal of our experiment was to evaluate whether gaze estimates based on limbal segments (LS) are as accurate as those derived from P-CR signals, and if so, to suggest that accurate gaze estimation is tractable for large eye movements. Methods. Twelve non-strabismic observers, head-stabilized with bite-bars, performed monocular fixations (5°- 30° in eccentricity along the 45° - 360° meridians) while P-CR signals (Eyelink 1000+) and ocular photographs were recorded. Limbal segments were extracted from photographs by first using a neural network to segment the iris and sclera and then defining the limbus as their common boundary. Next, an optimization routine transformed a reference limbus, obtained from photos captured in primary gaze, until the mean distance between points on the reference limbus and a given limbal segment was minimized. This approach produced transformation attributes (xSCALE, ySCALE, rotation, xTRANSLATION, yTRANSLATION) representing the unique shape and location of each observer's limbi across all conditions. Two calibration functions, one using the transformation attributes and the other using P-CR signals, were created by using third-order polynomials to map each set of data to known fixation locations. Distance between predicted and actual fixations was used as a measure of accuracy. Results. P-CR tracking produced more accurate gaze estimates for right- [0.0124 +/- 0.0028 mm (P-CR) vs. 0.0204 +/- 0.0030 mm (LS), $p = 0.04$] but not left- [0.0238 +/- 0.0046 (P-CR) vs. 0.0199 +/- 0.0031 (LS), $p > 0.05$] eyes. This did not result from a difference in mapping quality (all $R^2 > 0.88$) or from one method's bias for greater error in a specific location. Conclusions. Limbal segments are potential proxies for measuring large eye movements, overcoming the recording range limitations inherent to P-CR trackers and capturing the eye's full range of motion.

This research is supported by a President's Research Grant from Nova Southeastern University.

26.306 RELATIONAL MEMORY FOR OBJECTS IN SCENES: EYE MOVEMENTS REVEAL EFFECT OF THE EXTENT OF SCENE REPETITION

Josefine Albert[1,2] (josefine.albert@uni-bielefeld.de), Werner X. Schneider[1,2], Christian H. Poth[1,2]; [1]Neuro-Cognitive Psychology, Bielefeld University, Germany

When a scene changes between viewing instances, humans make relatively more fixations to, dwell longer on, and make more transitions in and out of the altered object's area. This is typically reflected in differences between novel, repeated, and manipulated scenes. Here, we investigated whether not only the existence but also the extent of learning history prior to the alteration of a scene affected viewing behavior. Experiment 1 adapted a paradigm by Ryan et al. (2000) to establish the memory effects with computer-generated scenes and high-resolution eye-tracking. Subjects viewed a set of scenes for a later memory test. Scenes were repeated, manipulated (object shifting or addition/deletion), or novel. Implicit relational memory was evident for shifting in all measures and for additions in most measures. In Experiment 2, we varied the extent of learning by repeating the scenes 1, 2, 4, or 8 times before introducing the manipulation. Based on the first experiment's results, all scenes underwent a shifting manipulation. We calculated the difference in the viewing parameters of the critical object between the presentation before the change and the change presentation. The number of repetitions prior to the scene's alteration had a significant effect. Participants increased their dwell time, fixation proportion, and number of transitions more when they had viewed the scene for 4 or 8 times than for 1 or 2 times. Eye movement measures correlated moderately, but insignificantly, with participants' explicit memory. This suggests that the investigated eye movement measures may capture memory strength rather than mere memory presence but their suitability for memory strength assessment is discussed.

26.307 EYE MOVEMENT PATTERNS SERVE AS INDICATORS OF IMPLICIT KNOWLEDGE ACQUISITION IN VISUAL ARTIFICIAL GRAMMAR LEARNING

Irina Lavrova[1], Anto J. Mohan[2], Miranda Scolari[3]; [1]Texas Tech University, [2]Texas Tech University, [3]Texas Tech University

Implicit learning is a phenomenon by which the extraction and acquisition of structural regularities occur unintentionally and without conscious awareness. This study examined whether eye movement patterns, which function as involuntary indicators of acquired knowledge (Silva et al., 2017), can be a reliable measure of implicit knowledge acquisition. A visual artificial grammar learning (AGL) task was used to explore the formation of implicit knowledge following a brief exposure to grammatical sequences. Sixty-five participants first made a same/different judgment for each presented pair of words created using artificial grammars of varying complexity. Importantly, they were not instructed to encode the words. Next, participants completed a surprise grammaticality judgment task, in which they decided whether each singly presented word followed the artificial grammar rules encountered in the first phase. Eye tracking revealed that participants exhibited eye movements patterns consistent with implicit learning. Dwell time was greater on target letters that violated the grammatical rules compared to non-target controls ($\chi 2 = 4.57$, $p = 0.032$), and targets were similarly associated with a greater number of regressions ($\chi 2 = 24.89$, $p < 0.001$). Grammar complexity did not significantly influence eye movements, despite previous studies showing it negatively impacts performance on behavioral AGL tasks (Van Den Bos & Poletiek, 2008). Furthermore, eye movement patterns were not linked to grammaticality judgment accuracy, indicating that dwell time and regressions to target locations reflect an earlier, pre-awareness stage of processing. These findings highlight that eye-tracking can effectively capture implicitly acquired knowledge, revealing participants' sensitivity to grammatical structures even in the absence of conscious discrimination between grammatical and nongrammatical sequences. This underscores the role of perceptual processes in implicit learning and positions eye-tracking as a powerful tool for studying the unconscious acquisition of structural regularities.

Object Recognition: Reading

SATURDAY, MAY 17, 2:45 – 6:45 PM, BANYAN BREEZEWAY

26.308 ARTIFICIAL FIXATION POINTS- ARE THEY GENERALIZABLE?

Ali Bhimani[1] (alin.bhimani@gmail.com), Jessica L. Parker[1], A. Caglar Tas[1]; [1]The University of Tennessee, Knoxville

We previously showed that including artificial fixation points (AFPs) help readers by decreasing reading speed and increasing reading comprehension (Bhimani & Tas, VSS 2024). In our previous study, AFPs were implemented in texts based on eye data collected during normal reading: We changed the font color of the first half of highly-fixated words to red, thus ideally cueing participants to the most efficient areas to fixate during reading. Texts that were altered with AFPs resulted in faster reading times and higher reading comprehension than texts that were not altered. In the present study, we tested the effectiveness and generalizability of AFPs. Experiment 1 used font color changes as in our previous study, but instead of modifying highly-fixated words, we changed the font color of the first half of randomly chosen words. If AFPs merely increased intrinsic motivation, rather than aiding reading, we should find similar accuracy and reading speed improvements as in the previous study. Experiment 2 used a different AFP manipulation, boldness, to test its generalizability. As in the previous study, we put AFPs on highly-fixated words. In both experiments, AFPs significantly decreased accuracy and slowed down reading speed (all ps <.001), suggesting that while text altering can benefit reading speed and improve reading comprehension, the type of AFPs used and whether they align with normal reading behavior is an important factor in their effectiveness.

26.309 DISSOCIATION BETWEEN THE PROCESSING OF SPATIAL FREQUENCIES VS SCALE IN VISUAL WORD RECOGNITION

Clémence Bertrand Pilon[1,2], Martin Arguin[1,2]; [1]Département de psychologie and Centre interdisciplinaire de recherche sur le cerveau et l'apprentissage, Université de Montréal, [2]Centre de recherche, Institut Universitaire de Gériatrie de Montréal, Montréal, Canada.

Two consensual theories in visual recognition focus on the temporal progression of spatial processing: the “global-to-local” theory, which suggests the human visual system processes large features first, followed by smaller ones (Navon, 1991), and the “coarse-to-fine” theory, which suggests that we first process low spatial frequencies (SFs), then higher ones (Bar, 2003). While these theories are often considered equivalent, spatial scale (SS) and SF are in fact theoretically dissociable : e.g. large features can rest on high SFs. The present study investigated the temporal progression of SF and SS processing in visual word recognition using random temporal sampling. Adult readers (16 per experiment) reported a word presented for 200 ms in a four-alternative forced-choice task. Target stimuli were an additive combination of the target word and a white noise field with a randomly varying signal-to-noise ratio (SNR) over exposure duration. In Exp. 1 stimuli were filtered according to four SF conditions (central frequencies of 1.2, 2.4, 4.8, and 9.6 cycles per degree). Exp. 2 were intact word images behind an occluding mask with gaussian apertures of varying sizes (center frequencies of 1.11°, 0.55°, 0.28°, 0.14°). Classification images (CIs) were computed to show the temporal progression of SF and SS processing efficiency. In Exp. 1, initial visual processing relied on high SFs, then followed by lower SFs. Exp. 2 revealed an initial dominance of the largest SS, then progressing orderly to smaller ones. Both experiments showed complex interactions of SF/SS processing over time with the frequency spectrum of SNR oscillations . In contrast to the "coarse-to-fine" order expected for SF processing in word recognition, Exp. 1 demonstrated a "fine-to-coarse" progression. The results of Exp. 2 , revealed a "global-to-local" order. These results highlight the importance of distinguishing between SF and SS, since the two are empirically dissociable.

26.310 FONT SIZE, X-HEIGHT, AND READABILITY: HOW TYPOGRAPHY AFFECTS READING SPEED ACROSS DIFFERENT SIZES

Ben Sawyer[1,2] (sawyer@ucf.edu), Md Manumur Rashid[1,2], Hilary Palmen[3], Mert Kucuk[1], Nilsu Atilgan[1]; [1]The Readability Consotium, [2]The University of Central Florida, [3]Google Fonts

The relationship between print size and reading speed is well-established, with larger sizes resulting in larger image on the retina. This benefits populations such as older adults and individuals with visual impairments (Legge et al., 1985). However, in reading research, "size" is sometimes discussed in terms of x-height, which refers to the height of lowercase letters in a font. Importantly, x-height and print size are distinct; different fonts can have varying x-heights even at the same nominal size, potentially influencing readability. To investigate this distinction, we examined the combined effects of print size and x-height on reading speed. Participants (N = 60) read passages at varying sizes and x-heights, followed by comprehension questions. A repeated-measures ANOVA revealed that, as expected, print size significantly impacted reading speed ($p < .05$), while x-height alone did not ($p = .134$). To explore the relationship further and account for individual variability, we conducted a linear mixed-effects model analysis. This deeper analysis revealed an interaction between size and x-height. Specifically, when print size was small, participants read marginally faster with larger x-heights (mean = 248 WPM) compared to smaller x-heights (mean = 232 WPM). This effect was not observed at larger print sizes. These findings confirm that print size is a critical factor in reading speed across all conditions. Additionally, x-height plays a role in enhancing readability when print size is small and reading becomes more challenging. This insight is particularly relevant in scenarios where increasing print size is not feasible; selecting fonts with larger x-heights may offer an alternative strategy to improve readability.

The Readability Consortium

26.311 IMPAIRED LOCAL PERCEPTION FOR VISUAL WORD FORMS IN DEVELOPMENTAL DYSLEXIA

Shao-Chin Hung[1], Hsiao-Lan Sharon Wang[2], Chien-Chung Chen[1]; [1]National Taiwan University, [2]National Taiwan Normal University

The classic Navon task demonstrates that humans identify global figures faster than local ones. This global precedence is challenged in dyslexics as they present a local before global perception. Here we investigated the global and local perception in 18 developmental dyslexic and 16 neurotypical Mandarin-speaking juvenile (aged 12-17 years). We employed a novel Navon task in which participants judged whether a global figure (in a global task) or local figures (in a local task) of stimuli presented on both sides of the central fixation were identical. The stimuli were large geometric shapes (i.e., circles, diamonds, or triangles) composed of smaller ones in Experiment 1. The stimuli in Experiment 2 were created similarly with three Chinese characters. In both Experiment 1 and 2, all participants showed a global precedence effect with increased accuracy in the global tasks than the local tasks. In Experiment 1, for either global or local tasks, both groups showed a congruency effect in which they responded more accurately while the local and global geometric shapes were the same than when they were different. In Experiment 2, both groups exhibited the congruency effect for the local tasks but not for the global tasks, indicating an interference from the global level to the local level, but not vice versa. Compared to the neurotypical juvenile, the dyslexics showed worse local perception possibly due to their susceptibility to visual crowding. Further, the inferior performance of dyslexics in local perception was more prominent when viewing Chinese characters than geometric shapes. In sum, we found no evidence for the local precedence either for geometric shapes or Chinese characters in dyslexia. The impaired local perception of Chinese characters observed in dyslexics may suggest specialized channels for visual word form processing.

NSTC 113-2410-H-002-251-MY3

26.312 LINKING FONT LEGIBILITY AND PREFERENCE

yeha cha[1], Kim Chai; [1]Korea University

Legibility plays an important role in font selection (Arditi & Cho, 2005, Bigelow, 2019, Minakata et al, 2023) and preference for fonts. (Grobelny & Rafał Michalski, 2014, Doyle & Bottomley, 2004). However, existing research has not systematically examined factors that increase preference, nor their relevance with legibility. To identify factors that can increase legibility and preference in fonts to examine their relationship we manipulated. The independent variables in this study are size, serif, and contrast between thin and thick strokes, which are systematically adjustable and decorative elements of a font (Amare & Manning, 2012). At two letter size levels (0.3º, 0.5º), three contrast levels of serif (0 [sans-serif], medium, long) and three levels of contrast (low, medium, high) were combined with variations of the Noto serif font, resulting in a total of nine fonts. A lexical decision task was applied to 6-character strings using these fonts, and legibility was measured by calculating the proportion of correct responses. Under the same conditions, preference was measured using a 4-point Likert scale by increasing the duration of the string presentation. Results showed that regarding legibility, larger fonts were more legible than smaller fonts. Regarding preference as well, larger fonts were preferred to the smaller fonts. For the small font of 0.3º, sans-serif was least legible at high contrast, followed by low contrast, and most legible at medium contrast ($p<0.05$). For medium and long serifs, with higher contrast, the legibility got higher ($p<0.001$). For the large fonts of 0.5º, there were no significant legibility differences for all serifs and contrasts. At both font sizes preference increased with contrast ($p<0.001$). Last but not least, legibility and preference showed a positive linear relationship (smaller; $r2= 0.8196$, larger; $r2 = 0.3354$). These results suggest that fonts with higher legibility are preferred, in case of small fonts.

NRF-2023R1A2C2007289

26.313 THE IMPACT OF FONT ON TYPO DETECTION: A NOVEL VISUAL SEARCH PARADIGM

Emily Heffernan[1], Benjamin Wolfe[1], Anna Kosovicheva[1]; [1]University of Toronto

In visual search tasks, we scan our environment to identify a target (e.g., looking for your suitcase at the Tampa airport baggage claim). A better understanding of the factors that lead to success and failure in visual search requires a task that mimics our lived experiences but permits a high degree of experimental control. Here, we developed a novel typo detection paradigm to uncover how attentional processes interact with visual properties of stimuli in a word-based visual search task. Participants (N=9) scanned pseudo-paragraphs comprised of random 5-, 6-, and 7-letter words to find typos (i.e., incorrectly spelled words), which were present on 50% of trials. Five types of typos were included: transpositions (two letters swapped), insertions (a letter added to a word), deletions (a letter removed from a word), repetitions (a letter repeated), and substitutions (a letter replaced with another letter). In addition, half of the trials were presented in an “easy-to-read” font (Arial) and half were presented in a “hard-to-read” font (a version of Roboto Flex with narrow width and a high stroke contrast). Font had a main effect on reaction time: participants responded more slowly when the stimuli were presented in the hard-to-read font. Font had no overall impact on accuracy. However, participants were worst at identifying transposition errors, and for these trials, font did have a significant effect, such that performance was lower for the hard- versus easy-to-read font. These findings also highlight substantial individual differences in performance and sensitivity to font manipulation. Taken together, these results indicate that the appearance of text does impact visual search for typos, but only for specific types of errors. This paradigm can elucidate how constraints in peripheral vision (e.g., crowding) impact visual search performance for text-based stimuli.

Funding Acknowledgments: This work was supported by a SSHRC Insight Grant to BW and AK.

26.314 PSYCHOPHYSICS OF VARIABLE FONTS: DO MULTIPLE FONT FEATURES INTERACT TO IMPACT READABILITY?

Silvia Guidi[1], Anna Kosovicheva[1], Benjamin Wolfe[1]; [1]University of Toronto Mississauga

When choosing a font, we have some intuitive understanding of why a particular font may feel easier to read, but what elements of a font actually affect readability? To answer this question, we used variable fonts, in which every element, such as the width or stroke contrast of each letter, can be adjusted on a continuous axis. Previously, we have shown that changes within a single axis can change saccade amplitude and reading duration thresholds (Guidi et al. VSS2024). In a new study, we examined how these axes impact readability in combination by manipulating text appearance on two axes, thin stroke and width, at three levels per axis across the full range, for a total of 9 conditions. Participants read a series of sentences in each font condition while gaze position was tracked, classifying each sentence as true or false. Sentence presentation duration was staircased and we calculated duration thresholds needed for 80% classification accuracy for each condition. Thicker thin strokes decreased duration thresholds across all width settings, while the thinnest thin strokes resulted in the highest duration thresholds (i.e., the slowest reading performance). These extreme thin strokes impacted reading speed regardless of the width of the text. Eye tracking data revealed that participants partially compensated for increased text width by increasing their saccade amplitudes. By understanding how different font elements interact with each other, we may be able to understand what parts of text presentation affect readability the most, which can then be used to help maximize reading efficiency.

This work was supported by a SSHRC Insight Grant to AK and BW.

26.315 QUANTIFICATION OF WHITE-MATTER CONNECTIVITY IN THE VENTRAL OCCIPITOTEMPORAL CORTEX

Yongning Lei[1] (t.lei@bcbl.eu), Pedro M. Paz-Alonso[1,2], Garikoitz Lerma-Usabiaga[1,2]; [1]Basque Center on Cognition, Brain & Language, [2]Ikerbasque, Basque Foundation for Science, Bilbao, Bizkaia, Spain

Reading is a rapid process. Within hundreds of milliseconds, perceptual input from the eyes is integrated by the early visual cortex and processed by higher-order visual and language networks, including the ventral occipitotemporal cortex (vOTC) and the inferior frontal gyrus (IFG). Electrophysiological evidence has shown that the IFG becomes active during reading within 200 milliseconds. This early-stage activation suggests the presence of direct white-matter connections between visual territories and the IFG. Previous research has identified four major white-matter tracts linking the vOTC and IFG, however, the detailed connectivity profiles between subregions of these visual and language circuits remain poorly characterized, partly due to the limitations of diffusion-weighted imaging (DWI). In this study, we used a high-quality diffusion MRI dataset (N=66, plus 31 retested) to investigate white-matter connectivity patterns between subregions of the ventral visual pathway and the IFG (pars opercularis, triangularis and orbitalis). To address methodological challenges, we applied a novel approach combining two tract reconstruction algorithms: ROI-to-ROI streamlines and whole-brain tractography. This approach allowed us to delineate subregion-specific white-matter connections without compromising their anatomical accuracy. Our results revealed that the pars opercularis is primarily connected to the anterior vOTC dorsally via the arcuate and superior longitudinal fasciculi. In contrast, the pars orbitalis exhibited exclusive ventral connectivity to the vOTC through the inferior fronto-occipital and inferior longitudinal fasciculi. The pars triangularis displayed a mixed connectivity pattern, with the posterior vOTC connecting ventrally and the anterior vOTC dorsally. In sum, our findings provide a detailed characterization of the structural connectivity underlying reading networks, offering a framework for exploring the white-matter connectivity and functional connectivity, and providing insights into the computational theories of reading.

BCBL acknowledges the support of 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

26.316 SIMULTANEOUS RESPONSE, WITHOUT SUPPRESSION, TO MULTIPLE WORDS WITHIN THE VISUAL WORD FORM AREA

Vassiki Chauhan[1], Alex White[1]; [1]Barnard College, Columbia University

The visual word form area (VWFA), within the left occipitotemporal cortex, is known to be critical for reading. We have learned much about this region's function by recording responses to single words. But when we read, we typically see multiple words simultaneously. We don't know how the VWFA's activity changes with the number of words being read, and whether the representation of each word might be affected by neighboring words. Here, we fill that gap by varying the number of words presented, while holding visual stimulation constant. We adapted the "simultaneous suppression" paradigm, which has shown that in other visual areas, the response to multiple stimuli presented simultaneously is smaller than the sum of responses to the same stimuli presented sequentially. Participants viewed rapid sequences of three frames (total duration = 700 ms). Each frame contained two character strings, one above and one below fixation. Most of those were unfamiliar pseudoletters with visual features matched to words in Courier New. There were four conditions: (1) Zero words; (2) One English word appearing at one position; (3) Two words presented sequentially (at different positions in different frames); (4) Two words presented simultaneously (in the same frame). We recorded activity with fMRI while 22 participants viewed these stimuli and reported how many words they perceived in each trial. The VWFA's response magnitude increased linearly with the number of words presented. The response did not significantly differ between the sequential and simultaneous two-word conditions, although participants performed worse at detecting two simultaneous words. Moreover, across all conditions, the BOLD response varied with the lexical frequency of each word, including for both of the two words presented simultaneously. In sum, the VWFA's activity scales with the number of words presented, and it is sensitive to the lexical properties of two words presented simultaneously.

26.317 THE IMPACT OF FONT WEIGHT ON EXPRESSIVENESS BEYOND LATIN: INSIGHTS FROM ARABIC AND ENGLISH

Nilsu Atilgan[1], Hilary Palmen[2], Mert Kucuk[3], Ben D. Sawyer[1]; [1]University of Central Florida, [2]Google, [3]Bogazici University

The influence of text properties like character spacing and font on readability has been extensively studied in Western languages using Latin scripts (Beier et al., 2022). Recent advancements in variable font technology now allow investigations into specific font features, such as weight, a highly impactful axis in typography (Dobres et al., 2016). Despite the global use of digital documents, non-Latin scripts remain underexplored. To address this gap, we examined how font weight affects the perceived expressiveness of reading materials in English and Arabic. Using Noto Sans, a typeface with extensive language coverage, 50 bilingual participants read sentences adapted from MNREAD (Mansfield et al., 1993) in both languages. After reading, participants selected adjectives from a list of 21 descriptors (e.g., loud, playful, calm, active, sophisticated) to describe the font. Noto Sans was presented at varying weights to examine the impact of font weight and writing system on adjective selection. A general linear model analysis revealed no significant main effect of the writing system ($p = .139$), indicating similar emotional responses across languages. Font weight significantly influenced adjective selection ($p < .001$). Lighter weights were associated with calmness ($p < .001$), while bolder weights were perceived as more active ($p < .001$) in both writing systems. Notably, an interaction effect was observed for certain expressions; for instance, bold Arabic text was rated as more childlike, a pattern not found in English. In conclusion, when designing typefaces or selecting fonts for digital materials, font weight can be a powerful tool to convey specific messages. Our findings reveal parallels between English and Arabic scripts in how weight influences expressiveness, while also highlighting unique interactions. This exploratory study emphasizes the need for the reading research community to investigate the impact of typefaces beyond the Latin alphabet, fostering a broader understanding of cross-script typographic design.

26.318 THE INFLUENCE OF UNFAMILIAR LETTER STRING ORIENTATION ON CROWDING EFFECTS IN JAPANESE READING

Kazushi Maruya[1] (kazushi.maruya@ntt.com), Miki Uetsuki[2]; [1]NTT Communication Science Laboratories, [2]Aoyama Gakuin University

Crowding in strings impairs letter identification (Whitney & Levi, 2011), but the effect is reduced for letter strings compared to symbol strings (Grainger et al., 2010). In Latin script readers, this reduction occurs only in horizontally oriented letter strings (Vejnović & Zdravković, 2015). In contrast, Japanese readers exhibit this reduction in both vertically and horizontally oriented strings (Uetsuki & Maruya, under review). These differences may reflect variations in reading experience with specific string orientations. This study investigated whether Japanese readers also exhibit crowding reduction for letter strings oriented in unfamiliar directions, such as diagonal strings. Nine native Japanese readers participated in the experiment. Stimuli included three types of letters: English alphabet letters, Japanese katakana letters, and symbols. Each trial began with a fixation point, followed by a target letter displayed for 100 ms, either alone (isolated condition) or flanked by two distractors (crowded condition). Targets appeared at one of three locations: left, lower-left, or below the fixation point, at eccentricities of 1.5 or 7 deg. The letter sizes were 0.44 deg for 1.5-deg eccentricity and 2.1 deg for 7-deg eccentricity, with inter-letter distances of 0.6 deg and 3.0 deg, respectively. Participants identified the target using a two-alternative forced choice task. The crowding effect sizes were calculated as the difference in performance between isolated and crowded conditions for each stimulus type and viewing condition, then averaged across participants. Results showed a crowding reduction even for letter strings oriented in unfamiliar directions under specific conditions, such as 1.5-deg eccentricity with right-diagonal string orientation and targets positioned left or below fixation. However, this reduction was inconsistently observed across conditions and less robust than for familiar (horizontal or vertical) orientations. These findings suggest that crowding reduction depends on reading experience with specific string orientations, highlighting the limited role of familiarity in reading.

A part of this work was supported by JSPS KAKENHI Grant Number 20K11927.

26.319 THE ROLE OF EXPERIENCE WITH VISUAL AND LINGUISTIC FACTORS ON SKIPPING BEHAVIOR DURING READING IN TEEN AND ADULT READERS

Anastasia Stoops[1,2], Sarn Mukhopadhyay[1,2], Volodimir Kindratenko[1,2], Jon Willits[1,2], Jessica Montag[1,2]; [1]University of Illinois at Urbana-Champaign, [2]National Center for Supercomputing Applications (NCSA), Urbana, IL, USA

During reading skilled readers skip and never fixate on between 20-30% of words. In controlled experiments, a combination of low-level visual factors such as word length and higher-level linguistic factors such as word frequency and lexical predictability affects skilled readers' skipping: shorter and lexically more predictable words are skipped more often. Recently, syntactic predictability (expectations about the upcoming word's part of speech and not the exact word) has been proposed as an additional source of linguistic predictability that affects skipping. To understand how skipping behavior emerges with experience, we investigated skipping behavior in both adults and adolescents, who are skilled readers but not yet adult-like in their behavior. The present study examined how visual (word length) and linguistic (word frequency, lexical and syntactic predictability) knowledge during reading affects skipping in skilled and adolescent readers. 113 college students and 52 adolescents (14-17 yoa) read 55 passages from PROVO corpus (Luke & Christianson, 2018) along with vocabulary (Shipley, 1941) and reading comprehension (shortened 10 minute version of Nelson&Denny, 1980) tests, combined into one reading experience composite score. Logistic mixed-effects regression examined the effects of reading experience, word length, word frequency, lexical and syntactic predictability controlling for the position in the sentence. Our results replicated prior findings from experimental and naturalistic work: readers skip predictable short words more often than predictable long words. Further, two novel findings emerged. First, word length interacted with reading experience such that better readers skip more longer words than poorer readers. Second, word length interacted with syntactic predictability such that longer, syntactically predictable words are skipped more than less predictable words. This work highlights the role of readers' experience with linguistic knowledge and word's syntactic predictability in eye movements. We show tighter than previously assumed coupling between readers' eyes movements and higher-level linguistic predictability beyond specific lexical items.

funding: NSF grant #2141268 to AS, JM, and JW

Object Recognition: Frames of reference

Saturday, May 17, 2:45 – 6:45 pm, Banyan Breezeway

26.320 MENTAL ROTATION: SEX AND STEM DIFFERENCES

Raevan Hanan[1], Brandon Eich[2], Tom Tomshe[3], Chloe Kindell[4], Hanane Ramzaoui[5], Heather Lucas[6], Melissa Beck[7]; [1]Louisiana State University

Mental rotation is a valuable skill for individuals in many STEM-related fields. Further, females are underrepresented in many STEM disciplines. Sex differences in mental rotation abilities may be a contributing factor, as evidence suggests that males outperform females on mental rotation tasks. In this study, we directly compared sex differences in mental rotation to differences in performance between STEM and non-STEM majors. Undergraduate participants (N=91) judged if two block stimuli presented side-by-side were either the "same" (non-mirrored objects) or "different" (mirrored objects). The angle of disparity (AoD), between the two objects ranged from 0-160°. For sex differences, we observed a main effect of sex for accuracy, such that male participants had higher accuracy than female participants. This main effect of sex was not observed in response times (RT), suggesting that when female participants accurately rotate the objects, they are doing so as quickly as their male counterparts. However, there was a three-way interaction between AoD, trial type (same/different), and sex for RT. The difference between same and different trials is smaller overall for the harder trials (AoD 120°, 140°, & 160°) and this decrease in the effect at larger AoDs is more pronounced for females. For STEM versus Non-STEM majors, there was no main effect of major for accuracy or RT. However, there was an interaction between trial type and major. Specifically, accuracy on same trials was consistently higher than accuracy on different trials for non-stem majors, but this difference was less consistent for STEM majors. STEM majors may use a different strategy than non-STEM majors on some trials that is not biased toward better responses on same trials. These results suggest that differences in mental rotation associated with sex impact performance in a different way than differences associated with STEM experience.

This work was funded by the National Science Foundation under Grant No. AM230273 for an ongoing investigation titled Enhancing STEM Success: A Multi-modal Investigating of Spatial Reasoning and Training in Undergraduate Education. The views expressed do not necessarily reflect those of the NSF.

26.321 WHEN TRACKING JUST ONE PREDICTABLE AND SLOW OBJECT IS VERY HARD: A CONSTRAINT FROM STRUCTURED REPRESENTATION OF AN ORBIT

Qihan Wu[1] (qwu30@jhu.edu), Jonathan I. Flombaum[1]; [1]Johns Hopkins University

The display and task are simple to describe, but in one condition, the task is surprisingly difficult to perform. In the easy condition, the perimeter of a square is constructed from adjacent red discs, four per side. A target disc turns white momentarily. The participant must reidentify the target at the trial end, which follows a period during which the square rotates around its center point. Average performance was 92%. The second condition is identical, except the red discs translate at a constant linear speed, tracing the perimeter of a stationary square. Performance is significantly worse, 84% on average. Two control experiments replicate the findings, and demonstrate that translational motion is the specific cause of the tracking deficiency compared to rotation. We explain the results through appeal to internally referenced representations, like those that underlie the perception of elongated objects. Specifically, the presence of a visible square causes the individual discs to be localized with (internal) reference to the square as a whole—something like "it's the second one from the left on the north side." Translating motion conflicts with this representation because the "north" side remains stationary though the target does not remain there. But with rotation, the internally referenced description remains accurate throughout the motion period, and tracking can progress by updating the square's changing alignment to an external reference. In a fourth experiment, connections between discs alleviate the challenge of translation, by exchanging the square for representations of barbell objects. We also report experiments with opposite effects: a cost for rotation compared to translation in situations where the representational conflict runs in the opposite direction. Together, the experiments reveal how composite and internally referenced representations support and constrain visual processing broadly, including the perception of objects, orbits, and groups.

26.322 WORLD-CENTRIC INDUCED CARDINAL BIASES IN 3D OBJECT ORIENTATION PERCEPTION

Anantaporn Sena[1] (anantaporn.s@chula.md), Rosanne Rademaker[2], Chaipat Chunharas[3]; [1]Cognitive Clinical & Computational Neuroscience Lab, Chula Neuroscience Center, King Chulalongkorn Memorial Hospital, [2]Rademaker Lab, Ernst Strüngmann Institute for Neuroscience, Max Planck Society, Frankfurt, Germany, [3]Chulalongkorn University Chula Neuroscience Center, King Chulalongkorn Hospital

Orientation perception of simple stimuli, such as lines or gratings, is affected by cardinal biases. These biases cause exaggeration of orientations away from the perfect vertical or horizontal axes. Previous research suggested that these biases may be caused by our world's natural statistics which shape our perceptual system to be more sensitive to cardinal orientations. Our previous study had revealed similar biases when perceiving a 3-dimensional object (an upright sitting cat) rotating in certain axes. This finding supports the idea that people use cardinal axes as a reference to enhance perceptual sensitivity. The question that remains unclear was whether these biases were driven by the object-centric axis or the world-centric axis. To explore this, our current study investigated the effect of object orientation using a 3-dimensional rotation task with a different major-axis stimulus (a bus). Six participants were briefly presented with a stimulus in pseudo-randomized orientations along three axes: —roll, pitch, and yaw. They were then asked to reorient the probe stimulus to match the rotation degree of the sample. The results revealed error patterns similar to those observed in our previous study. Specifically, a repulsive effect at vertical and horizontal orientations was found on the roll axis. On the yaw axis, repulsive effects were observed on the front and back sides of the object, while attractive effects were seen on the sides. On the pitch axis, however, only subtle biases were found. These findings, consistent with our earlier research, suggest that the error response patterns in major axes are influenced by the world-centric axis rather than the object-centric axis in the perception of an object's orientation.

26.323 THE IMPACT OF VISUAL OBJECT PERCEPTION ON THE DESIGN OF WRITING SYSTEMS ACROSS LANGUAGES

Roxana Ismail-Beigi[1], Yaoda Xu[1]; [1]Yale University

A diverse array of writing systems has been developed across human culture and history. After prolonged training, humans are capable of acquiring reading proficiency in each of these writing systems. Here we asked to what extent distinguishing features in visual object perception are incorporated into the design of writing systems. Bao et al. (2020, Nature) showed that visual objects are differentiated primarily by their animacy/curvature and spikiness in the primate infero-temporal (IT) cortex. The same two shape-defining features are also present in convolution neural networks (CNNs) trained for object classification. Are writing systems designed to take advantage of these shape-defining features and maximize letter/character dissimilarities

in curvature and spikiness to aid efficient letter learning and reading? Or are writing systems developed serendipitously without such considerations? To test this, we generated letter/character images from writing systems including alphabets (Latin, Greek, Cyrillic, Elder Futhark, Mongolian), abjads (Arabic, Hebrew, Phoenician), abugidas (Hindi Devanagari, Thai), and syllabaries (Japanese kana, Cherokee, Persian cuneiform), as well as a constructed abugida (Tengwar). We ran these images through the ImageNet-trained Alexnet layer FC6 and plotted the resulting activations in an object space defined by curvature and spikness. This object space was constructed following Bao et al. and closely resembled the object space obtained from the macaque IT cortex neuronal responses. We then obtained the average pairwise distances of the letters/characters for a given writing system in this space. Among the writing systems examined, Latin letters exhibit the greatest dissimilarity in terms of curvature and spikiness, with their overall distribution roughly expanding the object space. Other writing systems, such as kana, Devanagari, and Thai, tend to cluster densely in one area of the object space. Thus, while some writing systems incorporate the defining features of object perception, others appear to do so to a lesser extent.

NIH grant R01EY030854

Perceptual Organization: Serial dependence

Saturday, May 17, 2:45 – 6:45 pm, Banyan Breezeway

26.324 THE INFLUENCE OF TEMPORAL CONTEXT ON VISION OVER MULTIPLE TIME SCALES

Reuben Rideaux[1,2], Kacie Lee[1]; [1]University of Sydney, [2]University of Queensland

Past sensory experiences influence perception of the present. Multiple research subfields have emerged to study this phenomenon at different temporal scales. Effects associated with temporal context fall into three categories: the influence of immediately preceding sensory events (micro), expectations established by short sequences of events (meso), and regularities over long sequences of events (macro). In a single paradigm, we examined the influence of temporal context on perception at each scale. By integrating behavioural, electroencephalographical, and pupillometry recordings, we identify two distinct mechanisms that operate across all scales. The first is moderated by attention and supports rapid motor responses to expected events. The second operates independently of task-demands and dampens the feedforward neural responses produced by expected events, leading to unexpected events eliciting earlier and more precise neural representations. We further show that perceptual recall of sensory events exclusively reflects neural representations during this initial feedforward stage and that recall biases towards previous events (serial dependence) can be explained by expectation of sensory stability over time.

This work was supported by Australian Research Council (ARC) Discovery Early Career Researcher Awards awarded to RR (DE210100790) a National Health and Medical Research Council (NHMRC; Australia) Investigator Grant (2026318).

26.325 SERIAL DEPENDENCE UNDER UNCERTAINTY: THE ROLES OF SIGNAL-TO-NOISE RATIO AND POSITIVE EVIDENCE

Zoe Little[1], Colin Clifford[2]; [1]University of New South Wales

Assimilative serial dependence in orientation perception occurs where the response to an orientated stimulus is biased towards the orientation of stimuli seen in the recent past. It remains a matter of debate whether serial dependence is mediated at the sensory or decisional level. Manipulations that induce uncertainty in decision-making without affecting the quality of sensory information may help to distinguish between these explanations. Here, we made use of the positive evidence bias, wherein increasing the positive evidence of a stimulus (i.e., overall contrast) while keeping the signal-to-noise ratio (SNR) the same increases the confidence but not accuracy of responses. Previous research has demonstrated that serial dependence in orientation perception is greater to high positive evidence stimuli, suggesting that perception may be skewed towards more certain previous decisions (Samaha et al. [2019] J. Vis.). We asked how decisional uncertainty (positive evidence) and sensory uncertainty (SNR) in both current and prior stimuli modulate the size of the serial dependence effect. Participants (N = 33) viewed oriented Gabor patches presented under different levels of SNR and positive evidence and reported their orientation, as well as their response confidence. We found that serial dependence was greatest when the stimulus on the target trial had low SNR, low positive evidence, or was responded to with low subjective confidence, affirming that uncertain stimuli are more likely to be integrated with previously seen information. We found no modulation of serial dependence by any source of uncertainty on the inducer trial. This has implications for theoretical accounts of serial dependence as well as for identifying whether serial dependence is driven by previous visual inputs or previous decisions.

26.326 NAVIGATING UNCERTAINTY: THE ROLE OF PERCEPTUAL CERTAINTY IN SERIAL DEPENDENCE

Muzhen Ai[1]; [1]UCLA

Serial dependence effect, defined as the influence of past sensory input on current perception, is a phenomenon where prior stimuli bias current perceptual judgments. This study seeks to examine the factors that modulate serial dependence magnitude by aggregating and analyzing results across 8 studies (N=211) on serial dependence. We categorized trial-wise perceptual certainty defined by noise level, spatial frequency, stimulus contrast. Through this approach, we compared serial dependence effects in high- and low-certainty conditions. Our results demonstrate that, within current trial settings, low-certainty conditions consistently exhibit a higher reliance on past sensory information than high-certainty conditions, aligning with the Bayesian account. However, previous low perceptual certainty elicits a stronger serial dependence effect. We propose that in the face of ambiguous prior sensory information, individuals may adopt a perceptual strategy that places less emphasis on current sensory input, reflecting an adaptive mechanism for managing uncertainty in perceptual environments.

26.327 REPULSION FROM AND ATTRACTION TOWARDS PREVIOUS STIMULI AND RESPONSES DEPEND ON RESPONSE TYPE AND STIMULUS DURATION

Mert Can[1] (mert.can@u-paris.fr), Thérèse Collins[1]; [1]Integrative Neuroscience and Cognition Center, Université Paris Cité & CNRS, France

Perceptual reports can be attracted towards or repulsed from previously seen stimuli. We investigated the conditions in which attraction and repulsion occur with oriented Gabors by manipulating response type (continuous or categorical), response frequency (to every stimulus or every other stimulus), and stimulus duration (50 or 500 ms). When participants gave a continuous response by adjusting a response cue to match the orientation of the stimulus, repulsion from the previous stimulus occurred when the stimulus was presented for 50 ms and attraction to the previous stimulus when the stimulus was presented for 500 ms. These effects occurred whether participants responded to every stimulus or every other stimulus. When participants gave a categorical response by indicating whether the orientation was clockwise or counter-clockwise from a reference, there was attraction to the previous response and repulsion from the previous stimulus. Attraction to the previous response was weaker when participants responded to every other stimulus, suggesting a decay of response representation over time, and stronger when the stimulus was ambiguous (i.e., oriented close to the reference stimulus). Our results provide evidence of temporal differences between repulsion and attraction, which may relate to the involvement of feedforward and feedback processes in facilitating these effects. They also show that the processes resulting in repulsion or attraction were defined by response type. We suggest that attraction is both a decisional effect associated with categorical responses and a perceptual effect associated with continuous responses.

26.328 REPULSIVE SERIAL DEPENDENCE FROM A BINOCULAR MECHANISM

Thérèse Collins[1], Steven Shevell; [1]Université Paris Cité & CNRS, Paris, France, [2]University of Chicago, Departments of Psychology and Ophthalmology & Visual Science, Chicago, IL, USA

Stimulus history can shape perceptual experience by attracting or repulsing subsequent stimuli. The well-known phenomenon of serial dependence (SD) quantifies the degree to which an immediately previous stimulus alters the perception of a current one. We used monocular versus binocular stimulus presentation to investigate the neural level at which SD occurs. Participants indicated whether the direction of motion of a random dot kinematogram (RDK) presented for 500 ms was CW or CCW relative to a reference direction. In Experiment 1, RDKs were binocular and coherence varied (25%, 75%, 100%). Reported percepts were repulsed away from the previous direction (e.g., a previous counterclockwise stimulus increased clockwise reports for the current stimulus). In Experiment 2, RDKs were presented to observers either monocularly or binocularly at 25% or 100% coherence. With a current monocular stimulus, serial dependence was greater when the previous stimulus was binocular rather than monocular. This suggests a mechanism mediating SD is driven by binocular combination. Experiment 3 had three presentation conditions: 100% coherence presented to both eyes, 50% coherence presented to both eyes, or 0% coherence presented to the left eye + 100% coherence to the right eye. If SD depends on the binocularly fused coherence then the 50% binocular and 0/100% conditions should result in the same magnitude of SD. Repulsion was strongest from the 100% binocular RDK and weakest from the 50% binocular RDK, as expected; the mixed 0/100% RDK caused intermediate repulsion. Results from these three experiments show that repulsive serial dependence increases with the coherence of the previous stimulus, and is driven by a response at a binocular level of neural integration though not simply by the coherence of the fused binocular percept.

International Institute for Research in Paris, University of Chicago

26.329 SET SIZE AND DELAY DURATION INTERACT IN INFLUENCING SERIAL DEPENDENCE IN VISUAL WORKING MEMORY

Alan L. F. Lee[1] (alanlee@ln.edu.hk), Jenny W. S. Chiu[1]; [1]Lingnan University, Hong Kong

Serial dependence can be found in observers' responses in a series of visual working memory (VWM) tasks (e.g., Fritsche et al., 2017; Bliss et al. 2017; Pascucci et al., 2019). To assess how memory load influences serial dependence in VWM, we manipulated memory load by separately varying set sizes and delays in the present study. On each trial, observers viewed a set of 1, 3, or 6 Gabor patches (between-subjects, n=20 for each set size). Then, after a random delay (sampled from a normal distribution with M = 2,000ms, SD = 500ms), one of the Gabors' locations was cued as the target for response. Observers first reported the target Gabor orientation by matching it with an on-screen, rotatable line, then indicated their confidence rating on their orientation-judgment task on a slider bar. The difference in target Gabor orientation between every two consecutive trials was randomly chosen from 11 values, ranging from -75 to +75 degrees (15 degrees apart). We quantified serial dependence through bootstrapping: For each of the 1,000 bootstrap samples (20 observers resampled with replacement), we estimated the amplitude parameter by fitting a derivative-of-Gaussian (DoG) function to the grouped-average response errors, separately on the short- and long-delay sets of trials (median split) for each set size. Overall, consistent with previous findings, we found a significant attractive serial-dependence effect. Interestingly, attraction was the strongest for the set size of 3 (amplitude of around 2.5 deg), followed by 6 (~1.25 deg), and then by 1 (~0 deg), suggesting that serial dependence may be facilitated by an intermediate load in a memory task. Furthermore, set size and delay interacted in influencing serial dependence: for the set size of 3, attraction was stronger in short delays than in long delays, but, for the set size of 1, the effect direction reversed.

This work was supported by the Research Matching Grant Scheme (LU Fund Code: 185218) from the University Grants Committee of the Hong Kong SAR

26.330 AUTISTIC AND SCHIZOTYPAL TRAITS PREDICT WEIGHTING OF SENSORY EVIDENCE AND PERCEPTUAL PRIORS IN VISUAL CAUSALITY JUDGMENTS

Gianluca Marsicano[1,2] (gm3598@nyu.edu), David Melcher[1,2]; [1]Psychology Program, Division of Science, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, [2]Center for Brain and

Health, NYUAD Research Institute, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

Causality judgments, which involve perceiving one event as causing another, are fundamental to understanding our sensory environment. Here, we investigated how perceptual history and individual differences in autistic and schizotypal traits shape visual causality judgments in the neurotypical population. Participants (n=90) reported causality judgments (causal/non-causal) for varying collision lags. We also measured autistic and schizotypal traits via questionnaires, and a data-driven cluster analysis divided individuals into three groups: high autistic traits (ASD-like), high schizotypal traits (SSD-like), and low traits (LT). Causal response rates were modelled as a function of collision lags, and serial dependence analysis examined how prior trial judgments influenced current decisions. Overall, perceptual history significantly shaped causality judgments, with trials following causal responses more likely to be judged causal, and vice versa. Consistent with prior research suggesting that ASD and SSD can represent two poles of a predictive continuum, the SSD-like group exhibited a stronger tendency toward causality judgments and greater serial dependence, while the ASD-like group showed reduced serial dependence compared to the LT group. We characterized these individual differences in terms of encoding and decoding mechanisms using a hierarchical drift diffusion model. Results revealed that the ASD-like group required longer sensory encoding times and adopted a more cautious response strategy (greater decision thresholds) while displaying faster evidence accumulation (higher drift rate). In contrast, for the SSD-like group the encoding processes did not differ from LT, but they exhibited an initial bias toward causal responses and a lower drift rate, reflecting greater reliance on prior models instead of immediate sensory evidence. Together, these findings highlight distinct atypicalities for ASD- and SSD-like profiles in the encoding and decoding mechanisms underlying visual causality judgments, providing insights into how sensory encoding, evidence accumulation, and perceptual history interact to shape causality perception.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYU Abu Dhabi Research Institute grant CG012.

Perceptual Organization: Neural mechanisms

Saturday, May 17, 2:45 – 6:45 pm, Banyan Breezeway

26.331 PROBING THE NEURAL BASIS OF VISUAL ABSTRACTION: MACAQUES AND ANN MODELS ACHIEVE SIMILAR SKETCH RECOGNITION PERFORMANCE

Umael Qudrat[1], Shirin Taghian Alamooti[1], Judith Fan[2], Kohitij Kar[1]; [1]York University, [2]Stanford University

Visual abstraction is the process of distilling complex visual scenes into their essential components. Such abstraction is exemplified by the production and recognition of sketches, which can be effective in conveying the content of a scene while omitting many details. To what degree does visual abstraction also manifest in non-human primates, and what neural computations are responsible? To answer this question, we measured how well macaque monkeys (N=2) could identify the visual concept conveyed in human-drawn sketches and evaluated how well their behavior and ventral stream neural responses could be predicted by an artificial neural network (i.e., AlexNet). Both macaques performed a sketch-recognition task using stimuli (1000 images, 10 object categories) from the Google Quick Draw dataset. In each trial, we briefly presented (100 ms) a sketch, followed by a choice screen where monkeys selected which of two object images the sketch represented. Both monkeys achieved accuracies exceeding 70%, demonstrating that even simple sketches convey sufficient information for robust object identification by non-human primates. The monkeys' image-by-image recognition accuracies significantly correlated with those predicted by AlexNet ($R = 0.42$, $p<0.001$). This correlation matched the monkeys' noise ceiling (~0.4), indicating that AlexNet strongly approximates the abstraction strategies employed by the primate visual systems, given intrinsic variability in the current dataset. We also recorded population activity (384 sites) from inferior temporal (IT) cortex as monkeys viewed line drawings and sketches. Using linear classification on IT responses, we found that distributed neural activity patterns strongly predicted (accuracy~0.83, chance-level=0.5) object identity, providing evidence that IT cortex encodes the abstract visual features underlying sketch recognition. These findings establish macaques as a powerful model for investigating the neural computations that support sketch recognition.

KK was supported by funds from the Canada Research Chair Program, the Canada First Research Excellence Funds (VISTA Program), Google Research Award, and a NSERC-DG. JEF is supported by NSF CAREER #2047191, NSF DRL #2400471, and a Stanford Human-Centered AI Institute Hoffman-Yee award.

26.332 BORDER OWNERSHIP SELECTIVITY DIFFERS WITH CELL CLASS IN VISUAL AREA V4

Maryam Azadi[1], Fatemeh Didehvar[1], Tom P. Franken[1]; [1]Washington University in St.Louis School of Medicine

An important step in the segmentation of visual scenes into discrete objects is assigning each border to the correct side of foreground (border ownership). Object borders are owned by one side at a time, suggesting a role for inhibitory neurons in the computation of border ownership. While border ownership signals are known to be prominent in primate visual cortex, the relation between border ownership selectivity and cell type is unknown. Here we investigated this using Neuropixels probes in macaque area V4. We inserted the probes orthogonally in the cortex to record from populations of well-isolated neurons while a macaque was viewing visual scenes. We first mapped the aggregate receptive field of the columnar population of simultaneously recorded cells. We then presented a luminance-defined square object such that only one of its borders fell in this receptive field. We varied the side-of-ownership, luminance contrast polarity and square orientation. We identified border ownership-selective (BOS) cells by assessing the statistical significance of the border ownership index (BOI, difference divided by sum of spike rates to opposite sides of ownership) at the border orientation corresponding to the preferred square position. We used spike duration to separate the population into narrow-spiking (putative inhibitory neurons) and broad-spiking (putative excitatory pyramidal) cells. As in prior studies, spike duration was strongly bimodal, and narrow-spiking cells had higher mean firing rates and higher peak-trough ratios than broad-spiking cells. About a third of both narrow-spiking and broad-spiking cells were tuned for border ownership. Broad-spiking BOS cells had significantly higher BOI magnitudes than narrow-spiking BOS cells. Our data suggest that putative pyramidal cells have stronger border ownership selectivity than putative inhibitory cells in V4. Because cells that project to other areas are pyramidal cells, this suggests that border ownership signals are prominent in the projections that leave V4.

This work was supported by NIH grant R00EY031795 and a 2024 NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation.

26.333 SOMATOSENSORY RESPONSE TO VISUAL OBJECTS: CROSS-MODAL PROCESSING OF GRASP-RELATED FEATURES REVEALED BY MEG

Margherita Marchioro[1], Davide Tabarelli[1], Tommaso Currò[1], Fraser Smith[2], Carlo Miniussi[1], Luigi Cattaneo[1], Alessandra Dodich[1], Stefania Bracci[1], Luca Turella[1], Simona Monaco[1]; [1]Center for Mind/Brain Sciences (CIMeC), University of Trento, [2]University of East Anglia

Previous findings have shown that visual presentation of familiar but not unfamiliar objects elicits content-specific activity patterns in the early somatosensory cortex and triggers significant decoding in the mu rhythm response, known to be associated with tactile processing. However, it is not known what aspects of familiarity elicit these effects. Here we explored whether graspability, texture and grip size of viewed objects play a role in cross-modal activation in the somatosensory cortex. We used magnetoencephalography (MEG) to test 18 right-handed participants who viewed visual familiar stimuli presented in a block-design paradigm. The stimuli included graspable and ungraspable items, with graspable items classified by texture (smooth/rough) and grip size (small/large). We assessed graspability, texture, and size contrasts through univariate sensor-level analyses, computing event-related fields (ERFs) and time-frequency representations (TFRs). We hypothesized that early somatosensory effects in central and parietal regions, if reflecting prior experience, would show mu rhythm modulation across all contrasts and beta power variation for graspable items. We found significant ERFs in parietal, temporal, and occipital sensors for graspability (44-148ms) as well as for texture (164-680ms). Significant clusters emerged for size in parietal and occipital regions (34-162ms). The TFRs analysis on graspability showed beta power modulation in parietal and temporal regions, and mu-α power modulation in frontal and occipital sensors. Furthermore, we found mu-α power modulation for size in temporal and frontal sensors, and for texture in temporal and parietal sensors. Our results indicate that graspability of viewed objects activates regions beyond those typically associated with vision, particularly those involved in somatosensory processing, as suggested by mu rhythm modulation for graspable items. This cross-modal activation is likely related to previous experience with object manipulation and concurrent multimodal stimulation, i.e., vision and touch, and might allow retrieving the perceptual representation of prior experience to guide our behaviour.

This work was funded by the European Union - Next Generation EU - Mission 4, Component 2, CUP E73C22000490001, ID MSCA 0000051, and by the B2B project SPOKE 4. MNESYS grant PNRR - Next Generation EU.

26.334 SUBSTANCE-INVARIANT PROCESSING FOR JUDGMENTS OF FRONTAL PLANE DISTANCE, CENTROIDS, NUMEROSITY, AND LARGE-LETTER IDENTITY.

George Sperling[1] (sperling@uci.edu), Lingyu Gan[1]; [1]University of California, Irvine

Household measuring cups, weight scales, and measuring tapes, do not know what they are measuring, they simply make a real-number, substance-invariant measurement. Because there is a virtual infinity of visual substances, the brain also must make substance-invariant measurements but it uses a different mechanism: salience processing. The brain makes substance-invariant judgments by ignoring the nature of the substance to be measured, whether it is particular set of shapes, colors, textures, or whatever, and representing the just locations being occupied, and then measuring the locations versus the substance itself. The set of possible locations is called a salience map, occupied locations are designated by a positive number, salience. Originally, this salience architecture was used to describe priority in visual processing and visual search (Koch and Ullman, Human neurobiology, 1985). However, the same architecure can be used for other purposes (e.g., direction of feature motion, Lu and Sperling, Nature, 1995). Because what is judged is not the substance itself, but merely the locations occupied by the substance, the judgments are indifferent to the nature of the substance. Here we demonstrate three judgments that are substance invariant: Distance between two items in frontal plane, the centroid or numerosity of a group of items. Judgment accuracy is independent of whether the items are similar or different or whether they are isoluminant with the background, provided the items are distinctly visible. These three parametrically studied judgments are of novel stimuli. We also demonstrate partial substance invariance for text: Many different colors of large letters on an isoluminant gray background are easily readable but very small letters require high contrast black or white. Conclusion: The brain solves the problem of measuring many different substances by representing just the locations of the substances in a salience map and measuring the salient locations.

26.335 QUANTIFYING PHOSPHENE SIZE USING MRI-GUIDED TRANSCRANIAL MAGNETIC STIMULATION TO PRIMARY VISUAL CORTEX

Remy Cohan[1,2,3] (rcohan@yorku.ca), Stefania S. Moro[1,2,3], Jennifer K. E. Steeves[1,2,3]; [1]Centre for Vision Research, [2]Centre for Integrative and Applied Neuroscience, [3]Department of Psychology, York University, Toronto, Canada

Visual phosphenes are perceived flashes of light in the absence of retinal input and can be evoked by transcranial magnetic stimulation (TMS) to primary visual cortex (V1). Previous dose-response studies using direct electrical stimulation (DES) of the visual cortex have shown that higher intensities produce larger phosphenes, suggesting increased stimulation intensity affects a larger cortical area. Unlike DES, in TMS intervening tissues such as scalp, skull, and

cerebrospinal fluid can attenuate induced electric fields. Prior studies have examined stimulation intensity, but few have investigated how individual differences in phosphene size relate to biophysical factors. The current study examines TMS-induced phosphene size and its relation to phosphene thresholds, scalp-to-cortex distance, and modelled electric field strength measured using MRI-guided stereotaxic neuronavigation and our computer-based phosphene reporting tool. V1 stimulation evoked phosphenes in quadrants of the visual field corresponding to retinotopic region. Perceived phosphene size was negatively correlated with phosphene thresholds, indicating that higher phosphene thresholds led to smaller perceived phosphenes. Age and scalp-to-cortex distance were not correlated with phosphene size, consistent with the notion that scalp-to-cortex distance does not account for different intervening tissue types and corresponding properties. Phosphene size, however, was negatively correlated with electric field strength, indicating that intervening tissue properties may attenuate TMS intensity. These findings suggest that variability in phosphene perception may reflect biophysical factors, such as intervening tissue properties and highlight the importance of accounting for biophysical factors and electric field modelling to better understand variability in TMS-evoked phosphene perception. In addition, using a standardised method such as a phosphene mapping tool to quantify individual differences in TMS response is valuable to facilitate standardization of methods for non-invasive brain stimulation to advance TMS research and optimize its application in 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), and Connected Minds.

26.336 PERCEPTUAL FILLING-IN OF MOTION SPEED AT THE PHYSIOLOGICAL BLIND SPOT

Einstina Wang[1] (einstinawang@berkeley.edu), Dongcheng He[2], Susana T. L Chung[1]; [1]University of California, Berkeley, [2]Hefei Comprehensive National Science Center

The physiological blind spot, a photoreceptor-free retinal region, does not result in a lack of perception, due to the brain's filling-in process. Prior research highlights mostly V1-related features in filling-in. This study aimed to examine whether motion speed information can also "fill-in". After mapping out the blind spot of their right eye, seven participants with normal vision (age: 19-23 years) performed a speed judgment task monocularly, while fixating at a fixation target with eye movements monitored. Testing alternated between the two eyes. The stimulus comprised a vertical bar (width=1°) right of fixation, centered at the center of the blind spot. Two vertical patches (1°×6°) containing a horizontally-oriented grating were presented above and below the bar such that when viewed with the right eye, the two patches abutted the top and bottom edges of the blind spot, respectively. On any given trial (duration=1s), the spatial frequency (0.5, 1 or 1.5 cpd) and the motion direction (upward or downward) of the two grating-patches were identical; with one patch (randomly chosen) moving at 0.67 deg/s and the other moving at one of four speeds: 0.78, 0.97, 1.15 or 1.34 deg/s. Participants indicated which patch (top/bottom) moved faster. The point-of-subjective-equality (PSE) was determined from the psychometric function relating the proportion of upward responses with the speed difference between the two grating-patches. Our null hypothesis was that if motion speed did not induce filling-in within the blind spot, then the PSE for the right and left eyes should be the same. Across participants, we found a small effect of spatial frequency ($p=0.044$). Specifically, the PSE for the right and left eyes differed at 1 cpd, but not at the other two spatial frequencies. Our findings imply the possibility of a motion speed-induced filling-in of information within the blind spot, but the effect is apparently small.

UC Berkeley Rose Hills Summer Scholarship

26.337 TRAINING-INDUCED RECOVERY OF MOTION PERCEPTION AFTER OCCIPITAL STROKE DESPITE V1-V4 DAMAGE

Niki Lam[1] (niki_lam@urmc.rochester.edu), Berkeley Fahrenthold[1], Matthew Cavanaugh[1], Bryan Redmond[1], Jingyi Yang[1], Duje Tadin[1], Marisa Carrasco[2], Noah Benson[3], Elisha Merriam[4], Krystel Huxlin[1]; [1]University of Rochester, [2]New York University, [3]University of Washington, [4]National Institute of Mental Health

Cortical blindness is often defined by damage, usually from stroke, to the primary visual cortex (V1) in the occipital lobe of the human brain. Here, we asked to what extent such strokes also affect extrastriate visual areas. We then assessed the relationship between damage, and the size and severity of behavioral defects, and the efficacy of visual motion discrimination recovery attainable with perceptual training. We analyzed data from 35 subacute occipital stroke patients (mean age/SD: 52.3/11.5 years; time post-stroke: 3.3/1.1 months; male/female: 23/12) enrolled in clinical trial NCT04798924 (ClinicalTrials.gov). Patients underwent Humphrey perimetry and eye-tracker-enforced psychophysical testing to measure fine direction discrimination (FDD) thresholds in their blind and intact hemifields. This was done before and after training on the FDD task with endogenous, feature-based attention (FBA) pre-cues for 6-11 months. Structural MRI scans of each brain were analyzed using the Benson atlas (Benson & Winawer, 2018) to estimate whether V1-V3, V3A/B and V4 were damaged. Damage exclusive to V1 was found in only 11% of patients; the rest exhibited damage to multiple, early visual areas. Patients with V1-only damage had smaller visual defects than those with combined V1 and extrastriate cortex damage [$p = 0.01$], although this was potentially biased by the small sample for V1-only damage. Importantly however, severity of the deficit and efficacy of FDD training did not differ between those with V1-only damage versus combined damage. Instead, training-induced improvements in FDD thresholds were directly proportional to the number of days trained. In conclusion, while occipital strokes rarely injured only V1, the number and type of extrastriate areas damaged did not correlate with pre-training preservation of blind-field motion discrimination abilities pre-training, or with the potential to recover motion discrimination with training. Our findings suggest that both preservation and recovery likely rely on visual pathways that bypass V1-V4.

26.338 UNDERSTANDING THE REPRESENTATIONAL GEOMETRY OF PSYCHOLOGICAL AND NEURAL SPACES ACROSS MULTIPLE SIMILARITY DIMENSIONS.

Johan A. Gamba[1], Vincent Taschereau-Dumouchel[2], Megan A.K. Peters[3], Brian Odegaard[1]; [1]University of Florida, [2]University of Montreal, [3]University of California Irvine

Assessing similarities between objects is a fundamental cognitive process for humans. Multivariate neuroimaging techniques like Representational Similarity Analysis (RSA) have advanced our understanding of neural activation patterns underlying conceptual distances. In our study, we explored the relationship between eight psychological similarity dimensions and fMRI voxel-based RSA matrices across 11 regions of the ventral pathway and prefrontal cortex in Spanish- and English-speaking populations. Participants rated similarities of object pairs (280 in total) based on eight properties (general similarity, animacy, shape, color, category, dissimilarity, preference, and fear). We compared these behavioral ratings to neural data, employing metrics such as correlational distance (c.d.), c.d. post-Fisher transformation, and cosine similarity, alongside Pearson, Spearman, and Kendall Tau correlations, to identify amplified or generalized features across metrics. Our results revealed a behavioral representational distinction between “object-based” dimensions (e.g., shape) and “subject-based” ones (e.g., preference). Object-based properties showed a consistent representational geometry, whereas subject-based properties exhibited differences that modulated their alignment with neural data. Behavioral categories like “general similarity,” “color,” and “shape” were highly correlated with neural representations, particularly in occipital and fusiform regions. However, results were influenced by noise ceiling limits of each region of interest. Consequently, our results suggest a higher granularity in sensory cortices compared to multimodal areas and also inform methodological considerations when contrasting results across regions. Overall, our findings highlight consistent cross-population relationships between psychological and neural representational spaces, providing insights into how object-based and subject-based dimensions are encoded across different brain regions.

SATURDAY AFTERNOON POSTERS IN PAVILION

Attention: Neural, spatial

Saturday, May 17, 2:45 – 6:45 pm, Pavilion

26.401 ATTENTION MODULATES NEURAL REPRESENTATIONS FOR VISUAL FEATURES WITH MULTIPLE MANIFESTATIONS ALONG THE VISUAL HIERARCHY

Xiaoli Zhang[1], Taosheng Liu[1]; [1]Michigan State University

Much research has shown that spatial attention increases neural activity for attended stimuli. However, whether and how spatial attention modulates neural representations for visual features remains understudied. In this study, we comprehensively characterized attentional modulation using multiple measures of neural representations along the visual hierarchy. Inside an MRI scanner, a dot field in one of four colors was shown on each trial. Participants performed a luminance change discrimination task under two attentional conditions in separate runs: (1) at the dot field (i.e., the Attended condition), (2) at the fixation (i.e., the Unattended condition). We analyzed the blood oxygenation level-dependent (BOLD) activity in brain regions involved in visuospatial attention, including visual areas and frontoparietal regions. To assess the neural representations of color features, we compared a variety of metrics between Attended and Unattended conditions: univariate activity level, representational dimensionality, multivariate pattern classification, and noise correlation (NC) within each region. Both univariate activity and representational dimensionality confirm that attention increases neural signal strength to a greater magnitude in frontoparietal than visual regions. Attention also improves color decoding in prefrontal regions, although not in early visual regions. Interestingly, attention decreases within-region noise correlation in both early visual regions as well as frontoparietal regions. Our results are generally consistent with the model of frontoparietal network controlling spatial attention and reveal interesting connections from multiple measures of neural activity. Our results suggest that top-down spatial attention modulates neural activity along the visual hierarchy not only via increasing overall neural gain, but also by preserving neural population responses from sources of noise. Spatial attention appears to employ multiple mechanisms to enhance brain representations.

Funding: NIH R01EY032071

26.402 DOES HUMAN RIGHT FRONTAL EYE FIELD (RFEF+) PLAY A CRITICAL ROLE IN EXOGENOUS ATTENTION? A TRANSCRANIAL MAGNETIC STIMULATION (TMS) STUDY

Qingyuan Chen[1] (qc898@nyu.edu), Hsing-Hao Lee[1], Klara Hoxha[2], Antonio Fernández[1], Nina M. Hanning[1,3], Marisa Carrasco[1]; [1]New York University, [2]University of Bologna, [3]Humboldt-Universität zu Berlin

[BACKGROUND] Both endogenous and exogenous covert spatial attention improve visual sensitivity. Neurostimulation studies using transcranial magnetic stimulation (TMS) have shown that (i) early visual areas (V1/V2) are critical for exogenous but not endogenous attention, and (ii) the human homologue of the right frontal eye field (rFEF+) plays a central role for endogenous attention. It is unknown whether rFEF+ also plays a critical role for exogenous attention. Here, we used the same psychophysics-TMS protocol as in the previous studies to manipulate cortical activity in rFEF+ and measured its effects on the benefits and costs of exogenous attention. [METHODS] Observers performed an orientation discrimination task. Two oriented Gabor patches were preceded by an exogenous cue that was either valid, invalid, or neutral. We applied MRI-guided double-pulse TMS to the observers' rFEF+ during stimulus presentation. To measure contrast response functions (CRF), we varied stimulus contrast and analyzed sensitivity (d') in the target-stimulated and distractor-stimulated hemifields, contralateral and ipsilateral to the rFEF+ stimulation, respectively. [RESULTS] When TMS was applied to rFEF+, we found both benefits at the attended location and costs at the unattended locations at the high contrast levels, consistent with response gain, in both target- or distractor-stimulated hemifields. In contrast, when TMS was applied to V1/V2 in our previous study, the effects of exogenous attention remained at the distractor-stimulated location but were eliminated at the target-stimulated location. [CONCLUSION] This study reveals distinct roles of rFEF+ and V1/V2 in exogenous attention, emphasizing the critical involvement of rFEF+ in endogenous, but not exogenous attention. Together, our TMS findings establish a double dissociation between the two types of covert spatial attention and two critical cortical regions.

The Ministry of Education in Taiwan to HHL, Marie Skłodowska-Curie individual fellowship (MSCA-IF 898520) by the European Commission to NMH and NIH R01-EY019693 to MC

26.403 EMOTIONAL AROUSAL MODULATES THE GAIN OF CONTRAST RESPONSES IN HUMAN LGN

Chenanke Sun[1] (ankesun@bu.edu), Jasmine Pan[1], Louis Vinke[3], Joseph McGuire[1], Janneke Jehee[2], Sam Ling[1]; [1]Boston University, [2]Donders Institute for Brain, Cognition and Behavior, [3]Massachusetts General Hospital

Arousal is known to play a significant role in human behavioral performance and sensory processing. While its influence is believed extends to early vision, including as contrast sensitivity and spatial frequency processing, the neural mechanisms that give rise to these effects remain a poorly understood. To shed light on emotion's effects on vision, in this study we leveraged fMRI to assess how emotional arousal modulates the gain of the population contrast response function (CRF) in human visual processing. To do so, we measured BOLD CRFs within a participant's early visual cortex (V1-V3) and lateral geniculate nucleus (LGN), while they concurrently listened to auditory clips that were emotionally charged. To homogenize the population CRF, participants underwent a contrast adaptation paradigm, adapting to gratings of a fixed contrast level (16%) before viewing gratings with parametrically varied contrast levels (9 contrast levels, 2.67-96%). This paradigm was accompanied by sounds that were Negative or Neutral valence, taken from the International Affective Digitized Sounds (IADS) database. We found substantial increases in the response of the amygdala when participants heard the negative valence stimuli, supporting the efficacy of these stimuli. Interestingly, we discovered substantial changes in the CRFs within the LGN, but there were little-to-no modulatory effects of emotional arousal on the CRF in visual cortex. Specifically, we found an increase in the baseline response of the CRF in the LGN when participants listened to negative sounds compared to neutral sounds. Taken together, these findings suggest that emotional processing takes a predominantly subcortical route.

NIH RO1 EY035640; NIH RO1 EY028163

26.404 EXOGENOUS ATTENTION ALTERS THE GAIN OF THE CONTRAST RESPONSE IN HUMAN EARLY VISUAL CORTEX

Sanga Cho[1] (sangacho@bu.edu), Jasmine Pan[1], Sam Ling[1]; [1]Boston University

Our attention is involuntarily captured by sudden changes in our environment –a phenomenon known as exogenous attention. Previous studies have found that exogenous cueing of a location evokes a transient enhancement of visual processing for nearby stimuli. Although this has been well-established mostly in psychophysics and electrophysiological studies, we have relatively limited understanding of the neural mechanisms by which these effects arise. Indeed, while some neuroimaging work has found that exogenous attention boosts responses at attended locations, we lack a mechanistic account of the potential gain changes that underlie these boosted responses. Here, we examined the influence of exogenous attention on the BOLD population contrast response function (CRF) across early visual cortex (V1-V3) by pairing a contrast adaptation protocol (Vinke, Bloem & Ling, 2022) with a modified spatial cueing task. Specifically, participants began each scan with an initial adaptation period (60 sec) and each trial with top-up adaptation (4-8 sec) of two gratings (16% contrast), to promote population homogeneity in contrast response. After adaptation, participants were asked to identify the orientation of a target stimulus (200 ms), which was tilted clockwise or counterclockwise. Importantly, the target was preceded (pre-cue) or followed (post-cue) by a brief exogenous cue (50 ms). We found an increase in the gain of the contrast response with pre-cue (valid > invalid), particularly at high contrasts. This effect appeared most prominent in V1, decreasing from V2 to V3. The differential gain pattern of pre-cue was absent in post-cue trials, which rules out the possibility that the effect of the pre-cue was simply due to sensory summation by cue and target stimuli. These results suggest that exogenous attentional enhancement is regulated by response gain modulation within striate cortex, consistent with the predictions of the normalization model of attention.

R01EY028163; R01EY035640

26.405 SLEEP PROFOUNDLY ALTERS CORTICOTHALAMIC RESPONSES TO VISUAL STIMULI.

Nicholas G Cicero[1,3,4] (ngcicero@bu.edu), Michaela Klimova[2], Sam Ling[1,2], Laura Lewis[3,4,5,6]; [1]Graduate Program in Neuroscience, Boston University, [2]Department of Psychological and Brain Sciences, Boston University, [3]Department of Biomedical Engineering, Boston University, [4]Institute for Medical Engineering and Sciences, Massachusetts Institute of Technology, [5]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, [6]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital

While asleep, our brains withdraw from our external environment and presumably suppress the processing of external sensory information. Although sensory disconnection is necessary for our brain to transition into a sleep state, the mechanisms regulating sensory gating remain debated. While the eyelid is the first barrier to visual input while we sleep, with closed eyes, high luminance inputs still reach the lateral geniculate nucleus (LGN) and primary visual cortex (V1). Here, we investigated how arousal state impacts visual processing using simultaneous EEG-fMRI in humans resting with their eyes closed (n=15), across sleep and wakefulness. During an eyes-closed resting-state scan, we presented a temporal contrast modulated flickering stimulus while subjects naturally transitioned in and out of sleep. Subjects were also instructed to perform a self-paced behavioral task to track behavioral responsivity. We first found that while awake with closed eyes, flickering stimuli induced responses in LGN and in V1, whereas extrastriate cortical responses were suppressed. During non-rapid eye movement (NREM) sleep, the visual-evoked response pattern substantially changed: visually-evoked responses were largely preserved in the LGN, whereas we observed substantial stimulus-locked suppression in visual cortex. Specifically, the magnitude of stimulus-locked visuocortical suppression scaled with the stimulus intensity, wherein higher intensity stimuli induced even larger negative visually evoked responses. To further investigate when visuocortical deactivation occurred as subjects transitioned from wakefulness to sleep, we also separated trials by behavioral state and found that this visuocortical suppression occurred only when subjects were behaviorally unresponsive. During the fully unresponsive state, the EEG steady-state visual evoked potential was also significantly attenuated, and low frequency delta power, a signature of sleep depth, was greatest. Given that the LGN showed preserved visual- evoked activation across arousal states, these results suggest that cortical inhibitory circuits are a key mechanism by which visual cortex disengages from environmental stimuli during sleep.

26.406 INFLUENCE OF SPATIAL GROUPING ON MULTIPLE SALIENT DISTRACTOR INHIBITION DURING VISUAL SEARCH

Brandi Lee Drisdelle[1,2] (brandi.drisdelle@bristol.ac.uk), Martin Eimer[2]; [1]University of Bristol, [2]Birkbeck, University of London

In daily life, we encounter distractions that must be managed to successfully navigate our visual environment. Research has shown that we can avoid attentional capture by highly salient objects in visual scenes, for example, a red object among green objects. Most studies, however, focus on the inhibition of a single salient distractor. Recently, we demonstrated that inhibiting multiple salient distractors is possible (Drisdelle & Eimer, 2023), though it remains unclear whether these distractors influence each other during inhibition. We examined whether salient distractors could be spatially grouped by comparing displays with two salient distractors that were either adjacent or separated by a potential target (search item), using behavioural and electrophysiological measures. Observers searched for a shape-defined target among target-coloured distractors and two different-coloured (salient) distractors. Lateralised ERP activity associated with one salient distractor was isolated by placing one laterally and the other on the midline. The Pd component was used to track the time course of inhibition. Behaviourally, inhibition was measured using the capture-probe paradigm, where observers report probe letters superimposed on all shapes. Our results revealed no difference in the suppression effect (an impaired ability to report letters at the location of a salient distractor) based on whether salient distractors were adjacent or separated, indicating that salient distractors were effectively inhibited in both scenarios. During target search, electrophysiological results showed a reduced Pd amplitude for adjacent compared with separated distractors. We propose spatial grouping occurs when salient items are adjacent: neural activity is distributed across both distractors, eliciting a single Pd that is reduced because midline activity does not contribute to the Pd. In contrast, when salient distractors are separated, each elicits a distinct inhibition process, producing independent Pd components. Our findings demonstrate that salient distractors can influence one another, with perceptual grouping potentially altering underlying neural responses.

We acknowledge 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.

26.407 MICROSACCADES STRONGLY MODULATE BUT DO NOT NECESSARILY CAUSE THE N2PC EEG MARKER OF VISUAL-SPATIAL ATTENTION SHIFTS AS STUDIED IN PERCEPTION AND IN WORKING MEMORY

Freek van Ede[1], Siyang Kong[1], Baiwei Liu[1]; [1]Vrije Universiteit Amsterdam

The N2pc is a popular human-neuroscience marker of covert and internal spatial attention that occurs 200-300 ms after being prompted to shift attention – a time window also characterised by the spatial biasing of small fixational eye-movements known as microsaccades. To delineate the relation between these co-occurring spatial modulations, we conducted a combined EEG-eyetracking study where a central colour cue prompted covert or internal selection of a left/right visual target that was either visible (selection from perception) or held in working memory (selection from memory). We show how co-occurring microsaccades profoundly modulate N2pc amplitude during top-down shifts of spatial attention in both perception and working memory. At the same time, we show that a significant – albeit severely weakened – N2pc can still be established in the absence of co-occurring microsaccades. Thus, while microsaccade presence and direction strongly modulate N2pc amplitude, microsaccades are not strictly a prerequisite for the N2pc to be observed. Moreover, this relation holds no matter whether microsaccades also bring attended visual targets closer to the fovea (in perception) or not (in working memory).

This work was supported by an ERC Starting Grant from the European Research Council (MEMTICIPATION, 850636) and an NWO Vidi Grant from the Dutch Research Council (14721) to F.v.E.

26.408 NEURAL REPRESENTATIONS OF VISUAL STIMULI EXHIBIT A LEFT VISUAL FIELD BIAS

Shengxing Yang[1] (yang.shengxin@ufl.edu), Qiang Yang[1], Sreenivasan Meyyappan[2], George Mangun[2], Mingzhou Ding[1]; [1]J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, [2]Department of Psychology and Center for Mind and Brain, University of California Davis

The left visual field bias is a phenomenon where individuals tend to process visual information more efficiently when it is presented in their left visual field compared to their right visual field. The underlying neural mechanisms remain to be clarified. We recorded fMRI data in two experiments to examine whether neural representations of visual stimuli exhibited a left visual filed bias. In both experiments, a trial started with a cue, which instructed the participant to attend a location in either the left or the right visual field. In the first experiment, following a variable cue-target period, a grating appeared with equal probability in either the left or the right visual field, and the participant reported whether the spatial frequency of the grating was high or low if it appeared in the attended visual field and ignored the grating if it appeared in the unattended visual field. In the second experiment, at the end of the cue-target period, two rectangles appeared in the left and the right visual field, and the participant reported whether the rectangle in the attended visual field was horizontal or vertical. fMRI responses were estimated on a trial-by-trial basis and subjected to MVPA decoding analysis (high vs low spatial frequency in the first experiment and horizontal vs vertical in the second experiment). We found that in both experiments, when the attended stimuli appeared in the left visual field, the decoding accuracy in the retinotopic visual cortex was significantly higher than when the attended stimuli appeared in the right visual field, and this was the case both in the hemisphere contralateral as well as in the hemisphere ipsilateral to the attended stimuli. We have thus demonstrated that for the two experiments considered here the neural representations of visual stimuli exhibited a left visual field bias.

26.409 OSCILLATORY TRAVELING WAVES IN VISUAL SEARCH

Yue Kong[1], Kirsten Petras[1], Philippe Marque[2,3], Rufin VanRullen[4,], Laura Dugué[1,5]; [1]Université Paris Cité, CNRS, Integrative Neuroscience and Cognition Center, Paris, France, [2]University Hospital of Toulouse, Toulouse, France, [3]ToNIC (Toulouse NeuroImaging Center), Inserm, University of Toulouse 3, Toulouse, France, [4]CerCo, CNRS, Universite de Toulouse, Toulouse, France, [5]Institut Universitaire de France (IUF), Paris, France

Brain activity unfolds over both space and time, yet it is often approached from a single dimension with limited explanatory capacity. While the temporal properties of neural oscillations have been associated with visual attention, the potential functional relevance of their spatial propagation across the cortex –oscillatory Traveling Waves (oTW)– remains largely unexplored. A recent electroencephalography (EEG) study showed that alpha (8-12Hz) oTW propagate predominantly backward from frontal to occipital during sustained attentional orienting (Alamia et al., 2023; see also Fakche et al., VSS 2024). This raises the question about the functional role of oTW in inter-area communication, particularly during attentional exploration that requires interaction between frontal attentional regions and visual cortices (Fiebelkorn & Kastner, 2019; Kienitz et al., 2022). Here, we used Transcranial Magnetic Stimulation (TMS) to assess the causal link between attentional exploration, frontal eye field (FEF) activity and oTW. Human participants (N=16) performed a challenging visual search task, identifying the presence or absence of a target letter "T" among distractor "L"s (in left visual field), while their brain activity was recorded with EEG. At various delays after stimulus onset, a double-pulse of TMS (25 ms interval) was applied to the right FEF or to the vertex (control site). The behavioral results revealed a periodic pattern of FEF interference at ~6 Hz, i.e., TMS over the FEF (and not the vertex) modulated performance as a function of the stimulus-onset-to-pulse delay. We then characterized oTW properties (i.e., direction and speed) with optical flow analysis (Townsend & Gong, 2018; Gutzen et al., 2024), and observed that the patterns of oTW correlated with visual search performance. Together, our results contribute empirical evidence clarifying the causal role of oTW in inter-area neural communication during attentional exploration.

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) and was also supported by a doctoral fellowship from Université Paris Cité to Yue Kong.

Attention: Reward

Saturday, May 17, 2:45 – 6:45 pm, Pavilion

26.410 CAN FEAR LEARNING VIA MENTAL IMAGERY AFFECT SUBSEQUENT ATTENTION?

Laurent Grégoire[1], Leyla Ochoa[1], Shivam Pancholy[1], Liliana Hepburn[1], Steven Greening[2], Brian Anderson[1]; [1]Texas A&M University, [2]University of Manitoba (Canada)

Mental imagery plays a crucial role in emotions such as fear. Experimental research suggests that the content of visual experience is encoded similarly during perception and imagination. Whereas a visual percept is a mental representation elicited by an external stimulus, an imagined percept is a representation elicited internally that operates as a form of the corresponding visual percept and is associated with subjective vividness. Recent findings indicate that fear conditioning with imagined percepts generalizes to the corresponding visual percepts (as measured via skin conductance response and self-reported fear), despite the visual stimulus never being paired with the unconditioned stimulus. The purpose of the present study was to determine whether fear conditioning acquired via mental imagery could affect subsequent attention. Participants first completed a fear conditioning task in which an imagined CS+ (e.g., an imagined red square) was associated with shock and an imagined CS- (e.g., an imagined blue square) was neutral. Subsequently, they engaged in a visual search task. In Experiment 1, participants performed visual search for a shape-defined target. A singleton distractor was colored with a hue corresponding to either the imagined CS+ or CS-. No conditioning effect was observed at the group level, but the attentional effect toward the CS+ color was positively correlated with the ability to form vivid mental images as evaluated by the vividness of visual imagery questionnaire (VVIQ). In Experiment 2, we employed a more sensitive visual task and found that participants were unexpectedly biased toward the CS- color. This suggests that the CS- color was perceived as a safe signal, rather than a neutral stimulus as initially anticipated. Again, the attentional bias toward the CS+ color was positively correlated with VVIQ score, implying that attentional priority was linked to the ability to form vivid mental images during the conditioning phase.

26.411 DISSOCIATING REWARD AND AROUSAL EFFECTS BY EEG ALPHA POWER BUT NOT PUPILLARY RESPONSES

Yusuke Nakashima[1], Yuka Sasaki[1], Takeo Watanabe[1]; [1]Brown University

Reward and arousal, two major neuromodulatory signals in the brain, significantly influence various types of visual functions. The previously believed distinctions between reward and arousal have become less clear. Here, by manipulating the reward and sound levels, for the first time, we demonstrate that the effects of reward and arousal can be dissociated using EEG alpha power. To manipulate reward levels, participants in the higher-value group were deprived of food and water for four hours prior to the experiment, while those in the lower-value group were not. In each trial, a drop of water was given to both groups. To manipulate arousal levels, a clicking sound was either presented or omitted. All participants underwent the following three conditions: (1)

in the reward condition, water was provided. (2) in the arousal condition, a sound was presented, (3) in the reward+arousal condition, both water and sound were presented. During each trial, participants passively viewed a dynamic sequence of Mondrian patterns for 15 seconds, while pupil size and EEG alpha power in response to water and/or sound were recorded. Across all three conditions, pupil size increased similarly in both the lower- and higher-value groups, suggesting that pupillary responses do not reflect reward value. In contrast, EEG alpha power exhibited distinct patterns: in the higher-value group, alpha power significantly decreased from the pre-stimulus baseline in the arousal condition but increased in the reward condition. Conversely, in the lower-value group, alpha power decreased in both the arousal and reward conditions. Our findings indicate that alpha power effectively dissociates the effects of reward (water) and arousal (sound). Additionally, our results suggest that water induces not only a reward effect, reflected in increased alpha power, but also an arousal effect, reflected in pupil dilation and decreased alpha power.

NIH R01EY019466, R01EY027841, R01EY031705, NSF-BSF BCS2241417

26.412 EMOTION-INDUCED BLINDNESS EFFECT FROM EROTIC IMAGES IS NOT CAUSED BY DISTINCTIVE LOW-LEVEL IMAGE FEATURES

Jeff Saunders[1], Ho Ming Chan[1]; [1]University of Hong Kong

Erotic images produce strong temporal attention effects, which have been interpreted as emotion-induced blindness (EIB). However, the attentional effects of erotic images could alternatively be due to distinctive image properties. We distinguished these possibilities by comparing the effects of erotic images and neutral images with closely matched image properties. We used a convolutional neural network to generate pairs of erotic and non-erotic neutral images that have the same distributions of medium-level features. The erotic images showed nude people in a sexual context, while the non-erotic images depicted common objects without recognizably erotic content but with closely matched image features (e.g., 'nude chair'). We tested whether these two types of images produced similar distraction. We also varied the colors of the background images to test whether color distinctiveness affects initial capture of attention. Subjects performed an RSVP task with distractor images presented before target images with lag of 2 or 8. The distractor was either an erotic image, a matched neutral image, or a background image. Background images were presented with original colors (unmatched background) or were recolored to have the same color histogram as the distractor (matched background). We found that the erotic images produced a much larger decrease in target detection at lag 2 than the matched neutral images, and the effect was the same regardless of whether the background stream was color matched. The matched neutral distractors had a smaller but detectable effect when background images were landscapes, and no effect in a follow-up experiment that used background images from the same categories. Our results demonstrate that erotic images cause a strong EIB effect even when image properties are carefully controlled, and suggest that distinctive image features have little effect. We conclude that the attentional effects of erotic images are primarily due to the erotic content.

26.413 NO EVIDENCE FOR LOCATION-SPECIFIC MAIN SEQUENCE MODULATION BY MONETARY REWARD

Lukasz Grzeczkowski[1], Martin Rolfs[1]; [1]Humboldt-Universität zu Berlin, Germany

Saccades are rapid, ballistic eye movements with stereotypical kinematics: their duration increases linearly with amplitude, while peak velocity saturates exponentially, a relation known as the main sequence. While this relation holds generally true, individual saccades can differ in vigor, moving faster or slower than expected. Previously, we demonstrated that monetary reward for movement speed modulates saccade vigor, irrespective of saccade target location. This manipulation also altered the dynamics of presaccadic shifts of visual attention to the saccade target—before slow saccades, performance at the target rapidly declined. Here, we investigated if the effect of reward on saccade vigor can be specific to certain locations, and if so, if it affects presaccadic attention and visual saliency. In each trial, a cue indicated one of eight potential saccade target locations. A grating then briefly flashed at that location at varying times before saccade, probing presaccadic attention. Participants saccaded to that location, received feedback on their reward, and reported the grating's orientation. In two locations each (which remained constant for each participant), participants received monetary reward when their saccadic peak velocity reached 30% of their fastest or slowest saccades. This manipulation was unknown to the participants, implicitly incentivizing high or low vigor, respectively. Additionally, after each of the three experimental sessions, we measured reaction time and orientation discrimination performance at all locations to assess the saliency of these locations in a fixation task. Results showed that neither saccade kinematics nor presaccadic attention differed as a function of reward location, demonstrating that spatially localized incentives were unable to modulate behavior. However, we found a strong correlation between the peak velocity and presaccadic perceptual performance in the saccade task ($R = 0.82$). Performance in the fixation task improved across sessions but, again, did not differ between locations demonstrating no selective saliency manipulation by reward.

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) and the Heisenberg Programme of the Deutsche Forschungsgemeinschaft (grants RO 3579/8-1 and RO 3579/12-1) granted to MR.

26.414 PHYSICAL EXERTION AND SELECTION HISTORY MODULATE BEHAVIORAL AND NEURAL SIGNATURES OF SUSTAINED ATTENTION

Emily Machniak[1,2], Tom Bullock[1,2], Caitlin Gregory[1,2], Julia Ram[1], Barry Giesbrecht[1,2]; [1]University of California, Santa Barbara, [2]Institute for Collaborative Biotechnologies

Attention can be biased by arbitrary features of the environment (e.g., color) that we have prior experience with, especially if those features are associated with reward (Anderson et al., 2011). Attention can also be impacted by our physiological state, such that a transition from rest to physical activity can impact behavior and different stages of sensory and cognitive information processing (Bullock et al., 2015). While reward and physical exercise impact overlapping attentional mechanisms, the precise interplay is not well understood. Here, we investigated the impact of physical activity on experience-biased attention. Participants (n=35) completed a continuous performance task where they monitored a sequence of images (800 ms) surrounded by colored frames and responded to infrequent images with longer durations (1200 ms). During a learning phase, participants associated one color (e.g., red) with monetary reward and another (e.g., blue) with no reward. One week later, they returned for a test phase involving the same task, but with no prospect of reward and while riding a stationary exercise bike at different intensities (low and moderate). EEG data were acquired at the scalp throughout. In the learning phase, there was a steeper decline in target detection performance over the course of the trial blocks in non-rewarded relative to rewarded trials (p<.05). During the test phase, target detection was improved for previously rewarded colors (p=.003) but, the typical change in performance over time within the blocks was not observed (p=.33). EEG time-frequency analyses were computed for correctly detected target trials. Theta oscillations (~3-7 Hz), typically associated with cognitive control, and alpha oscillations (~8-14 Hz), typically associated with selective attention, before and after the critical target period were modulated by reward and physical exercise. Together, these results suggest that physical state and reward both influence neural signatures of sustained attention.

Research was sponsored by the National Institutes of Health (R21 DA043712-01) and U.S. Army Research Office under the contract W911NF-19-2-0026 for the Institute for Collaborative Biotechnologies.

26.415 **WHEN IT'S HARD TO SEE, WHAT MAKES US** LOOK HARDER? SHARED MECHANISMS BETWEEN PERCEPTUAL FLUENCY AND COGNITIVE CONTROL

Emma S. Wiedenmann[1], Rebeka C. Almasi[1], Myeong-Ho Sohn[1]; [1]The George Washington University

The conflict adaptation effect demonstrates flexible cognitive control, wherein the congruency effect is contextually reduced (Tae et al., 2022). Flexible deployment of cognitive effort is not limited to cognitive control tasks. Low perceptual fluency also requires effort. This flexibility has been observed in perceptual fluency through the Clarity Sequence Effect (CSE), where the difference between responses to clear and blurry stimuli is reduced after viewing a previously blurry stimulus (Dreisbach & Fischer, 2011). Contextual modulation has also been demonstrated for the Item-Specific Proportion Clarity Effect (ISPC). However, this contextual modulation has not yet been extensively and systematically explored with other cognitive control markers, such as the List-Wide Proportion Congruency Effect (LWPC). Furthermore, the explanation that these effects are likely subject to binding requires further investigation. The current study aimed to examine whether perceptual fluency exhibits similar contextual modulation to cognitive control. Participants performed a gender identification task using either clear or blurry stimuli. Experiment 1 sought to replicate the CSE by presenting an equal proportion of clear and blurry stimuli. Experiment 2 utilized a block design to investigate LWPC, with one block containing 80% blurry stimuli and 20% clear stimuli. Experiment 3 examined the ISPC using an intermixed design: Half the stimuli were mostly clear (80% clear, 20% blurry), while the other half were mostly blurry (80% blurry, 20% clear). Our findings revealed that the contextual modulation of perceptual fluency for the CSE and the LWPC was present only after a response repetition. In contrast, perceptual fluency was contextually modulated during the ISPC, occurring both after a response repetition and a different response. The other effects were attributed to binding. These results suggest that perceptual fluency and cognitive control share similar contextual modulation mechanisms when stimuli are modulated on an item-specific basis.

26.416 WITHIN-SUBJECT MANIPULATIONS OF PROACTIVE CONTROL DO NOT CHANGE NEGATIVE TEMPLATES BENEFITS: EXPLORING THE EFFECT OF REWARD ON NEGATIVE AND POSITIVE CUES

Matthew Tong[1] (mata22@lehigh.edu), Matthieu Chidharom[2], Nancy B. Carlisle[1]; [1]Lehigh University, [2]The University of Chicago

During visual search, knowledge of distractor features can form negative attentional templates, guiding attention away from distractors and resulting in improved performance. Previous research identified proactive control as a mechanism underlying negative templates, allowing for the anticipation of distractors to prevent attentional capture. In recent work, inter-individual differences in proactive control were associated with the response time (RT) benefits of negative templates, suggesting the importance of trait-level factors. In contrast, it is unclear whether state-level factors play a role. The goal of this study was to examine whether state-based changes in proactive control impact the benefits to RT from negative templates. To induce state-based variation in proactive control, we manipulated reward motivation in a cued visual search task across four experiments. Rewards for a correct response could either be low (75% of trials) or high (25%). The search cue, which varied in a blocked design, could match the target color (positive), the distractor color (negative), or neither (neutral). In addition to RT, we also measured the standard deviation of RT, given evidence that proactive control can be behaviorally measured using RT variability. In Experiments 1 and 2, using a point-based reward resulted in no significant effects on RT or RT variability. In Experiments 3 and 4, using a monetary reward instead led to faster RT and lower RT variability for high rewards compared to low in the positive cue condition, but not in the negative cue condition. We concluded that a more efficient engagement of proactive control based on reward was associated with positive templates, but not negative templates. Overall, these results provide new evidence that using negative templates is not modulated by within-individual fluctuations of reward motivation. Negative template use may be more related to a trait-based, inter-individual ability to engage proactive control efficiently, rather than state-based factors.

Attention: Divided, tracking

SATURDAY, MAY 17, 2:45 – 6:45 PM, PAVILION

26.417 SPATIAL DISTRIBUTION OF VISUAL ATTENTION IN CHILDREN AND TEENS

Valeria Reis do Canto-Pereira[1] (vrcantopereira@unb.br), Ronald Ranvaud[2], Luiz Henrique Canto-Pereira[3]; [1]University of Brasilia,

[2]University of São Paulo, [3]Brazilian Ministry of Development, Industry, Commerce and Services

Spatial visual attention is one of the key elements of visual perception, and its spatial allocation allow us to priority a given stimuli location (Carrasco, 2019). Although there are several evidences on how spatial attention is distributed in children and teens, there are some gaps about their ability to divide attention. This study was aimed to investigate how children and teens allocate their attentional resources by measuring reaction times under different experimental conditions. Participants were children aged 8 to 15 years (n=68) divided in different groups according to age. There were two different experimental conditions. In experiment I participants were asked to direct their attention towards a square frame subtending 4° of visual angle located in the center of the screen. In experiment II participants were instructed to attend, simultaneously, two square frames subtending 4° of visual located 10° to the right and left of the center of the screen. The task was to respond (key presses) to the onset of a target, a white dot subtending 0.2° of visual angle presented at 154 different positions, while always fixating a small cross in the center of the visual field. Stimulus duration was brief (100ms) to avoid eye movements and concomitant attentional shifts. Experiments were carried out in a counterbalanced way. Our results confirm previous findings of participants in the older group showed an increase in performance (faster reaction times and lower error rates). In experiment I, as expected, reaction times were faster in the attended region (central square frame). In experiment II (divided attention condition) younger participants were not able to disengage attention from fixation, whereas teen participants, in 15 years old group, showed divided attention. These results suggest that teens 15 years of age have reached the ability to divide attention as adults.

CNPq Grant # 142080

26.418 EXPLORING ATTENTIONAL REPULSION FOR FACES: TARGETS SHIFT AWAY FROM SIMILAR DISTRACTORS DURING SEARCH

Yifan Fang[1] (yifan.fang.gr@dartmouth.edu), Viola Störmer[1]; [1]Dartmouth College

Studies show that when a target stimulus is selected among similar distractors, attention is not tuned to the exact target feature but away from it, in order to increase the signal-to-noise ratio between targets and distractors. To date, most research investigating this “optimal tuning” account of attention have focused on relatively simple visual features, such as color or orientation (Chapman et al. 2023; Navalpakkam & Itti, 2007). In the present study, we investigated whether similar repulsion effects would occur for complex, higher-level visual representations such as faces. Participants (N=32) performed a visual search task in which they were presented with an oddball target face among three distractor faces for 1500ms and were asked to select the location of the target. On each trial, the target was chosen randomly from a circular continuum of emotional faces (ZeeAbrahamsen & Haberman, 2018). The distractors were selected to be either similar (±60°) or dissimilar (±120°) from the target. To test how search context modulates the perception of the target on correct search trials, participants completed a 2-alternative forced choice task reporting which of two faces – the target or the foil – was most similar to the target immediately after each trial. Critically, the foil was offset by 15° or 40°, either away or towards the distractor. If attention is tuned away from similar distractors, participants should choose the foil more often when it is rotated away from the distractor relative to when it is rotated towards the distractor. Our results showed that this was the case for the 40° foil condition ($p = 0.01$) and to some extent for the 15° foil condition ($p=0.13$). As predicted, there was no effect when search distractors were dissimilar. Our results suggest that the repulsion effect found for simple features can be generalized to faces.

26.419 FADING ENHANCEMENT? EXPLORING THE IMPACT OF TOUCH TIMING ON TARGET ENHANCEMENT IN MULTIPLE-OBJECT TRACKING (MOT)

Mallory E. Terry[1], Lana M. Trick[1]; [1]University of Guelph

Many everyday tasks such as driving a car or playing team sports require keeping track of the positions of several independently moving items among others. This ability to select and keep track of the locations of multiple targets among identical non-targets (distractors) is called multiple-object tracking (MOT) and is thought to provide critical location information for performing actions toward the tracked targets (e.g., Pylyshyn, 2001). In support of this, our lab found reduced MOT performance when participants had to touch targets in MOT while tracking as compared to touching distractors. Though the theoretical framework supporting MOT is debated, several studies have found support for attentional enhancement of target locations during MOT. In the present study, we sought to investigate the impact of touch on the attentional enhancement of MOT targets by modulating what item was touched (target, distractor in MOT) and when the touch occurred in the trial (early, late, or both early and late). We hypothesized that if the attentional enhancement of targets decreased over time and was impacted by touch, touches that occurred later in the trial would have a more considerable impact on tracking performance relative to those that occurred earlier in the trial. In support of this, error rates were significantly lower for targets that were touched later in the trial compared to earlier. Interestingly, the impact of touch timing differed based on the item in MOT that was touched. For touched distractors, error rates did not differ based on when the distractor was touched, but instead RTs to decide if the touched distractor was a target were slower when it was touched later in the trial. Taken together, these findings provide evidence of a shared mechanism employed in MOT and visually guided touch that may be differentially impacted the item touched and the timing.

26.420 THE IMPACT OF ATTENTIONAL FLUCTUATIONS ON PERFORMANCE IN AN IMAGE FLANKER TASK

Avalon Nisenbaum[1], Susanne Ferber[2]; [1]University of Toronto

Sustained attention is essential for tasks requiring prolonged focus. Yet, our attention fluctuates, even for simple tasks, as our minds shift in and out of focus. Such lapses in attention are called mind wandering (MW) and are often associated with performance declines. The flanker task is a well-established paradigm to study the ability to suppress irrelevant information and respond accurately to the target stimuli. When combined with MW assessments, it has been found that during MW periods, performance on incongruent trials decreases. Previous studies have often used symbolic stimuli, such as arrows, numbers, or letters, which may limit transfer to real-world situations. Here, we tested whether “in-zone” (high task engagement) and “out-of-zone” (low task engagement) periods affect performance in a task involving images of indoor/outdoor scenes and living/non-living items. Participants (N = 47) completed an image Flanker task designed to assess sustained attention through response times (RTs) and accuracy. Results showed significant differences between “in-zone” and “out-of-zone” conditions. Participants in the "in-zone" condition exhibited faster RTs and higher accuracy compared to the "out-of-zone" condition. In the “in-zone” condition, no significant differences in RTs were found between congruent and incongruent trials, whereas the “out-of-zone” condition showed significant RT differences, with incongruent trials slower than congruent ones. For accuracy, no significant differences were found in either zone. These results align with theoretical models, suggesting that attentional fluctuations directly impact task performance. Our findings underscore the dynamic nature of sustained attention and its significant influence on task performance. The clear distinctions between “in-zone” and “out-of-zone” conditions highlight the critical role of attentional fluctuations in shaping behavioural outcomes. Future studies will identify the ideal duration of attention-demanding tasks and effective break strategies that could inform interventions to enhance focus and productivity across various settings.

26.421 ATTENTION TO SHAPE OR LOCATION ENHANCES SPATIAL DISCRIMINATION IN VENTRAL AREAS: A 1-BACK FMRI STUDY

Sung-Mu Lee[1], Ishita Agarwal[1], Jinho Lee[1], Anne B Sereno[1]; [1]Purdue University

Attention is a fundamental cognitive function that allows humans to filter irrelevant information and focus on task-relevant stimuli. It plays a crucial role in how we process complex visual environments, influencing the neural representations of the visual input. Visual information is processed via two segregated pathways in the brain: the ventral stream, which is primarily responsible for processing object features (the “what” pathway), and the dorsal stream, which focuses on spatial location (the “where” pathway). However, recent explanations posit that both pathways process shape and space, with attention or task context influencing the differential representations observed1,2 and more so in the ventral pathway3. This study aims to investigate whether and how attention, directed either toward the shape or location of an image, modulates neural representations of spatial locations in these two pathways. Participants underwent MRI scanning while performing a 1-back task under two attention conditions using identical visual stimuli. In the “attend to shape” condition, participants were instructed to attend to the shape of the images and press a button when the same shape was repeated. In the “attend to location” condition, participants focused on the location of the images and pressed a button when the same location was repeated. Multivariate Pattern Analysis (MVPA) with Multidimensional Scaling (MDS) was applied to assess spatial encoding in both ventral and dorsal regions under both attention conditions. The MDS results revealed increased distance of spatial locations in “attend to shape” and “attend to location” conditions compared to a “passive viewing” condition in ventral but not dorsal regions. This suggests that while both the ventral and dorsal pathways are involved in spatial processing, attention (to either shape or location) selectively increases discrimination of the neural representations of spatial locations in the ventral pathway.

NIH CTSI and Purdue University

26.422 PRESSURE IN THE SPOTLIGHT: HOW MONITORING AND OUTCOME PRESSURES IMPACT TIME-SHARING PERFORMANCE

Lauri Oksama[1,3] (lauri.oksama@mil.fi), Niki Pennanen[2,3]; [1]Finnish Defence Research Agency, [2]Aalto University, Department of Computer Science, [3]University of Turku, Department of Psychology and Speech-Language Pathology

Performing under pressure, particularly in multitasking environments, is a critical challenge in both everyday life and high-stakes professions. This study investigated the differential effects of monitoring and outcome pressure on time-sharing performance and allocation of visual attention. Using a within-subjects design, 30 participants completed a recently devised time-sharing task requiring prioritization under three different pressure conditions. We hypothesized that in a high-demand time-sharing environment, outcome pressure would impair task performance and visual sampling of subtasks more significantly than monitoring pressure. To investigate our hypotheses, we recorded participants’ task performance metrics and eye movements during the task. However, our confirmatory analyses found no evidence to support either of these hypotheses. In contrast to the hypotheses, our additional exploratory analyses revealed that monitoring pressure, rather than outcome pressure, led to a statistically significant decrease in task performance. Notably, this effect occurred without detectable changes in visual sampling. This unexpected finding is likely due to the high sensorimotor demands of the task, specifically the need for precise and rapid mouse movements, which may have been disrupted by the participants’ heightened self-consciousness under monitoring pressure. Our findings contribute to the growing body of literature on the differential effects of monitoring and outcome pressure, with potential implications for real-world settings. For instance, in professions requiring precise motor skills—such as surgery, piloting, or competitive sports—heightened self-consciousness under monitoring pressure may impair performance, even without affecting attentional control. Similarly, everyday activities like driving under observation (e.g., during driving tests) or performing in front of an audience may be adversely affected. Understanding how monitoring pressure disrupts performance in such scenarios can inform training and support strategies to mitigate its impact.

26.423 MAGNITUDE OF TARGET-DISTRACTOR CORRELATION DURING RSVP INFLUENCES TARGET PERCEPTION

Michele Maslowski[1,2], David H. Hughes[1,2], Adam S. Greenberg[1,2]; [1]Marquette University, [2]Medical College of Wisconsin

During RSVP experiments (Chun & Potter, 1995), a target image appears among temporally presented distractors. The assumption is that items are processed independently, however evidence of an attentional blink (Raymond, Shapiro, & Arnell, 1992) suggests interactions between target and distractor items. Here, we hypothesize that if target image content is highly correlated with distractor images, target perception and task performance will be affected. Overlapping face and house images were embedded in noise and presented via RSVP. Each frame contained one face and one house image chosen from 32 possible images for each stimulus type; with two specific face and house images as temporal search targets. Subjects attended to either the face or house stream (cued via colored image frame) and identified targets. Frame duration (500 ms, 600 ms, 750 ms, 900 ms, 1000 ms, 1200 ms) was varied across 9 blocks. The distribution of correlations between target and distractors varied, though faces were more strongly correlated with distractors (Face 1: $\mu = 0.623$, $\sigma = 0.0567$; Face 2: $\mu = 0.476$, $\sigma = 0.0522$; House 1: $\mu = 0.3168$, $\sigma = 0.079$; House 2 : $\mu = 0.2884$, $\sigma = 0.0621$). We evaluated the correlation between target and subsequent frame (or frame to which subjects responded, for FAs) for the following response outcomes: hits (H), false alarms (FA), and misses (M). We then calculated the difference in correlation between each combination of response outcomes (FA-H, FA-M, H-M). We observed a significant interaction between target face and response outcome ($p < 0.01$). Significant positive difference scores were observed for FA-H ($p < 0.01$) and FA-M ($p < 0.01$) comparisons, suggesting that higher correlation values elicit a FA response. House targets failed to produce significant interactions, suggesting lower correlations limit influence on performance. Thus, RSVP target-distractor correlations affect target perception.

26.424 FUNCTIONAL CONTRIBUTION OF THE SUPERIOR PARIETAL LOBULE TO VISION

Laure Pisella[1], Tristan Jurkiewicz[1], Yaffa Yeshurun[2]; [1]Centre de Recherche en Neuroscience de Lyon, [2]University of Haifa

The Superior Parietal Lobule (SPL) is an associative area that is part of two identified functional systems: the dorsal visual pathway and the dorsal covert attention network. The dual visual pathway theory proposed that the occipito-parietal pathway is involved in mechanisms of "vision for action" but not in those of "vision for perception", highlighting that patients with optic ataxia (OA) are impaired for reach-to-grasp movements after SPL damage. More precise and recent investigations have shown that the visuo-motor impairments in OA are characterized by pointing errors in peripheral vision, underestimating the visual eccentricity of targets. Additionally, OA patients display, in their affected peripheral visual field, a perceptual underestimation of object size, and a deficit of covert attention revealed by slower visual processing despite valid cueing. We hypothesized that most of the deficits observed in OA patients in their contralesional visual field stem from a common dysfunction in the dorsal covert attention network, of which the SPL is a part. In two experiments we investigated the nature of peripheral pointing errors in the absence of covert spatial attention using a pointing task combined with central high-predictive and peripheral low-predictive attentional precues. These cues either indicated the correct pointing target location or a different location. We observed pointing errors mimicking the target position underestimations typical of OA through invalid endogenous cueing in control subjects. Moreover, in OA patients, valid cueing improved peripheral target-pointing performance. Low-predictive cues affected RT but not pointing accuracy. In a third experiment, we used similar predictive cues to modulate the perceived size of peripheral objects in control subjects. We found that invalid endogenous cueing shifted the responses toward a perceptual underestimation of object size. According to these results, we propose that SPL-based endogenous attention counteracts cortical magnification—i.e. actively re-magnifies the representation of the peripheral field.

This work has been supported by Fondation Neurodis. No commercial interests and conflicts to declare.

26.425 MONOCULAR ADVANTAGE FOR MULTIPLE OBJECT TRACKING

Guiping Zheng[1] (2021023767@m.scnu.edu.cn), Rong Jiang[1], Ke Zhou[2], Shuai Chang[3], Xinping Yu[3], Liqin Zhou[2], Ming Meng[1]; [1]Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), School of Psychology, South China Normal University, Guangzhou 510631, China., [2]Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, Beijing 100875, China., [3]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, Guangzhou 510060, China.

The Multiple Object Tracking (MOT) paradigm is widely used to explore the allocation and deployment of attention to multiple moving objects in dynamic scenes. Albeit extensive, previous research has led to conflicting views regarding the initial processing stage at which attentional modulation influences MOT — specifically, whether it occurs in early visual regions or at later processing stages. To address this question, we employed a dichoptic MOT paradigm, where 0 to 4 tracking targets and 4 distractors were presented to one eye, and the remaining targets (totaling 4) and another 4 distractors to the other eye. In a switch condition, the retinal projection of target stimuli swapped once between eyes, while in a non-switch control condition, targets presentation remained within the same eye, and visual perception maintained consistent across both conditions. In Experiment 1, such interocular target switching significantly impaired tracking performance when stimuli moved rapidly, revealing a monocular advantage for MOT. Experiment 2 examined whether this effect extended to distractor switching and found no significant results, suggesting that the monocular advantage might be specific to target switching. In Experiments 3a-3c, we introduced jittering of stimulus contrast to control for dichoptic contrast sensitivity differences, and further investigated how different proportions of the target allocation between the two eyes affected the magnitude of the monocular advantage. We found the monocular advantage was amplified when targets were equally distributed between the two eyes, with tracking accuracy in the switch condition decreasing by over 40% compared to the non-switch condition. These findings suggest that attentional modulation in MOT occurs at an early stage of visual processing, i.e. at or before monocular processing in the primary visual area (V1), and may involve subcortical structures, such as the lateral geniculate nucleus (LGN), which are mostly monocular.

This work is supported by (1) the National Nature Science Foundation of China (Grant No. 31871136), (2) the Science and Technology Ministry of China (Grant No. 2021ZD0204200), and (3) the Sino-German Center for Research Promotion (Grant No. M-0705).

26.426 CDA FLIP BETWEEN HEMISPHERES REFLECTS THE IMMEDIATE DEMANDS OF ATTENTIONAL TRACKING ACROSS VISUAL FIELDS

Huiqin Chen[1], Piotr Styrkowiec[2], Edward K. Vogel[1]; [1]University of Chicago, [2]Uniwersytet Wrocławski

The Contralateral delay activity (CDA) of the human EEG has been used to examine the handoff of attended targets as they move from one visual field to the other (Drew et al, 2014). Specifically, CDA is a sustained negative voltage over the hemisphere that is contralateral to the positions of the tracked targets. As tracked targets begin to cross the vertical midline, the CDA shows an apparent polarity "flip" (i.e., contralateral positivity), in which the originally ipsilateral hemisphere now has the tracked target in its contralateral field. This CDA flip has been proposed to reflect a handoff of the attended information between the hemispheres which was triggered by the immediate attentional demands of tracking the target as it moved through space. However, in those studies targets that crossed visual fields were always ultimately tested in the new field. Consequently, the hemispheric flip during tracking could instead be driven by the future demands of making a target discrimination and response in the new hemifield. Here we tested between these alternative accounts of the CDA flip. Specifically, we tested whether the CDA flip would still be observed if the subjects knew in advance that targets would ultimately return to their original hemifield before being tested. We found that the CDA flip occurred as items initially crossed visual fields and then also flipped back again as the tracked targets returned to the original hemifield. These results are consistent with the account of the CDA flip as reflecting an attentional handoff in support of the immediate needs of tracking the position of targets rather than the future needs of making a response.

Multisensory Processing: Audiovisual integration

Saturday, May 17, 2:45 – 6:45 pm, Pavilion

26.427 COMPLEXITY IS A COGNITIVE UNIVERSAL: EVIDENCE FROM CROSS-MODAL TRANSFER

Tal Boger[1], Shari Liu[1], Chaz Firestone[1]; [1]Johns Hopkins University

What connects a sharply twisted shape, a many-layered melody, and the multisyllabic string “animipatorun”? These items are unrelated in nearly every aspect; they span different modalities, arise from different domains, and have independent properties. Nevertheless, they seem unified by their *complexity*: Each is informationally dense relative to prototypical stimuli of its kind (cf., a square, major scale, or short string). Does the mind appreciate the complexity these stimuli share, even across dramatically different properties? Here, 4 experiments demonstrate *transfer* across these different stimuli, suggesting that a ‘universal’ representation of complexity exists in the mind. In Experiment 1, participants learned a reward rule for simple and complex shapes; selecting a complex shape was worth more (or less) points than selecting a simple shape. After this learning phase, participants saw new stimuli that also differed in their complexity: two arrays of colored dots, one uniform and the other highly varied. Without any further instruction, subjects transferred the reward rule to the dots, spontaneously selecting the more complex (or simpler) dot array. Experiment 2 generalized this pattern to audition: Subjects who learned that complex shapes were worth more points spontaneously selected complex melodies. Experiment 3 extended this result even further, finding successful transfer from shapes to letter-strings. In each case, this transfer arose bidirectionally. Finally, Experiment 4 tested the automaticity of such transfer. In a Stroop-like task, two shapes of differing complexity appeared above two letter-strings of differing complexity, and participants judged which shape (or letter-string) was more complex. Though only one stimulus class (either the shapes or the letter-strings) was task-relevant, participants were faster to judge the complexity of the target stimulus when the task-irrelevant stimulus was congruent in complexity. We suggest that visual, auditory, and linguistic complexity are ‘unified’ in the mind, supporting spontaneous and automatic transfer across modalities.

NSF BCS 2021053, NSF GRFP

26.428 AUDIO BLANKING EFFECT OF SACCADIC SUPPRESSION OF TARGET DISPLACEMENT

Hiu Mei Chow[1] (dorischm@gmail.com), Jialiang Ma[2], Satoshi Shioiri[2], Chia-Huei Tseng[2]; [1]St. Thomas University, Fredericton, Canada, [2]Research Institute of Electrical Communication, Tohoku University, Sendai, Japan

Human observers make saccades (large shifts of eye gaze) frequently as we explore the world visually, which creates motion blur or smears on the retina. However, we perceive a stable visual world and do not notice motion blur. That is because our brains selectively block visual processing during eye movements (i.e. saccadic suppression). One example of saccadic suppression is our inability to detect a target location change during saccades (i.e. saccadic suppression of target displacement), which might be explained by the assumption that the world is stable. As our world is multi-sensory and non-visual signals like sound are important elements for a stable world, we tested whether sound affected the visual perception of target displacement during eye movement in an experiment. Sixteen observers reported whether a visual target jumped to the left or right (0.33°) when they made a saccade (18°). Critically, a pure tone was presented synchronously with the visual target, with a 100-ms gap at the time of saccade (audio-blank), potentially signalling a change in the visual event. Other times, the sound was constant (audio-no-blank, signalling no change) or absent (silent). Perceptual performance analysis showed that observers’ ability to discriminate target shift direction was significantly higher in the audio-blank condition than in the silent (p = .002) and audio-no-blank (p = .037). Despite a shorter saccade latency in the audio-blank than in audio-no-blank conditions ($p < .001$), saccade latency did not predict perceptual task accuracy (p = .10) on a trial-by-trial basis after sound manipulations were accounted for. These results highlight an auditory counterpart of the visual blanking effect of saccadic suppression, where a brief gap in the visual target at the time of saccades leads to better visual discrimination. Our work supports the role of sound in modulating visual stability perception during saccades.

Part of this work was carried out under the Cooperative Research Project Program of the Research Institute of Electrical Communication, Tohoku University (R05/A17) and Discovery Grant Program of Natural Sciences and Engineering Research Council of Canada (RGPIN-2024-06028).

26.429 AUDIOVISUAL INTEGRATION IN MUSICIANS: IMPROVED PRECISION AND FLEXIBILITY IN AUDIOVISUAL TEMPORAL PROCESSING.

Subin Jeon[1], Hyun Ji Kim[1], Chai-Youn Kim[1]; [1]Korea University

Sensory recalibration allows the brain to compensate for temporal mismatches in sensory inputs to maintain a coherent perception (Vroonment & Keetels, 2010). While previous studies suggest that musical training may enhance recalibration through frequent training of complex rhythmic patterns (Jicol et al., 2018; O'Donohue et al., 2022; O'Donohue, 2024), the extent of this enhancement remains underexplored. To address this, we first replicated the double-flash illusion (Bidelman, 2016) to compare recalibration abilities between musicians (n=13) and non-musicians (n=10). Although the difference was not statistically significant, musicians demonstrated greater resistance to the illusion, more accurately reporting the number of flashes in illusory trials. Although this illusion measures audiovisual interactions, it primarily involves a visual task, with auditory stimuli being irrelevant. Thus, we conducted an additional study using the audiovisual correspondence detection task (Denison et al., 2013), where participants were asked to identify which of two Gabor patches oscillated in the rhythm of a given auditory reference stream. Stimulus onset asynchronies (SOAs) of the auditory tones and Gabor tilts were set within ±300ms. Our results showed that musicians demonstrated better sensory recalibration in the audiovisual correspondence detection task compared to non-musicians, as reflected in their higher performance accuracy ($p = 0.038$). This indicates that musicians effectively compensate for temporal misalignments across a broader range of SOAs. These findings suggest that musical training improves both the range and precision of sensory recalibration, enhancing the ability to adapt to temporal mismatches across modalities. In other words, this implies that musical training makes recalibration abilities flexible and adjustable to various task requirements.

NRF-2023R1A2C2007289

26.430 DIVIDING ATTENTION BETWEEN VISION AND AUDITION IN SPATIAL LOCALIZATION TASKS

Taylor J. Knickel[1,2] (knick071@umn.edu), Gordon E. Legge[1,2], Ying-zi Xiong[3]; [1]Department of Psychology, University of Minnesota, Minneapolis, MN, United States., [2]Center for Applied and Translational Sensory Science, University of Minnesota, Minneapolis, MN, United States., [3]Lions Vision Research and Rehabilitation Center, Wilmer Eye Institute, Johns Hopkins University

Spatial localization of environments often involves attending to distinct auditory and visual stimuli simultaneously. How effectively do observers attend to widely separated but simultaneous visual and/or auditory events? If spatial attention processes are separate for vision and audition (modality-specific), a minimal cost would be expected in attending to one visual and one auditory (bimodal) target pair than for two visual or two auditory (unimodal) target pairs. If spatial attention processes are shared for vision and audition (cross-modal), the expected cost for bimodal targets would be larger than unimodal targets. 19 participants with normal vision completed a spatial localization task, verbally reporting the direction (azimuth) of a single visual or auditory target, a pair of unimodal targets or a pair of bimodal targets. Auditory cues were a piano or violin G5 note presented at 60 dB SPL, while visual cues were an open or closed circle subtending 3 degrees, all presented for 500 ms in each trial. Reported error was computed as the absolute difference between the reported target location and the actual target location in degrees. When locating bimodal targets, both visual and auditory performance remain unaffected compared to single target conditions, regardless of spatial separation. When locating two simultaneous visual targets, there was no significant difference in reported error ($M = 6.86$) compared to bimodal conditions ($M = 6.12$). However, locating two simultaneous sounds led to significantly larger errors ($M = 28.28$, $p < 0.001$), that increased with spatial separation, compared to bimodal conditions ($M = 16.07$). The results point towards separate attentional processes for vision and audition (modality-specific hypothesis). Minimal cost was found in locating two simultaneous visual and auditory targets that are spatially distinct, regardless of locations and separations. Increasing the difficulty of the bimodal localization tasks may reveal a cross-modal cost in future studies.

NEI R00EY030145

26.431 EXPLORING THE NEURAL CORRELATES OF CROSSMODAL ILLUSORY PERCEPTION

Lara C. Krisst[1] (lkrisst@caltech.edu), Ailene Y. C. Chan[1], Noelle R. B. Stiles[1,2], Carmel A. Levitan[3], Shinsuke Shimojo[1]; [1]California Institute of Technology, [2]Rutgers University, [3]Occidental College

Previous research on multisensory integration reveals that participants often perceive an 'illusory' third visual stimulus when two visual flashes are presented at different locations, with three auditory tones (Stiles et al., 2018; Audio-visual Rabbit Illusion (AV Rabbit)). This crossmodal interaction was shown to be postdictive (wherein a later stimulus influences perception of an earlier stimulus). What is yet unknown is how this type of integration is represented in the brain. To this end, we use EEG in combination with multivariate pattern analysis to track the latency of the visual representations during the classic AV Rabbit paradigm. Participants are shown flashes moving along a horizontal/vertical trajectory and asked to report how many visual flashes they perceived. In the current task, participants are presented with two teardrop-shaped stimuli that rotate around a central origin, resembling the motion of clock hands. If participants perceive three teardrops (i.e., two real, one illusory in-between), they are instructed to report the orientation of the second (illusory) teardrop in a 360° space. We decoded the orientation information at each moment in time and compared trials in which participants reported seeing a third 'illusory' flash versus when they perceived only two flashes. Trials were physically identical (i.e., 3 beeps, 2 flashes), except for differences in the presence or absence of the 'illusory' flash. If the 'illusory' stimulus is postdictive in nature, we might expect this to influence the decoding accuracy of the visual stimuli. AV Rabbit behavioral effects were replicated with the teardrop-stimulus. Preliminary EEG findings show differences in decoding accuracy between the trials in which illusions were perceived and those in which perception was veridical, with earlier increases in decoding accuracy on illusory trials. Further investigation of the temporal dynamics of crossmodal illusory perception is required to fully elucidate mechanisms of postdiction in the brain.

26.432 HOW DO ARTIFICIAL NEURAL NETWORKS (ANNS) RESPOND TO AUDIOVISUAL ILLUSIONS SUCH AS THE MCGURK EFFECT?

Haotian Ma[1] (haotian.ma@pennmedicine.upenn.edu), Zhengjia Wang[1], John. F. Magnotti[1], Xiang Zhang[1], Michael. S. Beauchamp[1]; [1]University of Pennsylvania Perelman School of Medicine

Humans perceive speech by integrating auditory information from the talker's voice with visual information from the talker's face. Incongruent speech provides a useful experimental tool for probing multisensory integration. For instance, in the McGurk effect, an auditory "ba" paired with a visual "ga" (AbaVga) produces the illusory percept of "da". Artificial neural networks (ANNs) have made remarkable progress in reproducing human abilities and may provide a useful model for human audiovisual speech perception, prompting the question of how ANNs respond to incongruent audiovisual speech. To answer this question, we presented McGurk and congruent (control) stimuli to human observers and Audiovisual Hidden-unit Bidirectional Encoder Representations from Transformers (AVHuBERT), an ANN developed by Meta Corporation. Twenty McGurk stimuli were tested, consisting of a single "ga" video paired with different "ba" auditory recordings, all from the same female talker. Amazon Mechanical Turk was used to assess the perception of 128 human observers. To model individual differences in human perception, variants of AVHuBERT were created by adding Gaussian noise to the units in a transformer encoder layer of the model. Both human observers and AVHuBERT classified each stimulus as "ba", "ga" and "da". Performance was highly accurate for congruent syllables (mean of 94% for humans and 94% for AVHuBERT). For McGurk stimuli (AbaVga), there was substantial variability in the rate of McGurk "da" reports across different human observers (range from 25% to 65%) and different model variants (16% to 48%). The overall rate of illusory percepts was similar (46% for humans, 34% for AVHuBERT). The similar "perceptual" reports of human observers and AVHuBERT to the McGurk effect suggest that ANNs may be a useful tool for interrogating the perceptual and neural mechanisms of human audiovisual speech perception.

26.433 THE INTERPLAY BETWEEN ATTENTION AND AUDIO-VISUAL OBJECT STORAGE IN WORKING MEMORY

Ceren Arslan[1] (arslan@ifado.de), Daniel Schneider[1], Stephan Getzmann[1], Edmund Wascher[1], Laura-Isabelle Klatt[1]; [1]Leibniz Research Centre for Working Environment and Human Factors, Dortmund, Germany

To navigate a multisensory world, the brain must integrate and maintain information from multiple senses. However, current working memory (WM) research primarily focuses on the visual domain. Thus, our knowledge of how multisensory information is encoded and maintained remains incomplete. Here, we conducted two EEG experiments to elucidate how audio-visual objects are stored in WM using an audio-visual delayed-match-to-sample task. In Experiment 1, participants were presented with audio-visual objects while attending to auditory features (attend-auditory trials), visual features (attend-visual trials), or both (conjunction trials). Behavioral results showed that task-irrelevant features interfered at recall, suggesting bottom-up encoding into WM. However, traditional ERP measures of unisensory WM load showed they were not actively maintained. When the task required a deliberate feature integration (conjunction trials), there were greater attentional demands, indicated by stronger alpha power suppression at recall. In Experiment 2, we added unisensory controls (auditory- and visual-only) and manipulated the spatial arrangement of tones and orientations to be compatible or disparate. Performance declined in the attend-visual condition with spatially disparate features, while attend-auditory and conjunction conditions were unaffected by spatial manipulation. Representational similarity analysis indicated that task-irrelevant tones (in attend-visual trials) were reactivated at recall, while task-irrelevant orientations (in the attend-auditory trials) were more consistently filtered out. Contrasting the behavioral effects, this pattern was unaffected by the spatial compatibility between auditory and visual features. In sum, these findings highlight the complex dynamics of selective attention in modulating multisensory feature integration and audio-visual object storage in working memory.

26.434 TOUCHING SOUNDS: EXAMINING HOW THE AMOUNT OF AUDIO-VISUAL EXPOSURE INFLUENCES THE DEVELOPMENT OF AUDIO-TACTILE SOUND-SHAPE CORRESPONDENCES

Vivian M Ciaramitaro[1] (vivian.ciaramitaro@umb.edu), Rong Tan[2], Shibo Cao[1]; [1]University of Massachusetts Boston, Department of Psychology, Developmental and Brain Sciences, [2]University of Massachusetts Boston, Department of Counseling and School Psychology

How do we know what sensory information goes together? In addition to spatial co-localization and temporal synchrony, we rely on crossmodal correspondences: how features are shared across the senses. For example, nonsense words, such as "bouba", are associated with round abstract shapes and "kiki" with angular shapes. Such associations are observed for auditory and visual stimuli (AV), as well as auditory and tactile stimuli, which are touched but not seen (AT). Visual experience influences AT associations: AT associations are weak in early-blind adults (Fryer et. al., 2014) and in fully-sighted children with naïve visual experience, but are enhanced if fully-sighted children see the shapes first (16 trials of prior AV exposure; Chow et. al., 2021). Here, we examine how the amount of prior visual exposure influences AT associations. Sixty-one 6-8 year-olds completed 4 or 8 trials of AV exposure, seeing a round and spiky shape on a screen and judging which shape best matched a sound. Following exposure, children completed 16 AT test trials, feeling two shapes inside a box and judging which shape best matched a sound. No feedback was provided during exposure or test trials. We found that 8, but not 4, trials of prior AV exposure enhanced AT associations, while neither 4 nor 8 trials of prior AT exposure enhanced AT associations (Chow et. al, 2021 data re-analyzed). Thus, children did not benefit from repeated AT exposure, unlike blind adults where AT associations are enhanced, or blindfolded, fully-sighted adults where AT associations are diminished (Graven et. al., 2018). Children require enough experience with specific visual input to match sounds and shapes that they feel. These results complement our related work indicating that the type of prior exposure matters (see Cao et. al., VSS 2024). Future work needs to address mechanisms by which repeated exposure enhances crossmodal associations.

University of Massachusetts Graduate and Undergraduate Research Funds

26.435 VISUAL AUDITORY INTEGRATION IN SIMULATED AND NATIVE ULTRA LOW VISION

Yi xu[1] (yxu@sunyopt.edu), Navya Sreeram[1], Rob Chun[1], Roksana Sadeghi[2], Chau Tran[3], Will Gee[3], Gislin Dagnelie[4], Ione Fine[5,6], Arathy Kartha [1]; [1]Department of Biological and Vision Sciences, SUNY College of Optometry, [2]University of Berkeley California, Herbert Wertheim School of Optometry, [3]BaltiVirtual , [4]Wilmer Eye Institute, Johns Hopkins School of Medicine, [5]University of Leeds, Leeds, UK, [6]University of Washington, Seattle

Individuals with normal vision (NV) integrate information from their senses close to optimally. People with ultra-low vision (ULV) traditionally undergo blind rehabilitation, where they are taught to rely on tactile and auditory cues to compensate for their vision loss. This rehabilitation might help blind individuals to integrate optimally. Alternatively, being taught to ignore vision and rely on the other senses might lead to sub-optimal integration when there is remaining useful vision. Fourteen NV wearing ULV filters (sULV; 2.05 ±0.17 logMAR) and one ULV with advanced retinitis pigmentosa (2.0 logMAR, visual field <10 degrees) completed a spatial localization task in a virtual reality environment under unimodal (V = visual, A = auditory) and bimodal (VA = visual-auditory) conditions. Two flashing and/or ringing phones appeared in sequence in different locations of the central field (-4.5 to 4.5), and participants reported whether the first was to the left or right of the second. Percent correct as a function of the stimulus separation was fit using a cumulative Gaussian distribution. The unimodal standard deviations, σ, for the unimodal conditions (V and A), were used to predict optimal performance in the multimodal (VA) condition. For the sULV, mean σ was V = 1.5 ± 0.4, A = 7.6 ± 6.8, VA = 1.9 ± 0.5 with an optimal predicted VA = 1.5. There was a significant difference between predicted and empirical VA (paired t-test; $t(13) = 3.3$; $p=0.003$), showing that sULV were non-optimal in integrating visual and auditory cues. For the ULV participant, σ was V = 3.1, A = 6.1, and VA = 2.2, with an optimal predicted value of 2.7. These suggest that through years of living with her condition, the native ULV may have learned to integrate optimally, contrary to sULV, who had little time to adapt to filters.

26.436 VISUAL IDENTITY OF OBJECTS FACILITATES EARLY AUDITORY PROCESSING OF CONGRUENT SOUND

Mincong Wu[1,2], Andrew Marin[1,2], Jacob Momsen[2,3], Viola Störmer[4], Seana Coulson[2], Leslie Carver[2]; [1]Columbia University, [2]University of California, San Diego, [3]Yale University, [4]Dartmouth College

An object's visual identity can elicit expectations of contingent sounds. We asked whether auditory responses are sensitive to auditory contingencies that are linked to the visual identity of an object. We were guided by the hypothesis that behavioral and neural responses are facilitated when sensory input matches expectations relative to violations. 20 neurotypical adults were exposed to audio-visual (AV) contingences in an exposure- test phase design. During exposure, participants viewed pairs of shapes that spun to generate high- or low-pitch tones that were predicted by the object's shape. Participants engaged in a 2-alternative forced choice pitch classification task while high- density EEG was recorded. During test, participants were shown three conditions: audio-only, AV-match, and AV-mismatch (3-level between-subjects factor). AV-match trials maintained original shape-sound pairings from exposure, while AV-mismatch trials switched these pairings in 20% of trials. We conducted three linear mixed-effects models that used amplitudes associated with the P50, N100, and P200 ERP components as the dependent measures. Another mixed-effects model was used to assess reaction times from the 2-AFC task across individual trials of the test phase. An analysis of response time identified a main effect of condition ($p=.04$), such that RT's were faster for the AV-match condition compared to AV-mismatch, while the audio-only condition was not different from either. With regards to the EEG data, we observed a main effect of condition on the P50 amplitudes ($p<.001$), where amplitudes were greater for the AV- mismatch condition compared to AV-match ($p=.002$). We also observed a main effect of condition for the P200 amplitudes ($p<.001$), suggesting that the AV-match condition had smaller amplitudes compared to audio-only ($p<.001$) and AV-mismatch ($p=.02$). These findings show that an object's visual identity can facilitate early sensory processing of sound linked to that object, advancing our understanding of how visual cognition affects auditory perception.

Color, Light and Materials: Adaptation, constancy and cognition

SATURDAY, MAY 17, 2:45 – 6:45 PM, PAVILION

26.437 INDIVIDUAL DIFFERENCES IN COLOR PERCEPTION WITH NOVEL COLOR AXES

Jin Hirano[1], Masataka Sawayama[2], Kiyofumi Miyoshi[1], Shin'ya Nishida[1]; [1]Graduate School of Informatics, Kyoto University, Japan, [2]Graduate School of Information Science and Technology, The University of Tokyo, Japan

#TheDress has demonstrated significant individual differences in perceived colors, but the specific conditions that lead to similar variability remain unclear. Previous research has highlighted the importance of elongated color distributions along the daylight axis, but

#TheShoe questioned the generality of this rule. We examined the effects of color distribution axes on the individual variation in perceived color in various objects. We projected the color distribution of an object image onto a straight line along an arbitrary color axis passing through a white point in the CIE u*v* chromaticity diagram. By shifting the 1D color distribution along this axis, we generated a set of object images with continuously varying colors. A psychophysical experiment with the constant method was conducted with 30 participants, who answered perceived color from three options (e.g., blue, gray, yellow) at two distinct regions of the images. We estimated the magnitude of color shifts at which the response color was switched. This protocol was designed to quantify the state of individual color perception using a color-naming task that encourages surface matching (rather than appearance matching). Stimuli were four types of real-world objects, including Dress and Shoe, and control stimuli consisting of two color patches. There were six chromatic axis conditions, including the daylight axis. The results showed that the individual variation was larger for object images than for control stimuli. Within four objects, Dress and Shoe had larger variations than the other two novel objects. The large individual variation was observed not only for the daylight axis or the major axis of #TheShoe but also for other color axes (e.g., green–purple axis). The daylight axis showed a slightly larger variation, but only for Dress. Our results suggest that the individual difference in color perception is a more general phenomenon in object color perception than previously thought.

This work was supported by the KAKENHI Grant Number JP24H00721 and JP20H05957.

26.438 COLOR AFTERIMAGE IS BASED ON THE COLOR YOU PERCEIVE RATHER THAN THE ACTUAL COLOR OF THE OBJECT.

Neeranuch Kittikiatkumjorn[1] (neeranatk@gmail.com), Waragon Phusuwan[1,2], Anantaporn Sena[1], Arp-Arpa Kasemsantitham[1,3], Sedthapong Chunamchai[1]; [1]Cognitive Clinical & Computational Neuroscience Lab, Chula Neuroscience Center, King Chulalongkorn Memorial Hospital, [2]Medical Science, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, [3]Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Humans perceive color images through different wavelengths of light interacting with retinal cells, generating signals for the visual system, and subsequently creating color perception. However, prolonged exposure to a single color typically results in negative afterimages, or opponent colors, when the stimulus is removed. Whether these negative afterimages originate solely from retinal or jointly with cortical processes is still debated (Dong et al., 2017). One theory explains this phenomenon by retinal cone cell fatigue, where opposing pairs of cone cells inhibit the perception of the opposing color, resulting in the illusory perception of opponent colors after extended exposure to one color; meanwhile, statistical learning processes theorize the color illusions arise from the visual system's adaptation to prolonged exposure. Here, we investigated whether negative afterimages depend on the physical properties of color or on perceived colors influenced by color constancy — a phenomenon enabling varying color perceptions of the same physical color under different environmental conditions. We presented images of a Rubik's Cube with multiple-color squares illuminated by different lights; a grey square was placed alongside other colored squares to evoke perceptions of color constancy (yellow, blue, red, and green). After viewing these images followed by a white Rubik's Cube, eleven participants identified the afterimage color seen in the probe location. We found that the afterimage color of the grey square varied significantly based on color constancy ($F = 23.86$, $p < 0.001$). The colored squares that matched the light source color— appearing whitish due to color constancy – also resulted in a consistent white and black afterimage across different trials, demonstrating that the perceived afterimage color is more influenced by perceived color than by the actual physical color properties. These findings suggest that negative afterimage perception is driven more by cortical processes and statistical learning, rather than solely by retinal cell fatigue.

26.439 PINK ILLUSION, WHITE SHIFT

Stuart Anstis[1], Sae Kaneko[2], Patrick Cavanagh[3]; [1]UCSD, [2]Hokkaido U, [3]Glendon College, York U

A white or grey patch in a red surround looks green by simultaneous contrast. But a white disk centered in a rotating windmill of alternating red and white sectors looks faintly pink. And this pink spreads throughout the entire image. Reason: The trailing cyan after images of the red sectors quickly become the brightest regions in the image, so are taken as the anchor for ‘white’. They provide an index of the overall illumination that triggers a global recalibration of the achromatic point – a perceptual shift of the neutral white toward the color of the afterimage. Actual white areas then shift in the opposite direction and appear pink.

26.440 AN INVESTIGATION OF PERCEPTUAL EXPERIENCE IN HUE-ROTATED ALTERED REALITY

Yesesvi Somayaji Konakanchi[1,2,4] (yesesvi0808@outlook.com), Jenny Bosten[3,4], Anna Franklin[2,4], John Maule[1,4]; [1]Statistical Perception Lab, [2]The Sussex Colour Group & Baby Lab, [3]Sussex Vision Lab, [4]School of Psychology, University of Sussex

The visual system adapts to simple chromatic transformations (e.g., red filters), but whether adaptation occurs for more complex color transformations such as a hue rotation (e.g., blue sky turns magenta) is an open question. We conducted two experiments that exposed participants to real-time hue rotation in passthrough altered reality (AR) whilst doing normal everyday activities. A hue rotation of 120 or 240 degrees in HSL space was applied to a Meta Quest 3 AR headset’s look up table. In Experiment 1 we measured unique hue settings (yellow and blue) and categorical color constancy (colored chips were sorted into categories under different illuminations) before and after 4 hours of adaptation to the altered reality. In Experiment 2 we measured unique yellow and blue before, after, and during each hour of a 4-hour AR adaptation period. All psychophysical tests were conducted outside the headset in a dark room where all the color cues were removed to reduce de-adaptation. We also recorded participants’ introspections whilst immersed in a hue-rotated environment. Data were analyzed by comparing bootstrapped means in task performance pre- and post- adaptation, and Bayesian ANOVAs were used to investigate effects of adaptation on color constancy and unique hue settings. We found that adaptation to the AR environment did not disrupt categorical color constancy ($BF10(Exp1) < 0.001$) or settings of unique yellow and blue ($BF10(Exp1) < 0.3$; $BF10(Exp2) < 0.01$). Collectively, the results favor the null hypothesis that there is no adaptation to hue rotation, suggesting a limit to mechanisms of color adaptation. Directed content analysis of participants' introspections indicated that hue rotation causes confusion in object recognition, aversive responses to atypical food colors, and elements of derealization. The study highlights the potential of AR as a tool for understanding color perception and cognition.

Funding from the European Research Council (ERC) under the Horizon 2020 research and innovation programme (Project COLOURMIND: Grant agreement No. 772193, to A.F.)

26.441 THE ROSY-FINGERED DAWN: TWO PROBLEMS IN QUANTIFYING COLOR CONSTANCY

Adam Reeves[1]; [1]Northeastern University

As everyone knows, the environment appear 'rosy', or red-tinged, at dawn, even after full adaptation to the skylight. This alone demonstrates that colors are not constant, even if there is a tendency towards constancy. To capture this, two metrics are the constancy index CI=1−b/a and the Brunswick Ratio BR =c cos(θ)/a, where a, b, and c are sides of the color constancy triangle. Ideally, CI=BR=0 for no constancy and CI=BR=1 for perfect constancy. In the canonical Arend et al. (1991) study, LA made matches to 5 hues, each illuminated by an evening 4,000oK sky, a mid-day, bluish sky of 10,000oK, and a neutral sky of 6500oK. His mean CI, 78%, was close to the mean over 60 studies of 74% reported by Foster (2011). However, LA's mean BR was 95%, a gross exaggeration as his matches for 4 of the 5 hues fell well short. This problem is not unique to LA but typical. A further problem is that BR is indiscriminant: if the subject picks matches at random from RGB color space, the BR given (e.g.) a simulated 4,000 oK illuminant ranges across hues from +3.1 to -2.3, including +1.03 ('perfect color constancy') for a u'=.20, v'= .46 test patch. If the subject picks a single primary (R, G, or B) on every trial, BR ranges from −12 to +15, again straddling 1.00. In contrast, CI is always negative whether the subject picks at random from all of RGB space, from a limited region around white, or from any one spot, making it easy to filter data for bad responses. Recent papers have increased reported constancy towards 100%, but this is an artifact of employing BR, as can be easily seen by looking at the raw data. As a descriptor, CI is clearly superior to BR.

No funding, but inspiration from D. H. Foster

26.442 THE STRENGTH OF SIMULTANEOUS COLOR CONTRAST INCREASES WITH AGE AND REDUCED CHROMATIC SENSITIVITY: EVIDENCE FOR COMPENSATORY DEVELOPMENTAL PROCESSES?

Paolo Antonino Grasso[1], David Henry Peterzell, Federico Tommasi, Chiara Magnolfi, Linda Favillini, Rebecca Franconi, Elisabetta Baldanzi, Massimo Gurioli, Alessandro Farini; [1]University of Florence, Florence, Italy, [2]National Research Council, National Institute of Optics, Florence, Italy, [3]Fielding Graduate University, Santa Barbara, California, [4]National University, Pleasant Hill, California

During simultaneous color contrast, observers perceive 'illusory' color in a stimulus influenced by its surrounding context. This study investigated factors driving individual differences in these perceptions to elucidate underlying visual processes. Seventy-four individuals participated. In an achromatic test stimulus presented with eight different colored surrounds, a color matching task measured perceived hue and chromatic induction strength ('saturation'), while a color classification task assessed whether the stimulus appeared as a color or gray. Individual chromatic sensitivity was assessed using the Color Assessment and Diagnosis Test (CAD), which provided thresholds for red-green (RG) and yellow-blue (YB) channels. Individuals varied considerably in their color contrast perceptions. (1) Most participants saw colors complementary to the surround, though 10-30% indicated a gray appearance. (2) The strength of chromatic contrast saturation was inversely related to chromatic sensitivity, especially in the YB channel, where higher CAD thresholds correlated with stronger chromatic induction. Strength of saturation was positively associated with age, with older participants experiencing stronger illusions. (3) Saturation strength across hue angles showed very high intercorrelations, consistent with a broadly tuned factor mediating chromatic induction. A two-factor factor analysis identified highly intercorrelated factors for inducing reds and blues (F1: 180° to 360°) or yellows and greens (F2: 45° to 180°), not supporting mechanisms tuned to cardinal color axes or narrowly tuned chromatic channels.(4) Multiple regression analyses revealed that age mediated YB-saturation correlations, while RG-saturation correlations remained independent of age. To conclude, our results revealed that individuals with the lowest sensitivity in YB and RG channels experienced the strongest saturation, potentially reflecting a chromatic induction process compensating for reduced chromatic sensitivity.

European Union - PON Research and Innovation grant 2014–2020

26.443 INDIVIDUAL DIFFERENCES IN FOCAL COLORS AND COLOR NAMING

Camilla Simoncelli[1], Fatemeh, Charkhtab Basim[1], Erin Goddard[2], Michael, A. Webster[1]; [1]University of Nevada, Department of Psychology and Center for Integrative Neuroscience, [2]University of New South Wales Sydney, School of Psychology

Individual differences in color appearance (e.g. the stimulus that "looks" pure red) and color naming (e.g. the stimuli labelled "red") are well-established, but the extent to which these differences are related remains uncertain. We examined this relationship by comparing focal colors and color categories in 18 English-speaking color-normal observers, to ask whether differences in how colors are described or communicated can partly be predicted from how colors are perceived. Stimuli were defined in the LvsM and SvsLM cone-opponent plane and shown in a uniform field on a gray background. In the focal task, hue angle was varied at a fixed chromatic contrast to select the unique hues (pure red, green, blue, or yellow) or balanced binary hues (orange, purple, blue-green, yellow-green). In the naming task, stimuli spanning a wide range of hue angles and contrasts were each labelled with the same 8 terms plus gray. Variations across observers were significantly larger than within-observer variability for both tasks. However, with one exception (yellow), correlations between the focal hues and category boundaries were weak. Consistent with this, an algorithm designed to warp the color space to align appearance judgments (Simoncelli and Webster Color Res App 2024) did not result in increased consensus in color naming. These results suggest that inter-observer differences in appearance and naming reflect different underlying processes or criteria. Previous studies have emphasized the importance of correcting for both sensitivity and appearance differences to increase consensus in color rendering. However, the present analyses suggest that such corrections are unlikely to improve consistency across viewers in color naming, and thus point to the need to directly assess and include categorical variability as a third factor for predicting individual differences in color experience.

Funding: EY0108334 and ICC Danny Rich Memorial Scholarship Fund 2024

26.444 COMMUNICATING ABOUT COLOR AND TEXTURE

Angela Brown[1], Charles McDonald[1], Delwin Lindsey[1]; [1]Ohio State University

The surfaces of environmental objects generally have both texture and color qualities. Here we compare the naming of color to the naming of texture, then we use sorting game to ask which surface quality (texture vs. color) is more effectively communicated using single non-composite terms. 99 samples of materials varied in color and texture. 11 adults freely provided one or two terms for each sample. Information theoretic considerations suggested that color terms might be less effective for communication compared to texture, because each of 12 distinct color terms named 19 samples, whereas each of 17 distinct texture terms named only 6.5 samples. Each of 29 “other” terms named 32 samples. However, we found no evidence that texture and color are differentially communicated when participants actually interact. 46 dyads of adult participants played a sorting game with the 100 new samples. The first player of each dyad divided the samples into 2, 3 or 5 groups, based on similar color or similar texture, then named each of the groups they had created, following identical instructions for the texture sorts and the color sorts. The second player viewed the terms provided by the first player and sorted the samples into groups corresponding to the first player’s terms. Mutual Information (MI) quantified agreement between the two players’ sorts. Across all numbers of categories, MI for color sorting and texture sorting agreed well. The players of these dyadic games provided the same number of distinct terms per sample in the color and texture games. Our results emphasize the importance of “closing the loop” in studying communication about surface properties: simple naming results often miss important pragmatic features of dyadic interaction. Naming alone misses a lot of what communicators know about the properties of object surfaces.

26.445 DEVELOPMENT OF THE NBT-53 TEXTURE LIBRARY FOR THE STUDY OF TEXTURE SEMANTICS

Anna Chinni[1], Y. Ivette Colón[1], Kushin Mukherjee[1], Michael Gleicher[1], Karen B. Schloss[1]; [1]University of Wisconsin - Madison

Texture is useful for representing categorical data in information visualizations, especially when color display capabilities are limited (He et al., 2024). Prior work on texture for visualization mostly focused on which dimensions of texture could represent data effectively, with little focus on texture semantics—the meaning people ascribe to textures. To study texture semantics (cf. color semantics; Schloss (2024), we need a library of visual textures that are relatively uniform, are perceptually discriminable, and span dimensions of texture perception. To develop such a library, we began with the Normalized Brodatz Texture (NBT) database, a standard set of texture images that have been normalized for lightness (Abdelmounaime & Dong-Chen, 2013). We aimed to subset the 112 textures to select textures that were (1) uniform over the image (important for future use in data visualizations) and (2) perceptually dissimilar. Toward these goals, we first asked participants to rate the uniformity of each texture, and we excluded textures that were, on average, below the neutral point of the rating scale. To assess the similarity of the remaining 62 textures, we asked a second group of participants to make triplet similarity judgements. Using data from over 11,000 trials across 59 participants, we estimated a 3-dimensional embedding of the textures that best explained human judgements (Sievert et al., 2023). Although similar embeddings exist for a subset of the original non-normalized Brodatz textures (Ravishankar Rao & Lohse, 1996), we reasoned that the dimensions could be different for the normalized images. The 3-dimensions of our embedding were: fine/course, hard/soft, and random/non-random (named using data from different participants). Finally, we used k-means clustering on this embedding to identify highly similar textures and selected the most uniform texture within each cluster. This approach yielded a final set of 53 textures, which we call the NBT-53 texture library.

NSF award BCS-2419493 to K.B.S.

26.446 AFFECTIVE COLOR SCALES FOR VISUALIZATIONS OF CONTINUOUS DATA

Halle Braun[1], Kushin Mukherjee[1], Seth R. Gorelik[2], Karen B. Schloss[1]; [1]University of Wisconsin-Madison, [2]Woodwell Climate Research Center

Colors have affective connotations, which influence people’s evaluations of data visualizations (Bartram et al., 2017; Anderson & Robinson, 2021). For example, observers prefer visualizations in which affective connotations of the dataset (positive/playful vs. negative/serious) and colors representing the data are aligned. Previous studies of individual colors and palettes of discrete colors (e.g., palettes in categorical data visualizations) suggested that lighter, saturated colors were associated with positive concepts, whereas darker, desaturated colors were associated with negative concepts (Bartram et al., 2018; Schloss et al., 2020). For visualizations of categorical data, it is possible to make visualizations overall dark or light to convey affective connotations while maintaining discriminability of data categories. However, for visualizations of continuous data where spatial structure is key, this approach could pose a problem because lightness variation is important for revealing spatial patterns in datasets (Ware, 1988; Rogowitz & Treinish, 2009). We investigated if color scales can be created to have strong affective connotations while preserving lightness variation for visualizations of continuous data. Starting with grayscale maps that evoked similar affective connotations (established in a pilot), we applied 16 color scales to each map: 4 hues (red/yellow/green/blue) x 2 lightness levels (light/dark) x 2 saturation levels (saturated/unsaturated). The maps were obtained from a global raster dataset of aboveground biomass stock (Harris et al., 2021) and were selected to have a roughly uniform distribution of pixel values (normalized to scale from 0 to 1). The color scales were generated using Color Crafter (Smart et al., 2020). Participants rated associations between each colormap and eight affective concepts. Positive concepts (positive/exciting/playful) were associated with lighter, saturated, and bluer colormaps, whereas negative terms (negative/disturbing/serious) were associated with darker, unsaturated, and yellower colormaps. The results indicate colormaps can have strong affective connotations while preserving lightness variation important for observing spatial patterns in data.

NSF grant BCS-1945303 to KBS

26.447 ON THE ROLE OF COLOR IN THE AESTHETIC APPRECIATION OF PAINTINGS

Ombretta Strafforello[1], Johan Wagemans[1]; [1]KU Leuven, Department of Brain & Cognition

Why do we like some paintings more than others? Empirical aesthetics studies investigated factors like content, style, composition, harmony, order and complexity. The role of color, however, has received much less attention, although it is often discussed in art history and theoretical aesthetics. Here we explored the role of color in aesthetic perception through a data-driven approach. We computed statistics on the pixel distribution and on the dominant color palette of paintings, extracted via Modified Median Cut Quantization. We analyzed the average pixel brightness and saturation and the amount of color variation per painting, calculated as the average distance among the palette colors in RGB space. We used three datasets — LAPIS, JenAesthetics and BAID — comprising 11,790, 1,628 and 60,337 paintings labelled with aesthetic scores. LAPIS and JenAesthetics contain art style and subject matter annotations, enabling analysis across different painting categories. Our results show that top-rated paintings contain significantly more color variations than bottom-rated paintings in all datasets, indicating that the color diversity of a painting contributes to its liking. This partly relates to style: abstract paintings are rated lower than figurative paintings and also contain less color variations, as in the case of color field and minimalist art. We found a weak association (Pearson correlation) between the amount of color variations and liking in LAPIS' abstract paintings (0.266). Similar associations were found in JenAesthetics' figurative styles where color is a key feature, like Impressionism (0.308) and Rococo (0.330), in paintings of buildings (0.565), landscapes (0.317) and one-person portraits (0.206). However, no association emerges in paintings where spatial composition or shape dominates, such as portraits with many people and nudes. Finally, the amount of brightness variations correlates more weakly with liking (0.203 in LAPIS, 0.183 in JenAesthetics), suggesting that the effects of color variation cannot be reduced to brightness.

26.448 SELECTIVE ACTIVATION OF INTRINSICALLY PHOTOSENSITIVE RETINAL GANGLION CELLS (IPRGCS) MEDIATES N-BACK TASK PERFORMANCE AS A MEASURE OF WORKING MEMORY

Yuta Suzuki[1], Shigeki Nakauchi[2], Hsin-I Liao[1]; [1]NTT communication science laboratories, [2]Department of Computer Science and Engineering, Toyohashi University of Technology

Intrinsically photosensitive retinal ganglion cells (ipRGCs), which contain melanopsin as a photoreceptor that is maximally sensitive to short-wavelength light (460-480 nm; blue or cyan light), are known to mediate physiological responses such as circadian regulation and the pupillary light reflex. Additionally, previous studies have shown that exposure to blue light improves performance on working memory tasks compared with exposure to amber light. However, it remains unclear whether the improvement in cognitive task performance when comparing blue light with reddish light involves other photoreceptors or blue-specific factors, such as color perception. This study investigates the role of ipRGC activation on working memory performance using a silent substitution method that allows selective manipulation of ipRGC activity while minimizing the influence of LMS cone responses. For each subject (N=26 in total), we designed two metameric lights with different levels of ipRGC activation, both perceived as magenta. These lights were implemented using two projectors and optical filters (e.g., Allen et al., 2018). After completing the color matching protocol, participants performed two blocks of a working memory experiment. Each block consisted of several runs of a 1-, 2-back task under the different light conditions, along with two questionnaires on sleepiness and fatigue ratings. The results showed that the correct rate in the 2-back task was significantly higher under the high-ipRGC light than under the low-ipRGC light. Furthermore, the high-ipRGC light reduced subjective sleepiness and fatigue. Taken together, our results confirm that the effect of ipRGC on working memory performance does not necessarily require perceptual blue. This is consistent with the role of ipRGC in projecting to brain areas involved in working memory tasks. These findings highlight that ipRGC activation can enhance cognitive performance independent of blue light perception, broadening the potential applications of lighting environments to support cognitive function.

Color, Light and Materials: Surfaces and materials

SATURDAY, MAY 17, 2:45 – 6:45 PM, PAVILION

26.449 CHARACTERISING THE STATISTICAL IMAGE PROPERTIES OF MATERIALS IN THE STUFF DATASET

Emily J. A-Izzeddin[1] (ej.aizzeddin@gmail.com), Filipp Schmidt[1,2], Roland W. Fleming[1,2]; [1]Justus Liebig University Giessen, Germany, [2]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt

Critical to our interactions with the external world is our ability to make efficient inferences about the material properties of our environment. Indeed, humans are highly adept at performing such inferences, accommodating our expertise for hundreds of unique material categories. Broadly, the brain is thought to perform such inferences by relying on a subset of features to discriminate between material categories while minimising computational complexity. However, the specific features prioritised by the brain for such processing is subject to ongoing research. We recently released a comprehensive image database of 200 material categories (the STUFF dataset), providing a useful tool for identifying potential features that are prioritised by the brain for material processing and categorisation. However, there has yet to be a computationally-driven analysis of the basic statistical properties of images present in the STUFF dataset, which may contribute to human judgements, and may be beneficial to control for in future studies. We have therefore performed a series of analyses, focusing on image features such as luminance, oriented contrast, and spatial frequency content. We find that material images can be categorised computationally at above-chance levels based on these features alone. In addition, we find such simple features to be sufficient for explaining a significant proportion of human categorisations via observer models and generalised linear modelling. As such, our findings point to systematic low-level statistical material profiles present in the STUFF dataset images, highlighting the benefit of accounting for such features - either by eliminating systematic differences via image processing, or by acknowledging and partitioning the variance accounted for by such features before emphasising the role of more complex visual processing.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB/TRR 135 (project no. 222641018, project C1)

26.450 SPATIO-CHROMATIC CUES IN SHAPE AND MATERIAL PERCEPTION

Zoe R. Goll[1], Filipp Schmidt[1,2], Emily J. A-Izzeddin[1], Celine Aubuchon[1], Fatma Kiliç, Roland W. Fleming[1,2]; [1]Justus Liebig University Giessen, Germany, [2]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt

Every day, we interact with a variety of objects made from different materials and make inferences about their properties based on appearance. These inferences are complex, as the spatial and spectral properties of light reflected from these objects depend on illumination, shape, and material. Here, we focused on the role of colour in the perception of shape and material - and especially the role of spatial gradients of chromatic features - which remains poorly understood. We investigated how colour saturation and value gradients across an object's surface influence our perception of gloss and three-dimensional (3D) shape. We rendered images of 3D objects, each having identical shape and illumination, but differing in their material appearances. We then, in HSV colour space, removed either saturation or value gradients by replacing all values in that dimension with the mean. Participants were presented with four versions of each image simultaneously: the original rendering, versions with no saturation or value gradients, and a grayscale version. For a given image set, participants first rated the glossiness of each object version, followed by how 3D they appeared. Compared to the original, there was a significant reduction in perceived gloss for the versions with either saturation or value gradients removed. In addition, 3D ratings were lower for the version with value gradients removed. These results demonstrate the impact of saturation and value gradients on gloss perception and value gradients on perceived 3D shape. This suggests an important role for such spatio-chromatic cues in shape and material perception.

Supported by DFG, project number 222641018–SFB/TRR 135 TP C1, the ERC Advanced Grant "STUFF" (project number ERC-2022-AdG-101098225); and the Research Cluster "The Adaptive Mind" funded by the Hessian Ministry for Higher Education, Research, Science and the Arts.

26.451 EFFECTS OF LUMINANCE, TEXTURE AND BLUR ON CLASSIFICATION OF IMAGE PATCHES AS CHANGES IN MATERIAL OR ILLUMINATION

Eden E. Sterk[1], Madelyn G. Arena[2], Christopher DiMattina[2]; [1]Florida Southwestern State College, [2]Florida Gulf Coast University

To correctly parse the visual scene, one must detect edges and determine their underlying cause. Previous work demonstrated that image-computable neural networks trained to differentiate natural shadow and occlusion edges exhibited sensitivity to boundary sharpness and texture differences. Although these models exhibited strong correlation with human performance on the same task, we did not directly investigate whether humans actually make use of boundary sharpness and texture cues when classifying edges as shadows or occlusions. Here we directly investigated this using synthetic image patch stimuli formed by quilting together two different natural textures, allowing us to parametrically manipulate boundary sharpness, texture modulation, and luminance modulation. Observers were trained to correctly identify the cause of natural image patches (occlusion, shadow, uniform surface). These same observers then classified 5 sets of synthetic images defined by varying sharpness, luminance, and texture cues. These three cues interacted strongly to determine categorization. For sharp edges, increasing luminance modulation made it less likely the patch would be classified as a texture and more likely it would be classified as an occlusion, whereas for blurred edges, increasing luminance modulation made it more likely the patch would be classified as a shadow. Boundary sharpness had a profound effect, so that in the presence of luminance modulation increasing sharpness decreased the likelihood of classification as a shadow and increased the likelihood of classification as an occlusion. Texture modulation had little effect on categorization, except in the case of a sharp boundary with zero luminance modulation. Results were consistent across all 5 stimulus sets, showing these effects are not due to the idiosyncrasies of the particular texture pairs. Our results demonstrate that human observers make use of the same cues as our previous machine learning models when detecting edges and determining their cause.

C.D. is funded by NIH grant R15-EY032732-01

26.452 SEEING INVISIBLE MATERIAL FROM VISIBLE IMAGE FEATURES

Masataka Sawayama[1]; [1]The University of Tokyo, Japan

Humans are generally thought to perceive only the range of real-world information that their sensory systems can detect. For instance, visual sensory receptors are sensitive only to a limited range of light wavelengths. Similarly, spatial patterns beyond the resolution limits of visual acuity cannot be directly resolved. However, in natural environments, information beyond the sensory resolution, i.e., "invisible" information, does not exist in isolation from detectable sensory signals. Instead, they exhibit statistically co-varying characteristics, and computer graphics techniques, such as microfacet reflection models, intensively utilize these statistical relationships to render realistic materials. This observation suggests that humans might also leverage the relationships between detectable sensory signals and invisible properties to perceive such information in natural environments. This study aims to identify visible image features that co-vary with invisible properties using self-supervised learning and asks to what extent humans can estimate invisible information from these features. First, we trained two encoders with contrastive objectives: one encoder processed sensory signals, while the other took as input "invisible" super-resolution signals defined by spatial frequencies in images. A shared latent representation between the two signal types was obtained using a large-scale image dataset through

self-supervised learning. Using a neural style-transfer method, this shared representation was used to create style-transferred sensory signals that inherit styles of “invisible” signals. Next, we conducted psychophysical experiments to investigate whether humans could infer invisible styles from these transferred sensory signals. In these experiments, participants performed discrimination tasks to detect the target invisible styles in the non-target invisible styles. Results showed that participants correctly detected the target styles transferred onto sensory signals. Furthermore, when participants observed invisible signals with a resolvable distance after the discrimination task, they could detect the previously “unseen” target styles. These findings suggest that humans infer invisible information from sensory signals co-varying invisible properties.

26.453 MATERIAL APPEARANCE AFFECTS OBJECT RECOGNITION

Fatma Kilic[1] (fatma.kilic@psychol.uni-giessen.de), Filipp Schmidt[1,2], Celine Aubuchon[1], Emily A-Izzeddin[1], Zoe R. Goll[1], Roland W. Fleming[1,2]; [1]Justus-Liebig-Universität-Giessen, Germany, [2]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt

It is widely recognized that shape and contours serve as crucial cues for object recognition. Several studies have also highlighted the impact of color on our ability to recognize and classify certain objects, with diagnostic colors. Yet the role of broader material appearance in object recognition remains surprisingly underexplored. Here, we sought to measure the cognitive and perceptual impact of mismatches between an object's shape and its material in object recognition tasks. We rendered images of 3D scans of animals and vegetables that were outfitted with various natural or manmade materials (e.g., cardboard). Participants performed a Go/No-Go task in which they had to respond as rapidly as possible to an image of the specified superordinate category (i.e., Animal in one block, Vegetable in another) while ignoring stimuli of other categories. The results showed that people responded more slowly to vegetables in incongruent materials than those in natural materials, while there was no such effect for animals. The observed difference in identification processes between vegetables and animals may stem from the fact that vegetable shapes were less complex and therefore less diagnostic (e.g., carrot vs. fish). Thus the resulting ambiguity might be resolved through additional information on the material or texture to aid in accurate identification. In contrast, the more distinct and recognizable shapes of animals facilitate their recognition, reducing the reliance on external characteristics for identification purposes.

This research was funded by the European Research Council (ERC) Advanced Award ‘STUFF’ (ERC-ADG-2022- 101098225)

26.454 HUMAN AND MACHINE PERCEPTION OF MATERIAL SIMILARITIES

Maarten Wijntjes[1], Yuguang Zhao[1]; [1]Delft University of Technology

Large Multimodal Modals can be subjected to similar psychophysical paradigms as human observers, affording comparison between human and machine vision. In this context, we explored material perception. We created 32 stimuli of a constant 3D shape but with various material properties. Then we presented them in 1193 triplets in an odd-one-out task for both humans (N=18) and machine. The machine judgements were performed with gpt-4o, which has vision capabilities. Triplet data was both analysed directly, and also used to create perceptual embeddings using Soft Ordinal Embedding (SOE). The raw triplet data revealed an interesting commonality between human and machine judgements when we compared the ‘popularity scores’ of odd-ones-out: a group of 6 stimuli was substantially more different from the remaining 26 stimuli. Furthermore, we found that 47% of the triplet judgements were similar for the human and gpt-4o data, which is well above chance level (33%). The SOE analysis revealed that the accuracy (agreement between raw triplet data and multidimensional embeddings) was substantially higher for machine than human vision, indicating a higher degree of internal consistency. Also, we found a full saturation at 6 dimensions for the machine data: all triplets could be accounted for by the embedding. Besides various commonalities, the embeddings themselves revealed some peculiar differences. Firstly, translucent stimuli were close for humans but distant for the machine. Secondly, the machine embedding showed a clear cluster of achromatic stimuli, while this was entirely absent in the human data. This suggests that computers use colour for material perception, while humans do not. With some imagination, one could argue that human material perception partly prepares for physical interaction where colour is irrelevant, while the algorithm does not (yet) have a body to interact with the outside world.

26.455 MATERIAL IMAGE MORPH AND BINOCULAR INTEGRATION

Hua-Chun Sun[1,2] (hua-chun.sun@psychol.uni-giessen.de), Roland Fleming[1,2]; [1]Justus Liebig University Giessen, Germany, [2]Center for Mind, Brain and Behavior (CMBB), University of Marburg, Justus Liebig University Giessen and TU Darmstadt, Germany

Material perception involves both recognising materials, and determining their physical properties, even if they are unfamiliar to us. It is widely believed that material recognition involves representing materials (e.g., wood vs metal) in multidimensional feature spaces which cluster samples of each familiar material within a tight neighbourhood, facilitating recognition and discrimination. Yet a key open question is how we perceive the properties of materials that project to intermediate locations within such spaces (e.g., between wood and metal). Here, we employed deep learning-based image interpolation plus interocular fusion to probe material space, and measure how the brain integrates material information from the two eyes. We selected 20 images from the STUFF dataset (Schmidt, Hebart, Schmid & Fleming, 2023) to generate 10 cross-category morph pairs (e.g., sand-grass, moss-hair). Observers were presented with two types of stimuli: a reference stimulus, in which each eye viewed a different weighted combination of the morph pair (e.g., 30% sand + 70% grass in the left eye and 70% sand + 30% grass in the right eye), and a match stimulus, presented identically to both eyes. Participants adjusted the match stimulus along 49 morph steps (ranging from 2% to 98%) between the original images to achieve perceptual equality with the reference stimulus. Trials with unfusable or rivalry perception are removed from further analysis. Intriguingly, our results show that perceptually distinct materials presented to each eye can be integrated into a coherent novel material percept. The adjusted morph weight for the match stimuli was approximately the midpoint of the two images, with greater variance observed in material pairs with larger interocular weight differences. These findings suggest an interocular summation mechanism in the brain for combining material information from the two eyes.

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).

26.456 VISUAL ADAPTATION OF COMPLEX MATERIAL APPEARANCES

Chenxi Liao[1] (cl6070a@american.edu), Jacob Cheeseman[2], Filipp Schmidt[2], Roland W. Fleming[2], Bei Xiao[1]; [1]American University, [2]Justus Liebig University Giessen

Material perception is challenging to study because many intrinsic and external factors affect material appearances. How the visual system encodes complex materials (e.g., translucent materials) remains unknown. Here, we explore the representation of materials using visual adaptation—perceptual distortion after exposure to stimuli for a prolonged period. In each trial, participants were presented with a pair of images (spatially jittering to avoid retinal afterimages) of different materials to the left and right of the fixation (adaptor pair). After a total of 4 seconds, they were followed by a brief presentation (0.2 seconds) of a pair of test images (test pair). Tests were either identical or different materials sampled along a morphed sequence between the two adaptors. Participants judged whether the two test images were the same or different. Baseline judgments were collected without prior exposure to the adaptors. All stimuli were created by training generative neural networks on custom-made photographs of everyday translucent objects, like soaps and stones. We selected nine distinctive adaptor pairs with different perceptual similarity in material category, color, shape, and other material properties (e.g., translucency, smoothness). We morphed between adaptors (e.g., a soap and a stone) to sample sequences of fifteen test pairs equally around the midpoint of each morph sequence, spanning a range of variation in those perceptual dimensions. In comparison with the baseline, we found that adaptation enhanced sensitivity to subtle differences in similar stimuli, leading to more "different" judgments. Meanwhile, it reduced the discrimination for distinct stimuli, resulting in "same" judgments. Notably, the magnitude of the repulsive effect depended on the specific adaptor pair. Our findings show that adaptation can lead to powerful distortions of perceived material appearance, affecting multiple perceptual dimensions, and even changing the apparent category of materials.

Supported by NIH, award 1R15EY033512-01A1; DFG, project number 222641018–SFB/TRR 135 TP C1, the ERC Advanced Grant "STUFF" (project number ERC-2022-AdG-101098225); and the Research Cluster "The Adaptive Mind" funded by the Hessian Ministry for Higher Education, Research, Science and the Arts.

26.457 PERCEIVED ANIMACY FROM GLOBAL AND LOCAL IMAGE DISTORTIONS

Emre Türkmen[1] (turkmen.emre@metu.edu.tr), Görkem Baysal[1], Dicle Dövencioğlu[1]; [1]Middle East Technical University (METU), Ankara, Turkey

Visual shape cues to animacy are discussed along global and midlevel classes such as symmetry, head/limbs, and curvilinearity/rectilinearity. Local shape distortions are also shown to create rich percepts of transparent layers (e.g., a rainy windshield) but have not yet been tested for animacy. Here, we manipulated the contours of 2D shapes globally and locally to investigate perceived animacy. In Experiment 1, 11 animate-inanimate (e.g., bird-airplane) pairs were morphed into each other in seven steps. Participants were shown each shape (150 milliseconds, 20 repetitions) and asked to respond as "animate" or "inanimate." Psychometric curves fitted to the percentage of "animate" responses generated discrimination thresholds at 19.6-60.6% morph levels. Although animate shapes were significantly more curvilinear, we did not observe a systematic relationship between curvilinearity and the thresholds. In Experiment 2, we distorted the local disarray of 20 images to create their eidolons (equivalent appearance classes). We changed the reach-grain parameters of ten inanimate stimuli (Experiment 1) and ten insect-like stimuli with varying curvature (5) and symmetry (2) levels in their limbs. Participants were asked to adjust the reach-grain parameters for three modes: to make the stimulus look "underwater," "behind glass," or "animate." Group means for the reach-grain values for underwater and behind-glass conditions were compatible with previous findings, and reach values for animacy yielded a bimodal distribution. A repeated measures ANOVA revealed a main effect of mode on both reach and grain values, and reach (but not grain) values systematically changed with the curvature: As curvature increased, participants adjusted lower settings to reach parameters. Overall, this study provides further evidence for the perception of rich transparent layers by local distortions in the image. We also show that parametric manipulations in image distortions explain perceived animacy, and both global (curvilinearity) and local (reach) shape cues are important in discriminating animate from inanimate.

Research Universities Support Program Cooperation Protocol (ADEP Grant no 104-2023-11273)

Motion: Illusions

Saturday, May 17, 2:45 – 6:45 pm, Pavilion

26.458 DIVERSITY IN MOTION ADAPTATION: A CLINICAL PERSPECTIVE

Renell Rodrigues[1], Guang Yang, David Yu, Martha Munro, Adam Guastella, Frans Verstraten; [1]The University of Sydney

Do individuals with autism (ASD) tend to see the trees but not the forest? Research findings in local and global perceptual processing have been inconsistent. These atypical perceptual patterns become evident in experiments investigating motion detection using motion coherence thresholds. For example, if 10% of randomly positioned dots move coherently in one direction, neurotypical individuals indicate the global direction with ease, whereas participants with ASD find it challenging to indicate the global direction. One assumption is that the local dots interfere with the visibility of the moving dots. To investigate this, we used a motion adaptation paradigm with participants from a social research group who all reach clinical criteria for autism spectrum

disorder (DSM-5). Participants fixated on a dot on a screen, while being presented with a moving random pixel array (4 cardinal directions). After 30 seconds of adaptation to the motion stimulus, participants indicated the perceived aftereffect duration by pushing the space bar. The test pattern after adaptation was either a stationary noise pattern (a stationary random pixel array) or a dynamic noise pattern (the same random pixel array where the dots were continuously refreshed). Our hypothesis was that the dynamic test pattern will interfere with the ability to perceive the aftereffect. It was predicted that the aftereffect duration would be short or even absent for the dynamic test pattern. The results indicate that ASD participants are indeed unable to visually perceive the dynamic motion aftereffect. The aftereffect on a static pattern was successfully perceived. The preliminary conclusion is that the local noise inhibits the perception of the global motion signal and that there is potential to use this as a clinical assessment tool.

26.459 INCREASING WITHIN-HEMIFIELD MATCHING WITH ECCENTRICITY IN QUARTETS APPARENT MOTION

Eunhye Choe[1] (eunhye.choe.gr@dartmouth.edu), Patrick Cavanagh[1,2], Peter U. Tse[1]; [1]Dartmouth College, [2]Glendon College

In apparent motion quartets, dots at two diagonally opposite corners of an imaginary rectangle are replaced by two dots at the other two corners, causing subjective perception to shift between horizontal and vertical motion percepts. When the imaginary rectangle is a square (i.e. aspect ratio or AR = 1), there is typically a bias toward perceiving vertical more often than horizontal apparent motion. This may arise from a tendency to match successive dots within a hemifield, as such matching occurs within rather than between brain hemispheres. This bias can be counteracted by increasing the within-hemifield matching distance relative to the cross-hemifield distance (i.e. AR > 1). The within-hemifield matching bias is expected to be greater in peripheral vision, where hemisphere-specific processing occurs, than in foveal vision which is processed to some extent in both hemispheres. In this study, apparent motion quartets were presented on a virtual circle with radii of 0.25, 0.5, 1, 2, 4, or 8° of visual angle relative to fixation. Each quartet consisted of two diagonal dot pairs, with aspect ratios gradually shifting from 3 to 1/3 or the reverse. Participants reported the aspect ratio at which they perceived a change in motion direction. The results show that the threshold aspect ratio of horizontal to vertical distances increased significantly with greater eccentricity (0.25°, AR=1.12; 0.5°, AR=1.14; 1°, AR=1.16; 2°, AR=1.19; 4°, AR=1.22; 8°, AR=1.21). Our findings suggest that the bias to match within-hemifield becomes more pronounced with increasing eccentricity.

26.460 INVESTIGATING ATTENTIONAL REPULSION AS A MECHANISM FOR ANISOTROPIC POSITION SHIFTS AROUND MOVING OBJECTS

Mohammad Shams[1,4] (m.shamsahmar@gmail.com), Aurore Maloh[2], Peter Kohler[3,4], Patrick Cavanagh[1,4]; [1]Glendon College, York University, Toronto, ON, Canada, [2]Faculty of Health, York University, Toronto, ON, Canada, [3]York University, Toronto, ON, Canada, [4]Centre for Vision Research, York University, Toronto, ON, Canada

Moving objects shift the perceived position of nearby flashes (Whitney & Cavanagh, 2000). We presented bars rotating around the center of gaze and flashed dots at different locations and times relative to the bars. We found that flashes presented ahead of a moving bar were shifted in the direction of the bar's motion, peaking for probes 15 degrees of rotation ahead of the bar, while flashes presented behind the bar remained largely unaffected (similar to Watanabe et al., 2003; Durant & Johnston, 2004). It has previously been proposed (Shams, Kohler, & Cavanagh, ECVP 2023) that this anisotropic position shift is due to attentional repulsion (Suzuki & Cavanagh, 1997) where, in this case, the focus of attention leads the moving object. Here, we further tested the role of attention by varying the number of bars and flashes and by introducing a spatial cue at the center pointing to the physical location of the target, presented either before (pre-cue) or after (post-cue) the bar's motion. In the post-cue condition, we observed a large position shift (~1.7 dva) that did not differ between the one-bar and four-bar conditions. In contrast, in the pre-cue condition, the position shift decreased (~7%) in the one-bar condition and increased (~12%) in the four-bar condition compared to the post-cue conditions. The lack of effect of the number of probes in the post-cue conditions suggest that exogenous attention must be involved because exogenous attention can be unaffected by attentional load (Wright, 1994; Solomon, 2004). The pre-cue conditions suggest that endogenous attention may also play a role by either modulating the position shift by reallocating resources between the moving object and the probe location (as proposed by Adamian & Cavanagh, 2024) or by increasing the speed at which the probe location can be retrieved (Müsseler & Aschersleben, 1998).

26.461 MOTION INDUCED POSITIONAL SHIFT IS MODULATED BY ATTENTIONAL LOAD

Minwoo J.B. Kim[1] (cogminu@gmail.com), Oh-Sang Kwon[1]; [1]Ulsan National Institute of Science and Technology

The motion-induced position shift (MIPS) is an illusion where a stationary object in the periphery appears displaced due to local motion. According to a Bayesian model (Kwon, Tadin, & Knill, 2015), positional uncertainty causes the perceptual system to rely on local motion velocity to infer object positions, with attention potentially playing an important role in perceptual accuracy. Recently, Nakayama & Holcombe (2020) found that increasing bottom-up attention can reset illusory object positions and motion trajectories, highlighting the role of attention. This study tested whether increased attentional load reduces the illusory shifts of object position. Noise patches with Gaussian kernels were presented in the periphery (10–12° eccentricity) with local motion (9–11°/s) for 1 second. The number of objects presented simultaneously (1, 2, or 3) varied across blocks to manipulate attentional load. Participants fixated on the screen center and reported the final object position, and the speed and direction of the local motion. We mainly analyzed the first responded object within a trial to isolate effects of attention from working memory. Results showed that MIPS magnitude increased with the number of objects, indicating a trade-off between attentional availability and illusory positional shifts. Divided attention to multiple objects likely impaired the precision of positional judgments, forcing the perceptual system to rely more on motion cues. These findings align with theories suggesting attention causes a reset of illusory shift of object position.

This work was supported by National Research Foundation of Korea: [Grant Number NRF-2020S1A3A2A02097375]

26.462 PRIMING WITH FURROW ILLUSION DEMONSTRATES CONTRIBUTIONS OF CONSCIOUS MOTION PERCEPTION TO VISUALLY GUIDED ACTION

Marjan Persuh[1], Alyssa Costigan; [1]Department of Social and Behavioral Sciences, Manhattan University

One prominent proposal suggests a separation between the vision for action and vision for perception. Actions, however, depend on knowing and representing veridical object properties. The idea that veridical (as opposed to perceptual) information guides motor activity has been independently developed for the phenomenon of response priming. It has been proposed that motor responses in priming are based on the rapid feedforward stream of information processing, which extracts only basic, physical properties of the prime stimulus. We tested this hypothesis using a Furrow illusion, in which a veridical, vertical dot movement in the periphery is perceived as tilted when presented against a tilted background grating. In the first experiment, we asked participants to make speeded responses to the target consisting of a dot moving across a tilted path. Preceding the target, prime was presented in the periphery, which consisted of a dot moving vertically, but due to the Furrow illusion was perceived as moving across a tilted path. We asked whether participants' responses to targets would be affected by the consciously perceived movement along the tilted path or stay unaffected by the veridical, vertical motion of the primes. Contrary to theoretical expectations, we observed significant effects of primes on rection times to targets, demonstrating that motor responses were affected by the perceived and not veridical direction of motion. In the second experiment we first estimated the size of the illusion for each participant individually and then used that estimate to eliminate the illusion. With this manipulation, veridical motion was slanted, whereas perceived motion was vertical. Surprisingly, reaction times in this experiment were not affected by the veridical motion. In summary, our two experiments demonstrate that our visuo-motor responses use consciously perceived information about objects to guide motor responses.

26.463 THE MOTION AFTEREFFECT IN VISUAL SNOW

Samantha Montoya[1] (monto112@umn.edu), Anna Hillstrom[1], Karly Allison[1], Carter Mulder[1], Mike Lee[1], Michael-Paul Schallmo[1], Stephen Engel[1]; [1]University of Minnesota

Visual snow (VS) is the perception of tiny flickering dots covering the visual field. VS is estimated to affect ~4% of the population and, when combined with other symptoms, can be disabling. VS likely arises from spontaneous (non-stimulus driven) activity in the visual pathways, but the specific regions involved are unknown. We sought to determine whether the neural activity producing VS arises at or before motion selective areas. The motion aftereffect (opposite-direction illusory motion perceived after adapting to motion) results from neural adaptation in motion selective regions including area V5/MT. If the activity generating VS reaches these neurons, the VS symptom should be susceptible to the motion aftereffect. Eleven participants with VS viewed high-contrast, horizontally drifting gratings on the left and right of a central fixation point. The gratings (0.3 cyc / degree, 2 Hz) drifted towards or away from the center, which should produce opposing aftereffects. Adapter duration varied across trials (1.6, 5, 15, or 45 sec, each repeated on 4 trials). The gratings were replaced with blank circles in which participants judged the motion of their VS. Participants pressed a button when the motion of the snow inside the circles matched, indicating the duration of the effect. To ensure participants with VS experienced a typical motion aftereffect for external stimuli, stationary square-waves were shown in a control condition. Most participants (10/11) reported that after adaptation, their VS moved in the opposite direction of the adapting gratings. Longer adapter durations resulted in longer-lasting effects (ANOVA $F_{10,1} = 74.4$, $p = 6.1 \times 10\text{-}6$) following a power law, consistent with prior motion aftereffect literature. VS was susceptible to the motion aftereffect, indicating the neural activity responsible for snow reaches motion selective neurons, and therefore arises early in the visual pathways, at or before area V5/MT.

Supported by a UMN Grant-in-Aid (574483); grants from NEI, NIH (F31 EY034016, T32 EY025187, UL1TR0024940); NINDS (R25 NS117356); and the NSF (DGE 1734815 at the University of Minnesota).

26.464 THE ROLE OF LUMINANCE CONTRAST IN MOTION SILENCING ILLUSION

Hyerim Kim[1], Oakyoon Cha[1]; [1]Sungshin Women's University

Motion silencing refers to a visual illusion in which changes in visual features of small dots, such as color or luminance, become less noticeable when the dots move coherently as a group. Previous studies have typically presented dots changing color or luminance against a uniform gray background, which may have inadvertently caused switches in contrast polarity throughout the sequence, potentially disrupting the ability of simple cells in the primary visual cortex (V1) to continuously track these changes. Since simple cells are more sensitive to luminance contrast than complex cells, we investigated whether enhancing feature changes to align better with simple cells' sensitivity could increase the visibility of the changes, thereby weakening the motion silencing illusion. Participants viewed two groups of 200 dots each in the left and right visual fields. In the comparison field, the dots did not move, while in the distractor field, the dots moved upward or downward at random speeds. All dots cycled in color between bright yellow and dark blue. The rate of color change was fixed at 1 Hz in the distractor field, whereas it was adjusted in the comparison field using a 1-up, 1-down adaptive staircase procedure. This procedure allowed us to estimate the point of subjective equality for each background condition (white, mid-level gray, and black), tested in separate blocks. The dots' contrast was, on average, the least salient against the mid-level gray background, because its luminance was between that of bright yellow and dark blue. In addition, the gray background likely disrupted V1 simple cells' ability to track dots' luminance changes, resulting in stronger motion silencing. Nonetheless, Bayesian ANOVA revealed no differences across background conditions. These findings suggest that motion silencing illusion persists across varying background conditions, highlighting its robustness even when differences in luminance contrast are introduced.

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).

26.465 STROBOSCOPIC HALLUCINATION SPATIAL FREQUENCY CORRESPONDS TO STROBE STIMULATION TEMPORAL FREQUENCY

Nathan H. Heller[1] (nheller4@jh.edu); [1]Johns Hopkins Medicine, Center for Psychedelic and Consciousness Research

Stroboscopic (strobe) hallucinations are dynamic, kaleidoscopic patterns induced by flickering light in front of a person's closed eyes. They resemble geometric hallucinations reported during various clinical and pharmacological contexts (Billock & Tsou, 2012; Heller et al., 2023). Computational models attribute these percepts to waves of cortical excitation in V1 (Bressloff et al., 2001) and indicate that higher temporal frequency (TF) strobe corresponds to lower spatial frequency (SF) patterns (Rule, Stoffregen, & Ermentrout, 2011). However, anecdotal reports suggest the opposite relationship between TF and SF exists (Smythies, 1960). Here, we present the first attempt to quantify this relationship. First, participants reported which of 4 strobe frequency pairs (5&10Hz, 10&15Hz, 15&20Hz, and 20&25Hz) induced higher spatial frequency patterns, along with their confidence level (1-10). Then, by adjusting the spatial frequency of a checkerboard, participants matched the size of the checkerboard elements (dva) to the elements comprising strobe hallucinations for foveal (central 4 dva) and peripheral regions. Consistent with anecdotal reports, pairwise judgements revealed strobe TF and hallucination SF correspond positively. Confidence was moderate for the 20&25Hz pair ($M=6.5$, $SEM=1.4$) and otherwise high ($M=9.0$, $SEM=0.65$). Checkerboard adjustment showed that the size of hallucination elements (dva) decreased steadily until the two highest strobe TFs (5Hz: $M= 4.19$, $SEM=1.28$; 10Hz: $M=1.14$, $SEM=0.32$; 15Hz: $M=0.62$, $SEM=0.16$; 20Hz: $M=0.33$, $SEM=0.09$; 25Hz: 0.33, $SEM=0.11$). A repeated-measures ANOVA was conducted to assess the effects of visual region (foveal vs. peripheral) and strobe-stimulation frequency (5-25 Hz) on hallucination element size. There was a significant main effect of frequency ($F[4,12]=15.58$, $p<.001$, $\eta p2 =0.839$), corroborating the strong positive relationship between strobe TF and hallucination SF. The effect of visual region was marginal ($F[1,3]=9.64$, $p=.053$, $\eta p2 =0.763$), though there was a significant interaction between frequency and region ($F[4,12]=4.88$, $p=.014$, $\eta p2 =0.619$). Computational models of strobe hallucination should be amended to reflect these results.

Motion: Biological, self-motion

Saturday, May 17, 2:45 – 6:45 pm, Pavilion

26.466 NEURAL MECHANISMS OF SOCIAL CONTINGENCY PERCEPTION IN COMMUNICATIVE INTERACTIONS: AN FMRI STUDY

Ting Zhang[1], Rui Wang[1,2], Yi Jiang[1,2]; [1]Institute of Psychology, Chinese Academy of Sciences, [2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

Understanding social interactions is crucial for navigating human behavior in daily life, requiring the ability to recognize, interpret, and predict the actions of others. However, the brain mechanisms underlying the perception of social contingency in communicative interactions remain unclear. Here we combined behavioral and fMRI measurements to investigate the neural representation of social interaction by anticipating contingent actions. Participants viewed the point-light displays of two successively presented agents and performed a discrimination task (interactive or independent actions). This paradigm allowed us to test how the biological motion perception is shaped by the context of social interaction by quantifying neural activations in response to identical actions of the second agent under different social contexts, ruling out potential confounding factors such as low-level features or attentional orienting. Our results showed that people can identify communicative interactions from discrete body movements based on social contingency. fMRI analysis revealed decreased neural responses to interactive (compared to independent) motions in the right posterior superior temporal sulcus (pSTS), and the mentalizing network including the right temporoparietal junction (TPJ), bilateral precuneus and right superior frontal gyrus. However, such effects were absent in the low-level motion-selective middle temporal cortex (MT+) and body-sensitive extrastriate body area (EBA). Notably, correlation analysis showed that such social-context modulation of neural response in the pSTS was significantly associated with individual sensitivity to discriminating social interactions. Furthermore, psychophysiological interaction analysis showed increased functional connectivity of the right pSTS with the right medial prefrontal cortex, bilateral frontal gyrus, TPJ, and precuneus for the independent compared to interactive context. Taken together, our findings highlight that the pSTS serves as a hub for predicting social interaction, presumably involved in the integration of action perception and mentalization.

Acknowledgements: This research was supported by grants from the STI2030-Major Projects (2021ZD0203800), the National Natural Science Foundation of China (32430043), the China Postdoctoral Science Foundation (No. 2023T160678).

26.467 THE INFLUENCE OF REALISTIC OPTIC FLOW AND ECOLOGICAL SELF-MOTION STATISTICS ON OPTIC FLOW TUNING IN DEEP NEURAL NETWORKS

Alexander Lyon[1] (atlyon26@colby.edu), Oliver Layton[1]; [1]Colby College

Optic flow provides rich information about the self-motion of an observer and the layout of the surrounding environment. Over the past several decades psychophysical studies have characterized the accuracy with which humans perceive their self-motion from optic flow and neurophysiological studies have linked self-motion perception to primate brain area MSTd. Despite this progress, most studies rely on optic flow generated from minimal dot environments that differs substantially from the rich and complex patterns encountered during real-world self-motion (Matthis et al., 2022). Moreover, naturalistic self-motion is not uniformly distributed; some types of motion are more common than others. For instance, studies on optic flow experienced by infants and their mothers reveal distinct asymmetries, such as higher rates of expansion compared to contraction (Raudies et al., 2012; Gilmore et al., 2015). In the present work, we examined the extent to which ecological optic flow and biases in self-motion statistics may shape the neural mechanisms underlying self-motion perception. Specifically, we explored how well deep artifical neural networks (DNNs) capture optic flow tuning in MSTd when trained on optic flow generated from minimal dot or realistic virtual environments. Using the 3D game engine Unity, we created large-scale video datasets of simulated self-motion through realistic cluttered warehouse and outdoor scenes with ground-truth labels. We found that training DNNs on optic flow from realistic environments improves their alignment with MSTd optic flow tuning properties. Similarly, incorporating the nonuniform self-motion statistics from Raudies et al. (2012) enhanced the consistency between the DNNs and MSTd, for example with respect to translation tuning preferences. Our work takes a step toward characterizing how ecological self-motion statistics and optic flow generated from realistic environments may shape the neural mechanisms that underlie self-motion perception.

The Colby College Provost's Office and Department of Computer Science

26.468 EYE MOVEMENTS AFFECT HEADING PERCEPTION FROM OPTIC FLOW

Matthew Anderson[1] (matt.anderson@berkeley.edu), Emily Cooper[1], Jorge Otero-Millan[1]; [1]University of California, Berkeley

Self motion creates a global pattern of retinal motion signals called optic flow. The direction of self motion (i.e., heading) influences the focus of expansion of the flow, which is the retinal location from which motion vectors radiate. Previous work has reliably shown that humans use the focus of expansion as a cue for estimating heading. However, as we move through the visual world our eyes are continuously moving in order to monitor the environment (e.g., for hazards and navigational landmarks). These eye movements modulate the pattern of optic flow on the retina, and therefore the location of the focus of expansion. Although previous work has shown that heading perception is robust to some eye movements, the range of interactions between eye movements and heading perception are poorly understood. We manipulated gaze variables such as fixation position and smooth pursuit speed/direction to examine how they affect heading perception. Human observers viewed optic flow stimuli that simulated locomotion across a flat ground-plane. Meanwhile, observers pursued a dot that varied in position, speed, and motion direction relative to the optic flow's focus of expansion. We tracked observers' gaze with a video-based eye-tracker. After each stimulus was presented (for 1.5 secs), observers judged the heading direction. We found that heading judgements were strongly affected by fixation position, but weakly affected by the speed/direction of the fixated dot. As horizontal (azimuth) distance of fixation from the focus of expansion increased, heading accuracy significantly decreased, yet accuracy did not significantly vary with vertical (elevation) distance. We model these behavioral results with a Bayesian ideal observer that incorporates priors in heading perception and noise in both retinal cues (visual motion sensitivity) and extraretinal cues (efference copies of eye position and eye velocity).

This work was partially supported by Alcon.

26.469 EFFECT OF GAIN ADAPTATION ON POSTURAL SWAY

Xue Teng[1,2], Laurie M. Wilcox[1,3], Robert S. Allison[1,2]; [1]Centre for Vision Research, York University, [2]Department of Electrical Engineering and Computer Science, York University, [3]Department of Psychology, York University

In virtual reality (VR) environments, our movements often differ from those in the equivalent physical world. One common discrepancy is visual gain which is a scaling difference between visual and kinesthetic motion. Previous studies have shown that observers reliably perceived the point of subjective stationarity (PSS) in a gain discrimination task during active self-motion. However, the PSS did not shift following prolonged adaptation to non-unity gain. Here we asked whether postural response adapts to gain manipulations. Three adaptation gains were tested in separate, counterbalanced blocks. Each block consisted of a 10-minute initial adaptation, followed by four test segments interleaved with three 2-minute top-up adaptation periods. During adaptation observers were immersed in a virtual room and continuously walked to grab objects and align them with corresponding markers at other locations. Their virtual motion was scaled by 0.67, 1 or 2 times their physical motion. During testing, quiet stance postural sway was recorded while the surrounding environment oscillated sinusoidally at 0.2 Hz over a peak-to-peak distance of 0.5 m, in either the front-back (N=18) or the left-right (N=14) direction. To isolate the visual perturbation effects, we also included a 'stimulus absent' condition in which observers viewed a dark HMD screen. Prior to starting, 60 seconds of baseline postural sway data was collected for both visual stimulus present and absent conditions. Results showed that postural sway in dark was larger after adaptation to a gain of 2 (in either motion direction). Further, power analysis at 0.2 Hz suggested that visually-elicited synchronous postural sway was larger under both non-unity gains along the front-back direction, suggesting that gain manipulations resulted in destabilization. Collectively these experiments suggest that gain manipulations produce adaptation in postural responses, while perceived stability does not shift. This dissociation suggests that postural recalibration to gain adaptation operates independently of perceptual mechanisms.

The authors wish to thank NSERC Canada for support under an Alliance Grant #ALLRP 570802-21, in partnership with Qualcomm Canada.

26.470 GOOD INTENTIONS: OBSERVERS PREFER VIEWING MOVING SHAPES WHICH LOOK GOAL-DIRECTED

Hong B Nguyen[1] (nguyh376@newschool.edu), Benjamin van Buren[1]; [1]The New School

We often see others' behaviors as reflecting underlying mental states, such as beliefs and desires. Perceiving behavior as goal-directed may be computationally efficient: For example, when predicting whether your friend will walk to a store to buy eggs, it is simpler to do so in terms of their beliefs and desires than in terms of the complex physical forces causing their movements. Or, when predicting the orientation of a moving dart, it is simpler to predict that it will continuously turn in order to face another moving shape than to extrapolate its current angular acceleration. Computational efficiency provides a good in-principle argument for why we should see others' behavior as goal-directed—but this account has rarely (if ever) been tested. To investigate this, we measured whether displays featuring goal-directed movement produce a well-established signature of visual processing fluency—positive *hedonic* responses. In 10 experiments, observers viewed a moving shape (a dart, or an ellipse) which updated its orientation to face or remain aligned with a moving disc, producing a strong impression of intentionality. They also viewed closely-matched inanimate control displays (e.g. in which the shape moved identically, but updated its orientation to face away from the disc, or remain offset by 30°). Observers consistently preferred displays featuring goal-directed motion—and this result was mirrored in 10 further experiments measuring implicit associations. These experiments rule out low-level explanations (such as symmetry), and show that the preference to see goal-directed movement is not just a preference for more animate-looking displays, as goal-directed shapes were also preferred to those which looked animate but not strongly goal-directed (e.g. randomly moving darts which faced where they were heading, producing an impression of 'aimless' movement). These results suggest that seeing simple shapes in goal-directed terms allows us to efficiently predict their behavior, which in turn drives positive affective responses.

Temporal Processing: Duration, timing perception

Saturday, May 17, 2:45 – 6:45 pm, Pavilion

26.471 DEFINING A FUNCTIONAL HIERARCHY OF MILLISECOND TIME: FROM STIMULUS PROCESSING TO DURATION PERCEPTION

Gianfranco Fortunato[1] (gfortuna@sissa.it), Valeria Centanino[1], Domenica Bueti[1]; [1]International School for Advanced Studies (SISSA)

In the human brain, millisecond durations of visual events are processed across different regions, from early visual to parietal and frontal areas. Previous functional magnetic resonance imaging (fMRI) studies have shown that early visual areas respond to durations via monotonic tuning, while downstream areas employ unimodal tuning. However, whether and how the properties of unimodal tuning change along the hierarchy to mediate duration perception is still unknown. In this study, we recorded brain activity at 7T-fMRI while participants performed a duration discrimination task of visual stimuli. We modeled BOLD responses using the population receptive field approach for estimating vertex-wise tuning functions. Results showed that neuronal populations in occipital visual areas maximally respond to long durations, in line with the presence of monotonic tuning mediating duration encoding. In parietal, premotor and caudal supplementary motor areas, neuronal populations show selectivity to different durations, covering the entire presented range, and they are also spatially organized in maps. Both properties are compatible with the read-out of temporal information. In rostral supplementary motor areas, inferior frontal cortex and anterior insula, neuronal populations selectively respond to the mean of the presented durations. This selectivity also correlates with participants' point of subjective equality, suggesting a subjective representation of temporal information used to solve the task. Finally, our results showed the existence of specific dependencies between these neuronal populations, likely supporting the presence of distinct functional stages in duration processing and perception i.e., encoding, representation and task-related decisions. Overall, our findings suggest the existence of a putative functional hierarchy of visual temporal processing and highlight the role of tuning properties as neural mechanism of duration perception.

26.472 CRITICAL FUSION FREQUENCY DEPENDENCE ON STIMULI ORIENTATION IN THE PERCEPTION OF DUTY CYCLE VARIABLE FLASHING DOTS

Kotaro Oikawa[1], Ruggero Micheletto[1]; [1]Yokohama City University

Nowadays, video communication technology is widely used. Although many technological developments have improved video communication, the mechanism of perceptual effects such as flicker has not yet been clarified, and this remains a challenge for achieving better video quality. We focused on the curve of critical fusion frequency and duty ratio (C.F.F. - DR curve), which is the threshold of flicker generation by repeated flashing stimuli. In our previous work, we reported that the symmetry of the curve changed between single LED and multiple LED (3x3 grid) stimuli. In this study, we measured the C.F.F. - DR curve by decomposing the multiple-LED stimuli into vertical and horizontal stimuli. Specifically, an analog square wave generator was used to present flashing stimuli with different duty ratios. The vertical and horizontal stimuli were masked from each other using covers to align the light intensity. Subjects then adjusted the frequency of each flashing stimulus and recorded the C.F.F. value. The distance between stimulus and subject was 1.5 m and was done in a dark room. Asymmetric values were calculated by adding up the normalized vertical and horizontal C.F.F. - DR curves of the differences with a duty ratio of 50% as the boundary. The resulting asymmetry value for the vertical stimulus was approximately 25% higher than that for the horizontal stimulus on average for the subjects. It is suggested that this reflects the orientation between vertical and horizontal in the visual system. Furthermore, the experimental results are consistent with the parameters of the spatial filtering process in our previously reported model.

26.473 THE EFFECT OF ATTENTION ON SPEED AND TIME PERCEPTION

Garry Kong[1] (gkong@hiroshima-u.ac.jp), Akira Sarodo[2], Katsumi Watanabe[2]; [1]Hiroshima University, [2]Waseda University

We adapted a working memory paradigm to investigate how attention would affect perception of speed and time. Participants viewed two sequentially presented motion sequences, each consisting of a single ball that entered into a circular area at a random angle, turned and changed speed twice, then exited the area. Motion sequences ranged from 1.3 to 6.8 seconds (mean = 3.5) in duration. Participants were then asked to reproduce the motion sequences by tracing their finger across the circular area. Responses were scored with three methods: 1) by duration, 2) spatially and 3) temporally (e.g., reproducing the spatial midpoint of the sequence at 50% of the reproduction's duration = no temporal error). To manipulate attention, on 67% of trials, a 75% valid pre-cue was presented before the sequence to bias their attention towards the former or latter motion sequence. We found that participants accurately reproduced the speed and duration of the motion sequence, with mean rank-order correlation between actual and reproduction duration of the motion sequence of 0.566 (0.540 even for cue invalid trials). Unsurprisingly, we found that cue validity increased spatial accuracy and cue invalidity decreased it. More interestingly, we found no evidence that cue validity improved temporal error, only that cue invalidity decreased it. Similarly, no difference in reproduction duration was observed when the cue was valid, only that the reproduced duration was shorter when the cue was invalid. We conclude that inattention reduces the perceived duration of a motion sequence, while simultaneously decreasing the accuracy of the speed of that sequence.

Acknowledgements: All authors are supported by the Japanese Society for Promotion of Science Kakenhi (GK: JP22K13874; AS: JP23KJ2029; KW: 22H00090, 24K21505)

26.474 THE OBJECT OF TIME: TEMPORAL PERCEPTION OF OBJECTS IS IMPROVED BY PROXIMITY AND SINGULARITY

Alex Ma[1], Martin Wiener[1]; [1]George Mason University

The activities of daily life necessitate moving through spaces and visually processing objects at various distances. Research shows that visual responses for objects differs to visual responses for scenes containing many objects (Josephs & Konkle, 2020). Separately, the properties of visual scenes are known to dilate and contract subjectively perceived duration (Ma, et al. 2024). These experiments explored the possibility that the perceived distance of objects within an image could influence time perception, through a dataset which categorized images into portraits of singular objects, scenes full of objects, and "reachspaces", which contained objects that appeared at a reachable distance (Josephs, et al. 2021). The first experiment (n=20) used 216 images treated to reduce the effect of low-level image properties like color and brightness, while the second experiment (n=20) additionally used contrast normalized images. Both experiments employed sub-second temporal bisection tasks, and featured an eye tracking component to control for the effects of saccade density. The results for both experiments showed that the perceived distance did not significantly dilate or contract time perception, but that the durations of progressively more proximal, singular object images were linearly perceived more precisely and processed more quickly. This pattern of results could imply the processing of singular objects in images involves a mechanism of divisive normalization with correlated noise, in which a greater number of neural processes are required for a greater number of objects. The greater number of object processes for wider scene images would then combine their results together, leading to greater variance compared to singular object images. These findings suggest that the duration of images with a large number of objects and images with a single object are processed differently at a perceptual level, with implications for both object processing and models of time perception.

26.475 OCULOMOTOR ENGAGEMENT MODULATES MAGNITUDE-BIASED INTERVAL PERCEPTION

Zhenni Wang[1,2], Lihui Wang[1]; [1]Shanghai Jiao Tong University, [2]Shanghai Mental Health Center

Magnitude biases temporal perception, with a high amount of visual stimuli perceived as lasting longer than a low amount of visual stimuli, even though they are presented for the same length of time. Despite the well-documented phenomenon, it is yet unknown if eye movements play a significant role in the magnitude-biased interval perception. In a two-alternative forced-choice task, participants were asked to compare two intervals and judge which was longer or shorter. In Experiment 1, a sequence of dots was consecutively presented in each interval. The dot was presented at a random location on the visual field and participants were asked to follow the dot with eye movements. Importantly, the number of dots in each interval was manipulated in the way that one interval included more dots than the other, even though the two intervals had the same length. The results showed that the interval with more dots were accompanied by more saccades, and were judged as lasting longer. Experiment 2 had the same design as Experiment 1 except that the dots were simultaneously rather than sequentially presented. As a result, relative to the interval with fewer dots, the interval with more dots did not induce more saccades than that in Experiment 1. Although the interval with more dots was still judged as lasting longer than the interval with fewer dots, the magnitude-biased interval perception was significantly weaker than that in Experiment 1. Together, the results suggested a critical role of oculomotor engagement in modulating the magnitude-biased interval perception.

26.476 TIME FROM A DIFFERENT ANGLE: HOW THE ANGULARITY OF SHAPES AFFECT TIME PERCEPTION

Giuliana Macedo[1], Martin Wiener[1]; [1]George Mason University

Visual stimuli are known to vary in their perceived duration, with a variety of features, including both high and low level, shaping temporal responses. For example, the size, memorability, or numerosity of stimuli lead to longer perceived intervals, thus highlighting the malleability of time perception. However, beyond stimulus size, the influence of stimulus curvature or angularity has on time perception, has yet to be explored. To investigate this, we had human participants (n=35) perform a sub-second visual temporal bisection task, in which they categorized visual stimuli with different levels of curvature into short and long duration categories. For visual stimuli, we employed the well-known "bouba-kiki" shapes, characterized by high curvature and high angularity, respectively, along with three linearly interpolated shapes between them (5 levels total). Studies on the bouba-kiki effect, which is the association of naming curved shapes "bouba" and angular shapes "kiki", has been proven robust across languages and cultures (Cwiek, et al. 2021). Here, we hypothesized that more angular "kiki"-like shapes would dilate time. We additionally explored if these shapes with linguistic associations ("bouba" for curved, "kiki for angular) had any impact on time perception. As an additional control, we had subjects initially equalize the internal area of the shapes, to ensure no influence of size on time judgments. Finally, after the bisection task, we asked participants to name the shapes according to the "bouba-kiki" naming system to look for the influence of language on shape preference. Surprisingly, we found no dilation effects; instead, we discovered a linear increase in precision (as indexed by the coefficient of variation), as well as faster reaction times for progressively more angular stimuli, with no impact of naming preference. These findings suggest that the angularity of shape stimuli drives faster and more precise time judgments, which we interpret using an information-processing framework.

26.477 PREDICTED TIMING OF SOUND INFLUENCES AMBIGUOUS PERCEPTION IN THE BOUNCE/STREAM ILLUSION

Yuyeon Jung[1] (ch_y_y@naver.com), Seyoon Song[1], Chai-Youn Kim[1]; [1]Korea University

Previous studies have shown that our expectations can influence how we interpret ambiguous stimuli. For example, in the bounce-stream illusion, where two discs are perceived as either passing through or bouncing off each other, prior visual history influences perception, with increased trajectory switches enhancing the likelihood of perceiving "bouncing"(Grove et al., 2016). However, the role of temporal predictions in this process remains unclear. This study investigates how temporal predictive cues impact the perception of the bounce-stream illusion. A brief sound occurring near visual coincidence is known to enhance the likelihood of perceiving the discs as bouncing(Sekuler et al., 1997). To manipulate sound timing, stimulus onset asynchronies (SOAs) of 0ms (within the temporal binding window; TBW) and -300ms (outside the TBW) were used to create the two levels of the predictive condition. In this condition, predictive cues signaled when the sound would occur, allowing for the investigation of predictions in resolving perceptual ambiguity. Specifically, each SOA was paired with a fixation color (blue or pink of equal lightness), and these associations were established during the training session. In the main session, 75% of trials in the predictive condition presented sound at the predicted timing, while 25% were no-sound trials. The non-predictive condition followed the conventional bounce-stream illusion paradigm. In this condition, five SOAs (±300, ±150, 0ms) and no-sound trials were evenly distributed. Results showed that, in the absence of sound, 0ms predictive cues were associated with significantly greater bouncing perception compared to -300ms cues ($p = .036$) and marginally greater than non-predictive cues ($p = .055$). However, in conditions with sound, no significant differences were observed between predictive and non-predictive cues, suggesting that participants relied more on sensory information than expectations. These findings highlight the crucial role of expectations in resolving perceptual ambiguity, affecting the perception of either "bouncing" or "streaming.

Supported by NRF-2023R1A2C2007289

26.478 HOW ENSEMBLE TEMPORAL STATISTICS INFLUENCE DURATION PERCEPTION OF VISUAL EVENTS

Valeria Centanino[1] (vcentani@sissa.it), Gianfranco Fortunato[1], Domenica Bueti[1]; [1]International School for Advanced Studies (SISSA)

The duration perception of a sensory event is shaped by the statistical distribution of the durations at hand, namely the temporal context. For example, a given physical duration is reproduced differently when presented within distinct but partially overlapping duration ranges, systematically showing a bias toward the mean of the range at hand. Temporal context can also change when the frequency of presented durations varies within the same temporal range, but the effects of such manipulation are less well understood. At brain level, functional MRI studies suggested that the processing of brief visual events relies on tuning mechanisms whose properties change across the cortical hierarchy. How these tuning properties adapt to contextual biases, however, remains an open question. In this study, thirty participants reproduced eight visual durations presented under either a uniform or a positively skewed distribution. Results showed that, under the skewed condition, all durations were reproduced as longer, with the greatest bias observed for the shortest durations. However, this bias varied significantly across participants, highlighting a non-trivial and nuanced influence of temporal statistics on reproduction performance. To investigate the neural underpinnings of this effect, a separate group of eight participants performed the same task while brain activity was recorded at 7T-MRI. The blood-oxygen-level-dependent signal was then modeled with a unimodal neuronal response function. Preliminary findings showed that duration preferences in medial and lateral premotor areas and anterior insula shifted toward longer durations under the skewed condition. This shift was most pronounced in neuronal populations preferring shorter durations under the uniform condition, mirroring behavioral biases. In contrast, the sensitivity of neuronal responses remained stable across conditions. These findings suggest that temporal statistics shape our perception of visual durations by shifting neuronal preferences within higher-level regions of the cortical timing hierarchy.

26.479 A DISTINCT FUNCTIONAL ROLE OF PRIMARY VISUAL AND MEDIAL PREMOTOR CORTEX IN VISUAL DURATION PROCESSING

Domenica Bueti[1] (domenica.bueti@sissa.it), Alessandro Toso[1,2], Gianfranco Fortunato[1], Chiara Zanonato[1,3], Mathew Diamond[1]; [1]SISSA, [2]Section Computational Cognitive Neuroscience, Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg- Eppendorf, [3]Max Planck Institute for Biological Cybernetics

In humans, the duration processing of a visual event has been shown to entail a wide network of brain areas including primary visual cortex and supplementary motor area. The functional role of these brain regions in temporal computations is far from clear. A simple hypothesis is that V1, conveying the sensory input, encodes stimulus duration while SMA, at the other extreme of a hypothetical hierarchy, decodes it for task-related purposes. Here we used paired pulse transcranial magnetic stimulation (ppTMS) to test this hypothesis. Fifteen healthy volunteers were asked to discriminate the duration of pairs of visual stimuli while ppTMS was applied over right V1, SMA and the Vertex at different timings from the onset of the first stimulus of the pair i.e., 0, 60, 90 and 100% of the total stimulus duration. Compared to vertex stimulation, we observed a significant worsening of temporal discrimination thresholds after ppTMS over V1 at 60% and over SMA at 90% and 100% of the stimulus duration. A second experiment showed that the performance worsening after V1 stimulation was not due to an interference with low-level visual processing. Four different versions of a leaky integrator model were used to fit the data. All models assumed that visual time perception results from the leaky integration of a stimulus's luminance, with this process being inherently noisy. The models varied in their hypothesized effects of TMS, proposing that these effects could occur either at the perceptual or the input level, by altering the mean or variance of the noisy process. The two winning models suggested that ppTMS influenced the noise variance of the integration process, while V1 and SMA results were best explained by assuming TMS interference at the input and perceptual level, respectively. These results highlight a distinct functional role of V1 and SMA in duration processing.

26.480 TESTING THE LIMITS OF TEMPORAL PHASE PERCEPTION OF ISOLUMINANT COLOR MODULATION

Jaeseon Song[1] (jaeseon@udel.edu), Andrew Lisech[2], Josiane Mukahirwa[3], Keith Schneider[4]; [1]University of Delaware

Introduction: Temporal phase perception, which involves comparing modulations between temporal stimuli, is significantly slower than peripheral stimulus detection, suggesting a central temporal bottleneck. While phase perception thresholds for luminance flicker (7–10 Hz) have been studied, those for chromatic flicker remain largely unexplored. The critical flicker fusion frequency (CFF) for chromatic flicker (~30 Hz) is about half that of luminance flicker (~60 Hz), reflecting slower mechanisms in chromatic processing. It remains unclear whether this slower peripheral processing plays a greater role in central phase comparisons or if a general central temporal bottleneck, independent of peripheral stimulus characteristics, has a greater impact. To investigate, we measured phase discrimination thresholds for chromatic modulations across a range of temporal frequencies (TFs). Methods: Phase discrimination thresholds were measured using two chromatic circles rotating through isoluminant hue space (0–360° in HSV color space). Isoluminence was determined for R, G, and B channels using flicker photometry, and the isoluminant HSV values were calculated. Thresholds were measured across seven TFs (1–20 Hz) under two distance conditions (3° and 12°). Stimuli were displayed using a 480 Hz VPixx PROPixx projector at 4° eccentricity, with the circles positioned to the left and right of the screen center. Participants adjusted the test chromatic flicker's TF and phase to match a reference flicker. Results: Phase error magnitude increased with TF under both distance conditions, indicating greater difficulty in phase discrimination at higher frequencies, particularly at 3–5 Hz and above. Phase error was generally lower in the near condition compared to the far condition. Conclusion: Our findings show that phase perception for chromatic modulations operates at a temporal scale slower than phase comparisons in luminance modulations, challenging the idea of a constant central temporal bottleneck. Instead, this suggests that the temporal limitations of peripheral stimuli play an important role for central phase comparisons.

26.481 TIME-BOUND: SHARED TEMPORAL LIMITS IN VISUAL AND AUDITORY PHASE PERCEPTION

Andrew Lisech[1] (alisech@udel.edu), Jaeseon Song[1], Josiane Mukahirwa[1], Keith A. Schneider[1]; [1]University of Delaware

Introduction: The mechanisms governing central temporal limits in the brain remain poorly understood. Previous research suggests that our capacity to perceive high-frequency temporal flicker stimuli exceeds our ability to discern changes in their temporal properties. Here we investigated whether a central bottleneck constrains temporal phase discrimination across the visual and auditory systems. Methods: In separate visual and auditory experiments, fifteen participants judged whether sinusoidally oscillating stimulus pairs were in-phase or 180° counter-phase. In the visual experiment, dichoptically presented Gaussians were compared as they oscillated between light and dark states. In the auditory experiment, participants judged the phase of amplitude-modulated dichotic tone sequences. Using the method of constant stimuli, phase discrimination thresholds were measured across nine randomly-interleaved amplitude frequencies (1–30 Hz), with 25 repetitions per frequency. Accuracy data were fit to a cumulative normal function to derive 75% accuracy thresholds for each participant. Interstimulus distances were increased to minimize interference from low-level motion detection and other peripheral processes. Results: The study's main finding revealed that both visual and auditory systems exhibited phase discrimination limits between 7–12 Hz, indicating similar temporal constraints across modalities. Conclusion: Comparable frequency thresholds in vision and audition lend support to the hypothesis of a central processing bottleneck for temporal phase discrimination, rather than independent modality-specific limitations. These shared limitations may reflect the temporal properties of a unified temporal mechanism governing phase perception. To further investigate central temporal limits and their related mechanisms, future research should examine how this bottleneck operates in audiovisual integration and other multisensory contexts.

Temporal Processing: Clinical

SATURDAY, MAY 17, 2:45 – 6:45 PM, PAVILION

26.482 TEMPORAL CONTOUR INTEGRATION IN CHILDREN WITH AMBLYOPIA

Yan-Ru Chen[1] (2101110693@stu.pku.edu.cn), Shu-Qi Jiang[1], Xiang-Yun Liu[2], Jun-Yun Zhang[1]; [1]Peking University, [2]The Affiliated Tengzhou Hospital of Xuzhou Medical University

Purpose: The brain perceives global contours by integrating discrete but collinear stimuli. While previous studies have primarily explored contour integration deficits in amblyopia within the spatial domain, this study investigates how amblyopia affects contour integration over time, and examines the relationships between temporal contour integration deficits and visual functions. Methods: Nineteen amblyopic children (10.9 ± 2.7 years; 17 anisometropic, 2 anisometropic/strabismic mixed) and twenty-six visually normal children (10.5 ± 1.8 years) participated in this study. We measured the accuracy of detecting tilted contour paths formed by collinear Gabor elements under slit-viewing conditions for amblyopic eyes (AEs) and fellow eyes (FEs) separately at two spatial frequencies (1.5 cpd and 3 cpd). The slit width, orientation jitter of contour elements, and stimulus moving speed were varied in separate runs. Visual acuities for both eyes and Randot stereoacuity were assessed before the testing. Results: (1) The AEs exhibited deficits in temporal contour processing compared to FEs at average_interelement_distances (AIED) of 0.8, 1, 2, and 4. Slit width thresholds for AEs were significantly higher than for FEs, and were correlated with AE visual acuity at 1.5 cpd. (2) Increased orientation jitter of contour elements and changing stimulus speed reduced contour detection accuracy, eliminating the differences between AEs and FEs. (3) There were significant temporal contour deficits in AEs compared to control eyes, while no differences between FEs and control eyes. (4) Temporal contour integration deficits in amblyopia were not correlated stereoacuity, or spatial contour integration deficits. Conclusions: Amblyopic children demonstrate deficits in temporal contour integration in AEs, which appear to be independent of spatial contour integration deficits. Children with more severe visual acuity impairments in AEs tend to have more pronounced temporal contour integration deficits. These findings suggest that amblyopia is associated with temporal deficits in visual integration, in addition to known spatial deficits.

Supported by the Natural Science Foundation of China grants 31970978 and 32371079 (JYZ; Beijing, China).

26.483 THE INFLUENCE OF ACTION VIDEO GAME PLAY ON CRITICAL FLICKER FUSION

Colin R. Gardner[1], Lisa Renzi-Hammond[2], Drew H. Abney[1], Cassandra M. Mesh[1], Jessica Lin[1], Billy R. Hammond[1]; [1]University of Georgia - Department of Psychology, [2]University of Georgia - Institute of Gerontology

Video games are a common feature of modern life with over 66% of Americans and 40% of the global population playing some form of video game in 2022. Visual-motor tasks, in general, have been shown to produce concomitant changes in key central nervous system structures. Action video games (AVGs) require a mix of cognitive (e.g., attention) and visual-spatial skills that seem particularly well suited to promote changes in neural fundamentals like visual processing speed (often quantified as critical flicker fusion thresholds, CFF). In this study, we compared the CFF values of action video gamers (AVGPs) to non-action video gamers (NAVGPs). CFF was measured using a circular, 1°, 660 nm (20 nm half bandpass) (Nichia Corp., Mountville, PA) stimulus that flickered at 100% modulation. The target stimulus was surrounded by a 5.5°, 660 nm surround (average luminance, 25 cd/m2). Participants viewed the stimulus through a 3-mm artificial pupil (controlling for luminance variation) while focusing on a central 5' fixation point. Forty-four (22 in each matched group) young, healthy participants with normal vision were tested. CFF thresholds were derived from the averages obtained using the method of limits: 6 trials, 3 ascending and 3 descending. AVGPs ($M = 28.72$, $SD = 1.92$) had a significantly higher CFF than NAVGPs ($M = 26.72$, $SD = 1.64$), $F(1, 42) = 13.82$, $p < .001$. These results are consistent with the general hypothesis that AVG play leads to improved temporal processing speeds. This may prove useful for future interventions targeting neural pathologies that impact temporal processing ability. CFF is often considered fundamental to a number of other higher cognitive skills which may also be influenced by AVG play.

26.484 TEMPORAL NETWORK DYNAMICS IN THE INTACT HEMISPHERE FOLLOWING PEDIATRIC CORTICAL RESECTION

Sophia Robert[1] (srobert@andrew.cmu.edu), Michael C. Granovetter[2,3], Christina Patterson[2], Marlene Behrmann[1,2]; [1]Carnegie Mellon University, [2]University of Pittsburgh, [3]New York University

Many children demonstrate remarkable cognitive capacities following extensive cortical resections, underscoring the potential of neural plasticity to support recovery or maintenance of function during development. While previous studies have used static functional connectivity to explore neural differences in pediatric epilepsy patients after cortical resection, our approach characterizes the temporal network dynamics of the intact hemisphere of these patients. Blood oxygen-level dependent (BOLD) responses were recorded from patients with childhood cortical resections for drug-resistant epilepsy (n = 10, ages 15–37, 4 left- and 6 right-hemisphere resections) and controls (n = 9, ages 8–38) while viewing and listening to an 11-minute movie segment. BOLD responses were mapped to the cortical surfaces of patients' intact hemispheres and both hemispheres of controls and parcellated into 180 surface-defined regions using the Glasser HCP atlas. Inter-subject correlations of time-varying activity were computed across 180 regions and 22 functional networks to assess temporal dynamics of hierarchical processing. Results revealed significant network-specific reductions in inter-group synchrony in patients, regardless of resected hemisphere, compared to controls. Patients displayed divergent temporal profiles in many regions including visual, auditory, motor, and parietal areas. In contrast, some regions, such as the lateral temporal and inferior frontal cortices, exhibited similar temporal dynamics across patients and controls, potentially reflecting preserved connectivity in these areas. These differences could indicate adjustments to altered processing demands in patients, who rely on largely or entirely unilateral networks for functions typically distributed bilaterally. Our findings characterize the temporal dynamics of neural networks after cortical resection, with disrupted temporal dynamics in sensory and associative regions contrasting with preserved function in others, advancing our understanding of neural organization in pediatric epilepsy.

26.485 USING VIRTUAL PATIENTS TO PREDICT PERCEPTUAL PERFORMANCE AFTER OPTOGENETIC SIGHT RECOVERY

Vaishnavi B Mohan[1] (vbmohan@uw.edu), Ezgi Irmak Yucel[1], Nil Altinordu[1], Geoffrey M Boynton[1], Ione Fine[1]; [1]University of Washington

Introduction: Optogenetics is a promising technology to restore vision in individuals with blindness due to retinal disorders. Light-sensitive proteins (opsins) are introduced intravitreally and bind to surviving non-photosensitive retinal cells to elicit a neural response to visual stimuli. Much of this research is done with mouse models, where it is not obvious how a given opsin would translate to human visual perception. Our goal was to provide estimates of potential patients' spatio-temporal contrast sensitivity and perceptual experiences. We created 'virtual patients' consisting of neurotypical subjects viewing stimuli modified to match the temporal properties of two state-of-the-art opsins 4xBGAG12,460:SNAP-mGluR2 and ChRmine. Methods: A non-linear model based on single cell electrophysiological data describing the sensitivity and temporal dynamics of each opsin was used to spatiotemporally modulate a counterphase modulated Gabor stimulus to replicate the predicted retinal stimulus of optogenetic patients. Contrast detection thresholds for optogenetic and control stimuli were obtained using a 2AFC quick CSF Bayesian adaptive procedure across a range of spatial and temporal frequencies. Results: Contrast sensitivity for 4xBGAG12,460:SNAP-mGluR2 declined rapidly with increasing temporal frequency compared to the control stimulus, consistent with this opsin's slow temporal dynamics. Sensitivity to opsin ChRmine was lower overall, but there was little decline in sensitivity at higher temporal frequencies. Conclusion: Using this virtual patient paradigm, we demonstrate the trade-off between sensitivity and temporal resolution for two example opsins. 4xBGAG12,460:SNAP-mGluR2 would be expected to outperform ChRmine under many laboratory tests, such as a Snellen eye chart for a patient capable of maintaining stable fixation. However ChRmine would be predicted to outperform 4xBGAG12,460:SNAP-mGluR2 during outdoor navigation in daylight. Ultimately, our goal is to predict patients' performance across a wide range of clinically relevant tasks directly from an opsin's temporal dynamics, paving the way for informed optimization of sight-restoration therapies before clinical application.

NIH grant R01EY014645

26.486 REPRODUCIBILITY OF TEMPORALLY EVOLVING SEIZURE PATTERNS AND NETWORK CONNECTIVITY ACROSS MULTIPLE SEIZURE ONSETS IN HUMAN FOCAL EPILEPSY

Yaohong Wei[1], Wei Zhang[2], Sinclair Xinzhou Liu[2], Biao Han[1], Qi Chen[1]; [1]South China Normal University, Guangzhou, China, [2]Beijing Tsinghua Changgung Hospital, Beijing, China

Clinical semiology, waveform patterns, frequency band characteristics, and neuronal spiking activity exhibit notable reproducibility across seizures in the same patient. Drug-refractory focal epilepsy, recognized as a network disease, involves dynamic interactions across the epileptogenic zone (EZ), propagation zone (PZ), and non-involved zone (NIZ) during interictal, preictal, and ictal periods. However, whether these reproducible features distinctly delineate the EZ from the PZ and NIZ and the underlying temporal dynamics remain unclear. This study analyzed 14 patients with focal epilepsy characterized by low-voltage fast activity (LVFA), examining the reproducibility of raw signals, power spectra, and connectivity patterns during the interictal, preictal, and ictal periods across multiple seizures. Our findings reveal that neural activity and connectivity patterns in the EZ maintain high reproducibility across seizures. During the ictal phase, raw signals, power spectrum in delta (1–3 Hz) and gamma-to-ripple (30–250 Hz) bands, and connectivity patterns within the EZ demonstrated greater reproducibility than in the PZ, highlighting the stability of seizure generation and propagation mechanisms. Moreover, during the preictal phase, the power spectra of the beta band (14-30 Hz) and connectivity patterns within the EZ were both consistently more reproducible than those in the PZ. These findings enhance our understanding of epileptic network dynamics and may improve diagnostic precision and seizure modeling.

SUNDAY MORNING POSTERS IN BANYAN BREEZEWAY

Perceptual Organization: Ensembles

SUNDAY, MAY 18, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

33.301 ENSEMBLE CODING ABILITIES ARE GOVERNED BY A (MOSTLY) DOMAIN GENERAL MECHANISM

Greer Gillies[1], Keisuke Fukuda[2], Jonathan Cant[3]; [1]University of Toronto, [2]University of Toronto, Mississauga, [3]University of Toronto, Scarborough

Ensemble coding operates across a range of visual features (e.g., average orientation, average color) and stimuli (e.g., triangles and faces). However, ensemble displays usually contain stimuli from only one category. It is unclear if the visual system can extract summary statistics (1) across different stimulus categories to make comparisons, or (2) from single displays containing multiple stimulus types. In Experiment 1, participants compared average orientation (same/different) across two ensembles that were made up of either the same or different stimulus types (faces, shoes, triangles). We found that participants were just as accurate when the stimulus types were the same for both ensembles compared to when they were different, suggesting that ensemble abilities are governed by domain-general mechanisms. In Experiments 2 and 3, participants were shown a single ensemble that was made up of either one stimulus type (e.g., triangles) (pure condition) or two (e.g., triangles and faces) (mixed condition), and then determined whether the direction of a subsequent probe arrow was representative of the average orientation. There were no differences across stimulus types in the pure condition. However, in the mixed condition, stimulus specificities emerged. Specifically, when the ensemble contained triangles, participants had a hard time rejecting the probes that matched the orientation of just the triangle subset, and further evidence of this bias was found via a continuous-report task in Experiment 4. In summary, when every item within an ensemble is from the same category, the ability to extract summary statistics is similar across ensembles composed of different stimuli, but having more than one type of stimulus present in an ensemble interferes with the extraction/report of summary statistics. This suggests that ensemble coding abilities are governed by a largely domain-general mechanism, but specificities do emerge under certain possibly relating to the ease of computing shape-based orientation values.

33.302 EVIDENCE FOR PERCEPTUAL INTERDEPENDENCE BETWEEN NUMBER AND AREA

Moxuan Liu[1], Stella F. Lourenco[1]; [1]Emory University

That people readily perceive the number of objects in visual displays is uncontroversial. How number is perceived, however, remains the topic of ongoing debate, particularly the question of whether or not number is perceived independently of other (co-occurring) magnitude dimensions such as cumulative area (i.e., the total area of a set of objects). Some researchers argue that the influence of area on number judgments reflects post-perceptual "late-stage" decision-making. Others, however, have proposed that the relation between number and area may be perceptual in nature. To delineate between these two perspectives, we tested adult participants on a (delayed) match-to-sample task (DMTS, N = 39) designed to rule out decision-stage effects. Participants were instructed to select the choice display (from two options) that matched the sample display in number. Choice stimuli varied in either area or luminance, with differences between area and luminance matched for difficulty (confirmed by a separate group of participants, N = 39). Analyses of participants' responses (choices and RTs) on the DMTS task revealed greater interference by area than luminance on number matching, despite identical task demands for area and luminance conditions. This finding suggests that the relation between number and area is not simply due to task-based decision-making. To rule out an alternative explanation based on perceptual salience, a subsequent experiment with a triplet odd-one-out task was conducted (N = 24). Analyses revealed variation in the salience of area relative to luminance. Yet there was still greater interference by area than luminance on the DMTS task when area and luminance were comparable in salience. Altogether, our findings suggest that the perception of number is not independent of area information within visual displays. They also raise questions about the nature of the relations between number and other visual attributes such as luminance.

33.303 INVESTIGATING THE ROLE OF RETINAL AND PERCEIVED SIZE IN ENSEMBLE SIZE JUDGMENTS IN VIRTUAL REALITY

Katrín Fjóla Aspelund[1] (kfa1011@usnh.edu), Ömer Dağlar Tanrikulu[1]; [1]University of New Hampshire

This study examines whether the retinal size of objects influences ensemble size judgments in realistic contexts. Prior research has demonstrated that average size estimates are based on perceived size representations that are rescaled based on viewing distance. However, these studies often relied on explicit ensemble judgments on simplified stimuli that lacked naturalistic visual properties. To address this limitation, we use an immersive virtual reality (VR) environment with familiar objects and full-depth cues to investigate whether participants rely on retinal size or perceived size when averaging objects at distinct depth planes. Participants completed a two-alternative forced-choice (2AFC) task in a realistic supermarket VR environment, comparing the average sizes of two groups of virtual flour sacks placed on checkout counters at distances of 2.5m and 5m from the observer. We used a method of constant stimuli to calculate participants' Point of Subjective Equality (PSE) values. The closer group had a fixed mean size, while the farther group varied in mean size. Results showed a bias in participants' judgments, with the farther group needing to be slightly larger to appear equal to the closer group. PSEs were larger than the fixed mean of the closer group, indicating partial size-distance rescaling. Despite the availability of rich depth cues and realistic object properties, ensemble size judgments were still influenced by retinal size, indicating a measurable bias toward retinal size in average size judgments. Unlike prior studies assuming a linear relationship between retinal and perceived size, our 2AFC method avoided this assumption. The task was also implicit and realistic, asking participants which group contained more flour, simulating practical decision-making. Our results highlight the importance of studying ensemble processing in realistic environments to understand how it functions in daily life.

33.304 HOW DOES TASK-IRRELEVANT VARIABILITY INFLUENCE ENSEMBLE JUDGMENTS?

Suyeon Kim[1], Oakyoon Cha[1]; [1]Sungshin Women's University

Even a mundane visual scene, such as a crowd on a street, presents a vast amount of visual information. Facing this challenge, the visual system extracts statistical properties to reduce redundant information from groups. Last year, we found that variability judgments were influenced by local feature contrast, with this influence increasing as task-irrelevant variability increased. Building on Pascucci et al. (2021)'s finding that attention was unevenly distributed during ensemble judgments, we speculated that participants might have applied a narrower attention window when task-irrelevant variability raised the attention load. In the present study, we investigated whether task-irrelevant feature variability influences the distribution of attention during ensemble judgments. Participants briefly viewed a 5-by-5 array of oval shapes with varying orientations, and reproduced the average orientation of all ovals. We manipulated the color variability of oval shapes (Experiment 1) and the color variability embedded in the background (Experiment 2). Embedding colorful circles in the background allowed us to increase the task-irrelevant variability further in Experiment 2. For each trial, we calculated the average orientation in three different windows (one oval in the center, ovals in a 3-by-3 window in the center, and all 25 ovals) and assessed the contribution of each window to average judgments. In Experiment 1, task-irrelevant variability did not influence how the three windows contributed to the average judgments. However, in Experiment 2, task-irrelevant variability increased the contribution of the 3-by-3 window and reduced the contribution of the 5-by-5 window, suggesting that attention was less evenly distributed and shifted toward the center. These results suggest that task-irrelevant feature variability can influence the distribution of attention, thereby making ensemble judgments more reliant on fewer numbers of objects.

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).

33.305 OUTLIER DISCOUNTING IN TEMPORAL ENSEMBLES

Zhilin Zhang[1] (zzhang5@gradcenter.cuny.edu), Tatiana Aloi Emmanouil[1,2]; [1]The Graduate Center, City University of New York, [2]Baruch College, City University of New York

Ensemble perception has been shown to discount outlier items whose properties lie outside the distribution of the ensemble. Previous research investigated the mechanisms of outlier discounting for spatial ensembles that varied in orientation and found evidence for an iterative outlier exclusion process. In the current study, we examined whether and how outlier filtering occurs for ensembles in which objects are presented sequentially in time. In Experiment 1, participants reported their estimation of the average tilt in sets of 12 oriented lines (2 outliers) presented either sequentially (temporal ensembles) or simultaneously (spatial ensembles). To evaluate outlier filtering, we examined whether the standard deviation of their estimation to the local mean (average tilt without outliers) was smaller than that to the global mean (average tilt with outliers). The results showed outlier filtering of similar magnitude for temporal and spatial ensembles. Interestingly, for temporal ensembles, we observed reduced filtering for outliers presented later in the sequence. In Experiments 2 and 3 we further investigated this recency effect. Experiment 2 tested whether this effect is due to reduced processing time between the late outliers and the response probe. We did not find any evidence for this possibility since varying the time interval between ensemble and response had no effect on filtering. Experiment 3 tested the hypothesis that reduced filtering for late outliers was due to the lack of post-outlier context. The results supported this prediction, showing that outlier filtering was restored when appending additional non-outlier items in the sequence. Overall, this study demonstrates that outliers in temporal ensembles are discounted, although this effect is diminished for outliers appearing late in the sequence. Our findings suggest that temporal ensemble perception may involve a dynamically updating process that reinforces discounting of outliers whose deviation is confirmed by subsequently encoded items.

We thank Holland Brown, Elijah Burgos, Tanzifa Akther and Selly Kebbe for assistance in data collection. This project is supported by CUNY Professional Staff Congress grant PSC-TRADA-55-414 to Tatiana Emmanouil.

33.306 DATA VISUALIZATION CHOICES AFFECT PATTERN DETECTION IN 2X2 GRAPHS

Nestor Matthews[1] (matthewsn@denison.edu), Samantha Kozlowski[1], Megan Broderick[1]; [1]Denison University

Introduction: The COVID pandemic highlighted the need for health care professionals to create easy-to-read data visualizations for the general public. For example, plots showing COVID mortality rates by age (under 65 vs 65+) and vaccine status (vaccinated vs unvaccinated) could save or cost lives, depending on how easily general audiences perceptually organize the data visualization. We therefore conducted the present perceptual learning study to investigate how different data visualizations (bar graphs vs line graphs) affect naive participants' success in detecting visual patterns. Method: 416 naive Prolific 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 / right mean-height differences), Factor B Main Effects (black / white mean-height differences) or Interactions (slope differences). Before collecting data, we preregistered the prediction that failures to perform above chance would occur significantly more frequently for Factor A Main Effects than for each of the other two effects [https://osf.io/qwsfc]. We tested this prediction separately for line graphs and bar graphs. Results. For line graphs, failures to perform above chance occurred significantly more frequently for Factor A Main Effects than for Factor B Main Effects ($p = < .001$, Cramér's $V = 0.58$) and for Interactions ($p = < .001$, Cramér's $V = 0.66$). For bar graphs, failures to perform above chance occurred significantly more frequently for Factor A Main Effects than for Interactions ($p = < .001$, Cramér's $V = 0.44$) but not for Factor B Main Effects ($p = 0.340$, Cramér's $V = 0.13$, n.s.). Conclusion: General audiences benefit from 2x2 line graphs that show the most important main effect plotted as Factor B (black / white mean-height differences).

33.307 EXPLICIT ACCESS TO DETAILED FEATURE DISTRIBUTION REPRESENTATIONS

Vladislav Khvostov[1,2], Julie Golomb[1], Árni Gunnar Ásgeirsson[3], Árni Kristjánsson[4]; [1]Department of Psychology, The Ohio State University, USA, [2]HSE University, Russia, [3]University of Akureyri, Iceland, [4]University of Iceland, Iceland

The human visual system can quickly process groups of objects (ensembles) and build compressed representations of their features. The current thinking is that conscious access to this representation is very limited: observers cannot report any distributional aspects beyond simple summary statistics, such as the mean or variance. However, people subjectively feel they can perceive much more than that. Here, we introduced a new paradigm (Feature Frequency Report) to thoroughly probe this explicit representation. We showed that in this case, the subjective impression is closer to the truth than the scientific consensus. Observers viewed 36 disks of various colors for 800 ms and then reported the frequency of a randomly-chosen color using a slider (0-8 disks). In Experiment 1 (N=10), the colors of the disks had a Gaussian, uniform, or bimodal distribution with a random mean color. The distributions of averaged responses (both aggregated and at the individual-observer level) followed the shape of the presented distribution, revealing that people indeed represent more than summary statistics. Model simulations demonstrated that performance reflected integrated information from the whole set rather than the sub-sampling of a few items. In Experiment 2 (N=17), observers were presented with symmetrical and skewed Gaussian distributions and demonstrated that they could also explicitly represent distributional skewness. These results show that after only a brief exposure to a color set, the visual system builds detailed representations of feature distributions that observers can explicitly access. This necessitates a fundamental rethinking of how ensembles are processed. We suggest that distribution representation, not summary statistics, is the default way to consciously represent object groups. The summary statistics can be simple derivatives of these distribution representations rather than the core units of ensemble perception. Our finding helps to explain people's impression of having a rich perceptual experience despite severe attentional and working memory limitations.

Supported by grant #228366-051 from the Icelandic Research Fund (AK, AGA), NIH R01-EY025648 (JG), and NSF 1848939 (JG)

33.308 THE PARADOX OF CERTAINTY: GRAPHED ENSEMBLES CONVEY AVERAGES BETTER THAN GRAPHED AVERAGES

Yang Wang[1] (wangy0802@hotmail.com), Timothy F Brady[1], Sarah H Kerns[3], Jeremy B Wilmer[2]; [1]UCSD, [2]Wellesley College, [3]Dartmouth College

Data visualizations commonly use averages to enhance interpretability, but this may not always lead to accurate comprehension. Ensemble processing—the human ability to perceive summary statistics from groups of objects—is effective in various domains, yet its role in graph interpretation is underexplored. This study examines how individuals interpret averages using different graphical representations. We tested 157 UCSD undergraduate students across four conditions. In the 'explicit mean' conditions, bar graphs and dot plots explicitly displayed the average, omitting raw data. In the 'individual data' conditions, cloud plots and sinaplots depicted individual data points without explicitly indicating the mean. Surprisingly, the bar graph condition resulted in the least accurate average estimations. Despite clear labels and instructions, many participants misinterpreted the bar's height, often assuming the midpoint represented the average instead of the top edge. This aligns with the "Bar-Tip Limit" error, where viewers mistakenly perceive the bar's tip as the data's outer limit. Dot plots yielded more accurate average estimations, providing clearer guidance on average values, though this precision may impart a false sense of certainty by failing to convey the natural variability in the data. In the 'individual data' conditions, participants effectively estimated averages in cloud plots and sinaplots for normally distributed data, with variability reflecting natural uncertainty. For skewed distributions, estimates were somewhat less accurate, showing a bias toward the median. These findings suggest that visualizing individual data points leverages ensemble perception abilities, enabling viewers to extract meaningful averages while preserving data variability. Overall, the study shows that bar graphs are prone to misinterpretation when representing averages, whereas dot plots offer clearer guidance. More importantly, visualizing individual data points, as in cloud plots and sinaplots, appears superior for conveying meaningful averages with inherent data uncertainty, emphasizing the need to select appropriate graphical representations for effective data interpretation.

33.309 CONTEXTUAL PERCEPTION OF EMOTIONAL FACES IN DYNAMIC CROWDS

Görkem Er[1] (gorkemerr@gmail.com), Timothy D. Sweeny[1]; [1]University of Denver

Perception of facial expressions is highly contextual. For example, when an individual face is seen around other faces, people tend to say its expression looks like the others in the crowd. However, this effect of perceptual averaging has only been shown when faces are static and perceivers must attend to many faces at once. Here we examined (1) whether this effect persists in more rich and realistic contexts, when faces are dynamic and attention can be biased to a single face in a crowd. Additionally, when a face is dynamic, perceivers may overestimate its emotional intensity in the direction of its changing expression (e.g., perceiving the expression at an intensity it appears to be heading towards rather than its actual intensity), which we refer to as expression-trajectory bias. We examined (2) whether crowds exaggerate this effect. Observers rated the emotional intensity of a target face that appeared at a random location, dynamically expressing increasing or decreasing intensities of anger or happiness for 333-ms. This target face appeared among three non-target faces, which transitioned from neutral to full expressions of happiness or anger across 1,666-ms. We found two distinct effects of the dynamic crowd context on ratings of individual's expressions. First, ratings were pulled towards the emotion of the crowd; angry faces appeared less intense among happy crowds and happy faces appeared less intense among angry crowds. Second, we found clear evidence of expression-trajectory bias in judgments of individuals' dynamic expressions. This bias, however, was mitigated in the context of crowds. Crowds may thus serve as a reference when judging an individual's dynamic expression. These two simultaneously occurring effects extend prior findings, demonstrating that contextual effects in the perception of facial expression can occur when faces are dynamic and even when selective attention is engaged.

33.310 VISUAL ADAPTATION TO ARBITRARILY ASSIGNED VALUE

Sam Clarke[1], Sami Yousif[2]; [1]University of Southern California, [2]University of North Carolina, Chapel Hill

Visual adaptation has been described as a "powerful tool for dissecting vision by exposing the mechanisms that are adapting" (Webster, 2015; p. 547), an "intrinsic feature of visual processing" which "reaches the status of a universal law" (p. 548). Others have described it as "a largely unambiguous and uncontroversial way to identify visual processing" (Kominsky & Scholl, 2020; p. 3). In recent years, the scope of visual adaptation has expanded dramatically, to include not only things like color and motion but also higher-level visual features like size, number, causality, and variance. Here, we expand the reach of visual adaptation even further: We demonstrate robust, spatiotopic visual adaptation to arbitrarily assigned value. In our task, observers were introduced to displays of fake "coins" with values ranging from one to five. They were told that they will be completing a quantity discrimination task in which their goal is to select the side with more total value. Prior to making each judgment, however, participants adapted to separate displays of coins. In three separate experiments, we found robust evidence of value adaptation: When participants adapted to a high-value display, they were subsequently more likely to indicate that the contralateral side was greater in value. *Visual* adaptation to value seems not to be possible, however: Value is not — and could not be —a perceptual dimension. The choice that observers make depends on their knowledge of the coin values, which is arbitrarily and rapidly mapped onto the visual information. Yet according to the agreed-upon standards for what constitutes visual adaptation, we have documented robust evidence of value adaptation. These findings thus apply pressure to our understanding of visual adaptation as a phenomenon, specifically the (popular) argument that adaptation is somehow uniquely indicative of perceptual processing.

3D Processing: Shape

SUNDAY, MAY 18, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

33.311 UNFOLDING SPATIOTEMPORAL REPRESENTATIONS OF 3D PERCEPTION IN THE HUMAN BRAIN: AN FMRI-EEG FUSION STUDY

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

Although visual input is initially recorded in 2D on our retinas, we live in a 3D world, and our visual systems must integrate 2D representations with various depth cues to achieve 3D perception. However, it remains unclear how exactly the brain integrates 2D and depth information into 3D representations during visual processing in time and space. In this study, we collected fMRI and EEG data from participants viewing 3D stimuli with red-green anaglyph glasses. Participants first completed a behavioral session involving depth judgment tasks and a 3D cube adjustment to quantify and equate across individual differences in depth perception. They then participated in two EEG sessions and two fMRI sessions in which they viewed peripheral stimuli presented across 64 possible 3D locations (4×4×4 grid centered on fixation). With a large number of trials (>7,000 trials per participant) and by integrating EEG, fMRI, and computational methods via representational similarity analysis, we address several previously unexplored questions. We first successfully tracked the spatiotemporal dynamics of neural representations of object spatial information (horizontal, vertical, depth, radius, polar angle, etc.) in various coordinate systems (Cartesian, Cylindrical Polar, and Spherical Polar coordinates), finding that neural representations reflect different types of spatial information across different brain areas and points in time. Second, we revealed an overall preference for representations consistent with a 2D polar coordiante system and a 3D cylindrical coordinate system, though there was evidence for Cartesian preference in some brain areas. Finally, we asked whether there is evidence of 2D and/or 3D spatial integration (integrated representations of space). Partial correlations revealed unique representations of 2D and 3D integrated representations (Euclidean distance), beyond the individual spatial components. These findings provide valuable insights into the neural mechanisms underlying our ability to perceive the world in 3D.

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

33.312 SHAPE PROCESSING ALGORITHMS IN V4 DERIVED FROM NEURAL NETWORK MODELS

Allen M. Chen[1] (allenmuhanchen@gmail.com), Ramanujan Srinath[1], Charles E. Connor[1]; [1]Johns Hopkins University

Studies of neural coding in intermediate and higher level stages of the ventral/object pathway of primate visual cortex have revealed that neurons represent 2D and 3D shape in terms of object fragments, their geometric properties, and their object-relative spatial configuration. However, almost nothing is known about the neural algorithms that generate this information from lower level inputs. Here, we used analysis of visual neural network processing to derive hypotheses about these algorithms in area V4, an intermediate stage where individual neurons encode orientation, curvature, and object-relative positions of 2D contour fragments, 3D surface fragments, and 3D medial axis fragments. Our aim is to test these algorithms with experiments in V4, V2, and V1, to discover how this geometric information about object fragments is generated from the 2D Gabor-like filter signals that tile visual space in area V1. We have analyzed AlexNet, a visual network trained on the ImageNet database. Previous work has shown that V4 responses are most closely modeled by layer 3 of such convolutional vision networks. We analyzed how layer 3 response patterns depend on differential inputs from convolutional neurons in layer 2 (conv2) modulated through, max pooling, rectification, and connection weight patterns to the layer 3 neuron. In turn, we analyzed how the responses of these conv2 neurons depended on differential inputs from layer 1 Gabor-like filters (conv1). One hypothesis emerging from these analyses is that V4 neurons that encode 3D shape fragments should be driven almost entirely by Gabor-like signals for achromatic contrast and low spatial frequencies. In contrast, V4 neurons selective for 2D shape should be driven more strongly by signals for chromatic contrast. Our initial tests of this hypothesis contrast the sensitivity of 3D- vs. 2D-responsive neurons in V4 to achromatic vs. chromatic contrast and low vs. high spatial frequencies.

33.313 A BAYESIAN HIERARCHICAL MODEL FOR RECOVERING 3D NATURAL SHAPES FROM PERSPECTIVE IMAGES

Mark Beers[1]; [1]University of California, Irvine

Mathematically, 3D shape reconstruction from a 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, I showed that subject's perceived 3D shape from a perspective image was often closer to veridical, and never less veridical, when perspective information was more reliable. Here, I elaborate on last year's experiment to test how the human visual system incorporates symmetry and perspective information and propose a Bayesian model that explains how constraints (aka priors) are combined with the visual image. On each trial in the experiment, a static perspective image is shown on the computer monitor, as well as an adjustable rotating 3D shape. The subject adjusted the aspect ratio of the rotating 3D shape until the shape matched the 3D percept produced by the static 2D image. From trial to trial, object shown and simulated distance varied. As the simulated distance of an object to the observer increases, the perspective information becomes less reliable and the subject's percept often becomes less veridical. All shapes used in the experiment were symmetrical or approximately symmetrical, real-world objects selected from the ModelNet-40 dataset. A hierarchical Bayesian model is introduced which performs 3D shape reconstruction. Several versions of the model are compared. The best models include the following shape constraints: compactness of the convex hull of the 3D shape and a second mirror symmetry. In addition, the model uses perspective information to aid in reconstruction but the contribution of perspective information is mediated by a measure of its reliability. The reliability of perspective information is affected by relative angles and lengths of visible symmetry line segments. The model biases the 3D reconstruction towards veridicality when perspective information is reliable.

33.314 A CONVEXITY-BIAS MODEL CAN EXPLAIN NON-RIGID PERCEPTS OF RIGIDLY MOVING STRUCTURE-FROM-MOTION STIMULI

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

In previous work, we demonstrated a violation of the rigidity assumption in Structure-From-Motion (SFM): a rigidly rotating plane, with one concave and one convex part protruding from each of its vertical halves, is perceived as a surface with two convex parts, moving non-rigidly (VSS, 2023, 2024). Under orthographic projection, observers predominantly perceived the non-rigid interpretation, while under perspective projection, non-rigid percepts were less frequent but still significantly above zero. We view these results as illustrating a competition between priors for rigidity and convexity, where convexity typically "wins" even in the presence of perspective cues — leading to non-rigid percepts despite all stimuli being rendered as rigid. We developed a model to construct a 3D shape re-interpretation where both parts are convex. The model "flips" points on the concavity across a mirror, along the ray from each point to the observer, thereby maintaining projective consistency, and transforming it into a convexity. For orthographic and perspective stimuli, the model differs in two important ways to maintain projective consistency: (1) the rays to the observer, reflecting differences in the image formation process; and (2) mirror location, as perspective provides cues to preserve the rigidity of the square plane, such as foreshortening. The model correctly predicts the non-rigid motion percepts that arise when observers perceptually reinterpret concave parts as convex. It also captures the qualitative change in non-rigid percepts between orthographic and perspective projections that we observed in our previous work. On the other hand, this pattern of results could not be explained by a simple motion-based heuristic that uses relative motion to determine the depth of each point. The success of the shape-based model underscores the significant role of the convexity bias in interpreting SFM: a shape prior is capable of predicting non-rigid motion percepts on a rigidly rendered SFM stimulus.

33.315 BILATERALLY SYMMETRIC 3D RECONSTRUCTOR (BIS3D): A COGNITIVELY INSPIRED MODEL OF HUMAN 3D SHAPE PERCEPTION

Doreen Hii[1] (doreen.hii@uci.edu); [1]University of California, Irvine

The first fully autonomous, theory-driven model of 3D object reconstruction from a single 2D orthographic image, is presented. The theory on which my model is based (Pizlo & de Barros, 2021) states that the formation of the mental representation of a 3D shape is possible because the human visual system implements powerful a priori constraints such as symmetry, compactness, and planarity. My model, Bilaterally Symmetric 3D reconstructor (BiS3D), operates in two stages. In the first stage, the model reconstructs a template 3D object through the following steps: (a) extracting pairs of similar 2D contours, (b) reconstructing these pairs in 3D by applying the symmetry prior, (c) propagating the 3D information to all visible 2D edges using planarity of faces, and (d) recovering the back, occluded part of the object by mirror-reflecting one half of the reconstruction with respect to the plane of symmetry. The last step guarantees a volumetric output with all elements of a 3D combinatorial map: vertices, contours, surfaces and volume (Brun & Kropatsch, 2003). In the case of orthographic projection, symmetry and planarity priors only reconstruct a template 3D object without specifying its aspect ratio. Therefore in the second stage, the model optimizes the aspect ratio of the template object by maximizing compactness, producing a unique 3D interpretation. I tested the model in a shape constancy experiment where the task was to respond if two successive presentations of 2D images were from the same 3D polyhedral shape. Shape constancy performance of the model was evaluated by Intersection over Union of the two reconstructed 3D shapes. To demonstrate the effect of the compactness prior, I varied the degree of differences between the slants of the symmetry plane used in the two presentations. The performance of the subjects and the model was systematically affected by this manipulation.

33.316 EFFECTS OF DISTANCE AND GEOMETRIC CONTEXT ON ORIENTATION DISCRIMINATION

Ruth Rosenholtz[1]; [1]NVIDIA Research

Many researchers have studied the human ability to perceive 3D shape and depth from a single 2D picture. How accurate are these judgments, what cues do they rely on, and what ambiguities are inherent to a given cue? Why are people often insensitive to viewing a picture from the wrong center of projection, and when do they perceive distortions? Such 3D perception depends upon lower-level judgments that often involve comparisons across space. For example, utilizing simple perspective cues requires comparing the angle between two separated lines, parallel in the world but converging in the image. If we understood how well humans can make such comparisons, that would inform our understanding of the accuracy and ambiguities inherent in picture perception. Two experiments measured orientation discrimination across space, both for a pair of isolated lines and for two lines forming opposite sides of a quadrilateral. Ten observers were asked to judge which line was tilted more clockwise. The lines had an average length of 2.9 deg, and the spacing between them was 0.95, 1.9, 3.8, or 7.6 deg of visual angle. On each trial, a staircase determined the difference in orientation, with the base orientation either near-horizontal or oblique. Performance degraded significantly as a function of distance between the lines, and was worse when the lines formed part of a quadrilateral than when isolated. In addition, we found a typical oblique effect. In the easiest condition (isolated lines, near horizontal, smallest spacing) thresholds were comparable to prior work (1.2 deg). However, in the hardest condition (quadrilateral, oblique, largest spacing), thresholds rose to approximately 6.9 deg. This may strongly constrain our ability to perceive shape from perspective and notice distortions.

33.317 EXPERIENTIAL3D: FOUR ILLUSIONS CHALLENGE OUR UNDERSTANDING OF 3D VISUAL EXPERIENCE

*Paul Linton[*1] (paul.linton@columbia.edu); [1]Columbia University*
[*]Equal Contribution

Four illusions suggest that many 3D visual processes are cognitive rather than perceptual: 1. LINTON STEREO ILLUSION: A back circle (at 50cm) and a front circle (at 40cm) appear to move rigidly together in depth (whilst their angular size is fixed) when their separation is kept constant in retinal disparity (moving to 40cm and 33cm) rather than physically constant (40cm and 30cm). This suggests that perceived stereo depth reflects retinal disparities, and depth constancy is cognitive rather than perceptual. 2. LINTON SCALE ILLUSION: Increasing interpupillary distance using a telestereoscope makes scenes seem miniature. We develop a VR telestereoscope that decouples increases of vergence and vertical disparities from increases in horizontal disparities. When horizontal disparities are increased, but vergence and vertical disparity are normal, the scene looks miniature. When vergence and vertical disparity are increased, but horizontal disparities are normal, the scene looks normal. This suggests that visual scale relies on a cognitive association between accentuated stereo shape and closer distances. 3. LINTON UN-HOLLOW FACE and LINTON MORPHING FACE ILLUSIONS: Using VR, we argue perceived depth is not inverted in the Hollow-Face illusion: First, objects placed in the hollow of the Hollow-Face illusion can be seen veridically in depth against the Hollow-Face illusion. Second, if we put points on the tip and base of the nose, and gradually switch the position of the two eyes (morphing from hollow to protruding), the change in relative depth between these points is seen veridically. This suggests that depth cue integration is cognitive rather than perceptual. 4. LINTON SIZE CONSTANCY and LINTON SHAPE CONSTANCY ILLUSIONS: Placing a solid rectangular frame around instances of pictorial size constancy, stereo size constancy, and stereo shape constancy all demonstrate that these constancies can be experienced without affecting perceived angular size, suggesting that these constancies are cognitive rather than perceptual.

'New Approach to 3D Vision' Grant from NOMIS Foundation to PL at Italian Academy for Advanced Studies, Columbia University; Presidential Scholars in Society and Neuroscience, Columbia

University; Art, Humanities, and Neuroscience Fellowship, Italian Academy for Advanced Studies, Columbia University

33.318 LARGE-SCALE HORIZONTAL-VERTICAL ILLUSION MEASURED OUTDOORS WITH SUPINE OBSERVERS

Chung Sze Kwok[1], Katelyn Becker[1], Ya Min Phyu[1], Frank H. Durgin[1]; [1]Swarthmore College

Vertical extents typically appear larger than horizontal extents of the same size (horizontal-vertical illusion: HVI). The HVI may be as high as 10% for small objects viewed on a screen. But for very large objects (buildings, light poles), the magnitude of the HVI can be much larger (15-25%). Previous research using virtual environments (VEs) has revealed that medium-sized poles (0.5 -1.5 m tall) elicit a large (15%) HVI when compared to horizontal extents that are in a different depth plane, whilst only eliciting a small (10%) HVI when in the same depth plane. The present study tested whether the effect of depth separation would be observed with real objects outdoors. It additionally tested for the HVI with the objects viewed through a large overhead mirror at 45° while lying supine, using the same effective eye-height as when standing. A total of 48 participants were tested, in both upright and supine postures, using a method of adjustment: They adjusted the frontal ground distance between two balls to match the height of poles (up to 1.3 m tall). The balls were either in the same depth plane as the pole or in a nearer or farther depth plane. When in a different depth plane, matches overestimated the actual vertical by about 15%. When in the same depth plane, overestimation was not greater than 10%. Thus, these results replicate the findings previously observed using VEs, and show that the depth-separation effect can be obtained even with supine observers. An additional angle estimation task, which directly followed the HVI task, found that explicit estimates of azimuthal angular direction were more exaggerated when participants were supine than when they were standing. This latter finding is consistent with previous observations in VEs, and has now also been replicated in supine observers viewing the real world.

33.319 THE CONTRIBUTION OF BINOCULAR DEPTH INFORMATION TO THE PERCEIVED SIZE OF 3D SHAPES

Iroshini Gunasekera[1,2] (iroshini@yorku.ca), Romina Abadi[1,3], Faruq Afolabi[1,3], Xue Teng[1,3], Robert S. Allison[1,3], Laurie M. Wilcox[1,2]; [1]Centre for Vision Research, York University, [2]Department of Psychology, York University, [3]Department of Electrical Engineering and Computer Science, York University

Studies of size perception have tended to focus on the effect of distance and used simple 2D stimuli. Factors affecting the perceived size of volumetric 3D shapes have received less attention. Here, we evaluate the contribution of monocular and binocular depth information to size perception, during passive viewing and active interaction with the stimuli. Using virtual reality (VR) we presented a virtual shape-posting toy. Targets were 3D shapes (triangle, pentagon, square, and quatrefoil) and on each trial, one target appeared alongside a box with identically shaped slots. Participants indicated if the target was larger or smaller than the corresponding slot in a 2AFC task. Shape size varied according to the method of constants; the psychometric data were fit with cumulative normal distributions to compute JNDs and PSEs. We assessed the effects of stereopsis and object motion in two experiments. In Experiment 1 (N=28) the shapes were viewed monocularly or binocularly and were stationary or could be picked up and rotated. In Experiment 2 (N=26) we evaluated the effect of two types of motion (passive rotation vs active movement) and monocular vs. binocular viewing. In Experiment 2, positional data for both head movements and object trajectories were collected. We found that observers discriminated size accurately across all test conditions; PSEs showed no consistent bias. Overall, binocular judgements were more precise (smaller JNDs) than monocular judgements. Surprisingly, in the monocular conditions, discrimination performance worsened when participants interacted with the object. Analysis of target positions during trials showed that observers did not adopt different strategies depending on the depth cues available (e.g. they did not bring monocularly viewed objects closer). Our results underscore the importance of binocular depth information in perception of 3D object size, even when motion or active interaction could theoretically enhance size judgements.

This research was supported by grants RGPIN-2019-06694 to LW and ALLRP 570802–21 to RA and LW from the Natural Science and Engineering Council (Canada), and CF-REF VISTA funding to I. Gunasekera.

Scene Perception: Ensemble

Sunday, May 18, 8:30 am – 12:30 pm, Banyan Breezeway

33.320 SPATIAL WEIGHTING OF ORIENTATION ENSEMBLES IS SHIFTED TOWARDS HIGHLY CORRELATED REGIONS

Richard Bailey[1], Jefferson Ortega[1], Andrey Chetverikov[2], David Whitney[1]; [1]University of California, Berkeley, [2]University of Bergen

Observers can easily report the average orientation of an array of oriented lines with a randomly generated mean orientation. Previous papers showed that one can calculate a spatially weighted map from these ensemble judgments, which reveals specific elements or regions in the display that matter most for ensemble orientation perception (Tiurina et al., 2024). In the experiments here, we used this approach to test how correlations in the visual environment affect ensemble perception. Observers judged the mean orientation of an ensemble of lines. On a given trial, the orientations were drawn from the same underlying distribution with a common mean, and the only manipulation was whether the orientations were more correlated in one half of the display (left or right). In Experiment 1, the correlated side of the display was held constant on one side of the screen on every trial. Our results showed that the spatially weighted maps shifted toward the correlated side. This hints that correlated information in the world might modulate ensemble perception. One possibility, though, is that when there is more correlation there is also lower variance in the display. So, observers might use variance as a proxy for correlation, and the spatially weighted maps may simply be biased toward lower variance in the display. To address this, in Experiment 2, the correlated side was randomized in either left or right visual field, from trial to trial, unpredictably. Interestingly, we found that the spatially weighted maps did not shift toward the lower variance (correlated) side. This suggests

it is not a trial-by-trial shift in spatially weighted maps toward lower variance in the display. Instead, it seems to be a spatially weighted bias toward higher correlations in the environment, which builds up over time.

"Supported in part by the National Institute of Health (grant no. R01CA236793) to D.W and (grant no. 1F99NS141343) to J.O."

33.321 ASYMMETRIC VARIANCE IN ENSEMBLE ORIENTATIONS MODULATES SERIAL DEPENDENCE

Theo Schmit[1], Jefferson Ortega[1], Andrey Chetverikov[2], David Whitney[1]; [1]University of California, Berkeley, [2]University of Bergen

Serial dependence stabilizes our perceptual experience by biasing our current perception toward previously seen stimuli. While this effect is well documented for simple features, less is known about how variance in ensemble presentations influences serial dependence. We investigated this question by manipulating stimuli variance in an ensemble orientation judgment task. Participants were presented with 36 oriented lines and tasked with reporting the overall, average orientation of all lines. There were two types of trials in our experiment. In asymmetric variance trials, 16 orientations on the left side of the screen were more correlated to each other than the 16 orientations on the right side of the screen, leading to an asymmetry in orientation variance across the screen. In symmetric variance trials, stimuli variance was uniform across the entire screen. Participants completed 800 trials (400 for each trial type) and trial type was interleaved throughout the experiment such that asymmetric variance trials always followed symmetric variance trials and vice-versa. For asymmetric variance trials, we found that observers made larger overall errors and they exhibited a significant serial dependence effect, suggesting that ensemble judgments were biased towards their perception of the previous trial orientation. We found significantly less serial dependence and more accurate reports for symmetric variance trials. The results suggest that when performance levels are low, potentially indicating higher uncertainty of the estimated ensemble, observers rely more heavily on past experiences to inform their current judgments.

Supported in part by the National Institute of Health (grant no. R01CA236793) to D.W and (grant no. 1F99NS141343) to J.O.

33.322 HIDDEN MARKOV MODELS REVEAL ATTRACTIVE AND REPULSIVE PERCEPTUAL MODES DURING AN ENSEMBLE JUDGMENT TASK

Katrina A. Wolters[1], Veith Weilnhammer[1], Jefferson Ortega[1], Andrey Chetverikov[2], David Whitney[1]; [1]University of California, Berkeley, [2]University of Bergen

Cognitive modes are distinct patterns of conscious attention and mental engagement that shape how individuals perceive their external environment. Recent evidence suggests that human perception alternates between two distinct modes: an internal mode, where perception is biased toward previously seen stimuli, and an external mode, where perception is aligned more closely to the external world (Weilnhammer et al., 2023). Hidden Markov models have previously been used to identify these dynamic modes of perception during orientation judgment tasks, and have found that they can identify states of attractive, dynamic templates (i.e. serial dependence; Weilnhammer et al., 2024). In the current study, we aimed to investigate whether similar dynamic states can be identified while observers perform an ensemble judgment task. Additionally, we expanded upon previous studies by using a novel modeling approach by fitting hidden Markov models to metathetic (circular) response data. Observers completed 800 trials of an ensemble judgment task, where 36 orientations were presented on the screen for 1000ms, followed by a 500ms delay and a response period. During the response period, participants were instructed to reproduce the average orientation of the 36 presented lines. Interestingly, our analyses revealed contrasting patterns of perceptual biases. We found an internal mode-associated state that exhibited a serial dependence effect, in which observers' ensemble judgments were attracted toward the previously presented stimulus. We also found an external mode-associated state exhibiting a repulsive bias away from the previously presented stimulus. These results indicate that distinct patterns of perceptual bias (either attractive or repulsive) can arise during cognitive modes.

Supported in part by the National Institute of Health (grant no. R01CA236793) to D.W and (grant no. 1F99NS141343) to J.O.

33.323 EXTRACTION OF AVERAGE TEMPERATURE FROM VISUAL SCENE ENSEMBLES WITHOUT HIGH SPATIAL FREQUENCIES

Vignash Tharmaratnam[1], Dirk B. Walther[2], Jonathan S. Cant[1]; [1]University of Toronto Scarborough, [2]University of Toronto

We have previously demonstrated that participants can rapidly extract average scene temperature (i.e., how hot or cold scenes would feel on average; VSS, 2024), without reliance on color, contrast, or low spatial frequencies. Furthermore, this occurred without utilizing visual working memory (VWM) resources, a hallmark of ensemble processing. In the present study, we furthered this investigation by examining whether average scene temperature could be extracted without reliance on high spatial frequency content. Given the established importance of low spatial frequencies in the rapid formation of scene gist representations (Oliva, 2005), we predicted that average scene temperature could be extracted when high spatial-frequency information was filtered out of scene images, and, similar to our previous results, this would occur without reliance on VWM. Participants rated the average temperature of scene ensembles that were gray-scaled and had a low spatial frequency filter applied (< 1 cycle/degree). 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, with all 6 scenes being integrated into their summary statistics. This occurred without relying on VWM, as fewer than 0.9 scenes were remembered on average. These results reveal that computing cross-modal summary statistics (i.e., average temperature) does not rely on high spatial frequency information or VWM resources, and that abstract multisensory information can be rapidly retrieved from complex visual stimuli.

33.324 SET SIZE AND ENCODING TIME INFLUENCES INDIVIDUAL OBJECT AND ENSEMBLE PERCEPTION IN NATURALISTIC SCENES

Yanina Tena Garcia[1], Bianca Baltaretu[1], Dominik Endres[2], Katja Fiehler[1]; [1]Justus-Liebig-Universität Gießen, [2]Philipps-Universität Marburg

Individual and ensemble perception are essential processes for effectively interacting with objects in our environment. Individual object perception focuses on the individuation and identification of single objects, while ensemble perception extracts summary information from groups of objects. The underlying mechanisms of these processes remain an active area of investigation, particularly regarding how they are influenced by the number of target objects and the duration of their presentation. Although previous studies have predominantly used simplistic stimuli, it is unclear how these findings extend to naturalistic settings. To address this question, we conducted a computer-based experiment in which participants were tasked with remembering and indicating either the position of a single object or the average position of all presented target objects within a naturalistic scene. We examined differences in individual versus ensemble perception across two key variables: scene set size (3, 6, or 10 objects) and encoding time (50, 100, or 800 ms). Our results show that encoding time had a similar effect on locating accuracy for both processes, with accuracy improving as encoding time increased. However, the set size manipulation had distinct effects: Individual perception showed higher locating accuracy for smaller set sizes, especially for longer encoding times. In contrast, ensemble perception was associated with reduced accuracy for smaller set sizes, especially at the shortest encoding time (50 ms). These findings suggest that individual and ensemble perception are differentially influenced by the amount of task-relevant information, highlighting the complementarity of their roles for real-world actions. We are currently applying two Bayesian models to explore whether individual and ensemble perception operate independently or are interdependent, as a way of gaining deeper insight into their underlying relationship.

German Research Foundation (DFG), project number 222641018 – SFB/TRR 135 TP A4

33.325 ENSEMBLE PERCEPTION AND LOC: A FUNCTIONAL MRI STUDY

Noam Khayat[1] (noamkhayat@gmail.com), Shaul Hochstein[1]; [1]Hebrew University of Jerualem, [2]ELSC Edmond & Lily Safra Center for Brain Sciences, [3]Life Sciences Inst

Ensemble perception refers to the visual system's ability to extract summary statistical information from groups of similar objects, commonly assessed through the efficiency of perceptual averaging. While this phenomenon has been extensively studied and characterized behaviorally, its underlying neural mechanisms remain largely unexplored. Using functional magnetic resonance imaging (fMRI) and Multi-Voxel Pattern Analysis (MVPA), we investigated how the neural representation of ensemble averaging—specifically for line orientations—can be decoded across cortical regions. Our experimental paradigm was designed to identify regions that encode explicit ensemble average information, with participants instructed to actively perceive and passively report the average orientation in a subsequent two-alternative forced choice (2-AFC) task. While some studies have suggested that the Parahippocampal Place Area (PPA) may specialize in ensemble perception and others have proposed that the neural representation of ensemble percepts is formed by pooling signals across multiple levels of the visual processing stream, our findings reveal that the Lateral Occipital Complex (LOC) also encodes ensemble average information, as evidenced by significantly above-chance classification accuracies. This result is novel and surprising, as the LOC is traditionally associated with object recognition and perception, independent of low-level visual features. We demonstrate that, beyond its well-established role in recognizing individual objects, the LOC also processes information of groups of objects perceived as unified wholes.

Israel Science Foundation (ISF)

33.326 VISUAL STRATEGIES FOR TREND DETECTION: COMPARING THE EFFECTIVENESS OF SINGLE-HUE AND MULTI-HUE COLOR PALETTES ACROSS CONTEXTS

Amelia C. Warden[1] (acwarden@umich.edu), Jessica K. Witt[2]; [1]University of Michigan, [2]Colorado State University

Effectively conveying time-series data is critical for decision-making in many domains, like climate science, economic forecasting, and epidemiology. Oftentimes, designers opt to use colors that are semantically compatible with the underlying data, such as red for higher and blue for cooler temperatures. An alternative method is to use colors that better exploit ensemble processes, which refer to the visual system's innate ability to extract summary statistics, like the mean, from a set of similar objects. Our prior work examining trend detection in visualizations found that single-hue color palettes, which engage visual system processes, better convey temperature trends than semantically compatible color palettes. While assessing how color-coded visualizations impact trend detection in a politically and emotionally significant context is important, this context carries inherent biases that may influence perceptions and interpretations. To further explore the efficacy of color palettes that engage the visual system, the current work examines trend detection for data representing a more impartial context with fewer inherent biases, specifically red and blue car sales over time. Participants viewed stripplot graphs and indicated whether trends were increasing when the underlying data was encoded with either a single-hue color palette exploiting ensemble processes or a semantically intuitive multi-hue color palette. Using signal detection theory, the results revealed significantly higher sensitivity (d') to trends presented with a single-hue color palette than a more semantically compatible multi-hue color palette. These findings highlight the generalizability of using color palettes that better engage properties of the visual system, suggesting they improve trend detection independent of the underlying context of the data. Additionally, these findings further demonstrate the advantages of single-hue displays in enhancing ensemble perception. The results have implications for broader applications of ensemble-based color schemes used in information visualizations, which can improve public comprehension and decision-making when adopted for time-series data.

Visual Search: Eye movements, scenes, real-world stimuli

33.327 FROM SALIENCE TO MEANING AND SCENE GRAMMAR: PREDICTING VISUAL SEARCH EFFICIENCY IN NATURALISTIC SCENES

Antje Nuthmann[1], Anton Janser[1]; [1]Kiel University

Efficiently locating objects within visual scenes is crucial for everyday behavior and involves eye movements to direct our attention. We report results from a large-scale project employing a quasi-experimental approach to investigate scene guidance during visual search in naturalistic environments. Using 170 real-world scene images, each featuring a single target object, we examined how naturally varying object and scene properties influence temporal measures of search efficiency. Object-based predictors included the object's distance from the center of the scene, its size, and its visual salience, derived from saliency map computations. To gauge the spatial distribution of meaning within the scenes, a meaning map was generated for each scene image, aggregating crowd-sourced responses. Object meaning was represented by the mean value over the search object, while the mean value across all pixels in the map served as a proxy of the information density in the scene. The object's relationship to the scene was captured through human ratings of semantic fit, reflecting the likelihood of encountering the object in the scene, and syntactic fit, reflecting its positional plausibility. Central to the study was an eye-tracking experiment in which over 50 observers located the target object in each scene by directing their gaze to it. The main dependent variable was the latency to first fixation on the target, measuring how efficiently attention was guided toward it. Linear mixed-model analyses revealed independent effects of object size, object meaning, and syntax, indicating shorter latencies for larger targets, higher-meaning targets, and targets appearing at more plausible locations. Moreover, latencies increased as the information density increased. First-pass gaze duration for the target object showed opposing influences of object salience and meaning: it was longer for more salient objects but shorter for those with higher meaning. We discuss our findings in the context of existing research using experimental manipulations.

33.328 SHAPED BY MEANING: TOP-DOWN FACILITATION OF SEMANTIC COGNITION IN TOOL VISUAL EXPLORATION

Luigi Valio[1] (luigi.valio@studenti.unisob.na.it), Gaia Diglio[1], Antimo Buonocore[1], Maria Antonella Brandimonte[1], Giovanni Federico[1]; [1]Università degli Studi Suor Orsola Benincasa

Tool use is fundamental to human interaction with the environment. Emerging evidence highlights the role of semantic cognition—the ability to represent, understand, and apply prior knowledge—in complementing sensorimotor and mechanical processing of tools. While semantic cognition is known to guide visual exploration by aiding in the inference of tool identity, purpose, and functionality, the precise cognitive mechanisms and contextual influences remain poorly understood. Here, we investigated how semantically congruent, incongruent, and neutral contexts modulate the temporal allocation of visuospatial attention during tool exploration. Tools were displayed against three types of backgrounds: neutral (e.g., a tool on a plain grey table), congruent (e.g., a whisk in a kitchen), and incongruent (e.g., a whisk in a recording studio). Twenty-eight participants (female/male: 20/8; mean age = 21.04 ± 3.08 years) viewed each scene for six seconds, following a 0.5-second fixation cross and preceding a 4-second blank screen. Eye movements were recorded to examine fixations on functional (goal-related) and manipulative (grip-related) tool components. Participants predominantly fixated tools on their functional areas (e.g., the head of a whisk), particularly during the initial 1500 milliseconds of visual exploration, with fixation patterns strongly influenced by contextual information. A repeated-measures ANOVA revealed significant effects of the visual-perceptual context: in neutral and incongruent conditions, participants directed attention toward functional parts, likely to resolve perceptual and conceptual ambiguities. Conversely, in congruent conditions, the alignment between tool and background facilitated a more balanced exploration, with a preliminary trend toward increased fixations on manipulative areas, thus suggesting a context-driven readiness for action. These preliminary findings underscore the pervasive role of semantic cognition in guiding visuospatial attention and offer new insights into how top-down processes adapt dynamically to environmental demands, shaping human-tool interactions across a variety of contexts and scenarios.

33.329 THE MORE YOU KNOW, THE SLOWER YOU SAY NO: BROAD AND NARROW TEMPLATES IN VISUAL SEARCH FOR REAL-WORLD OBJECTS

Susan Ajith[1,3] (susan.ajith@psychiat.med.uni-giessen.de), Daniel Kaiser[1,2], Lu-Chun Yeh[1]; [1]Justus Liebig University Gießen, [2]Center for Mind, Brain and Behavior (CMBB), Philipps University Marburg, Justus Liebig University Gießen, and Technical University Darmstadt, [3]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig

When searching for an object, search templates derived from prior knowledge guide attention towards likely targets. Objects with consistent visual appearance, such as basketballs, are likely to activate narrow search templates. In contrast, objects with high variability, like bags (which feature various shapes and colors), are likely to activate broad templates that account for their real-world diversity. To investigate how such narrow or broad search templates affect search for everyday objects, we conducted a visual search study, where participants viewed a word cue followed by a search display that did or did not contain the target. Targets were selected from objects spanning a range of variability in appearances, from small (e.g., basketball or tennis racket) to large (e.g., bag or lamp) variations. To quantify the variability in participants' mental representation of the objects, participants additionally performed a drawing task, where, before the search experiment, they drew four exemplars for each of the target objects. By assessing dissimilarities among the drawn exemplars, this task allowed us to gauge the variability within each object. Results showed that response times in the search task varied systematically with the variability in the drawings: When participants produced less variable exemplars during drawing (evaluated using a deep neural network model), they were faster during search, while more variable drawings led to slower search. This was true for target-present and target-absent trials. Crucially, differences in search performance linked to template variability were significant even when only the initial trials per object were considered. This suggests that

real-world object variability impacted search from the outset, before participants could estimate the variability of the target objects in the experiment. Together, our findings demonstrate that search templates for real-world objects are inherently shaped by the variability of an object in the world.

SA is funded by the JLU graduate scholarship. LCY is supported by the MSCA programme (101149060). DK is supported by the DFG (SFB/TRR135,222641018; KA4683/5-1, 518483074, KA4683/6-1, 536053998), “The Adaptive Mind”, and an ERC Starting Grant (PEP, ERC-2022-STG 101076057).

33.330 THE ROLE OF SCENE CONTEXT IN THE GUIDANCE OF ATTENTION BASED ON OBJECT-LOCATION ASSOCIATIONS

Justin Frandsen[1], Brian Anderson[1]; [1]Texas A&M University

Relationships between objects and their locations (e.g., clock on the wall, pot near the stove) can be used to guide attention in visual search through naturalistic scenes. Due to these associations being formed throughout the lifespan, the mechanisms that underlie the development of such learning-dependent attentional guidance are unclear. By pairing arbitrary stimuli, for which participants lack semantic knowledge, with locations in scenes, unique insights can be gained concerning the role of experience in scene-based attentional guidance. Previous work found that associations between arbitrary shapes (objects without semantic information) and location could be formed via statistical learning mechanisms. This was done by associating a given shape with a region of the scene (i.e., wall, counter, floor) where it would most likely appear over trials. It was found that search was faster when the target appeared in the high-probability location relative to a low-probability location. In the current study, we tested whether this learning was context-dependent. We used three scene categories (kitchen, bathroom, living room) and three arbitrary shapes with different location associations in each scene category. Specifically, each shape was associated with a different region (wall, counter, floor) in each category of scene, with 75% contingency. While eye position was tracked, participants were cued to search for one of the three target shapes in a subsequently presented scene, where they would respond via keypress to the orientation of a small T within the shape. Results showed no effect of when the target appeared in the associated high-probability location relative to an unassociated low-probability location. This suggests that the learning of the object-location relationships that guide attention is not dependent upon context and rather occurs in a context-general manner across categories of scenes.

33.331 UTILIZING GENERATIVE AI FOR SCENE SEARCH: ENHANCING FLEXIBILITY AND CONTROL IN VISUAL SEARCH EXPERIMENTS

Kerri Walter[1], Peter Bex[1]; [1]Northeastern University

Traditional scene search tasks require finding or creating images that satisfy the demands of the search task, which limits generalizability and adaptive designs. Generative AI enables the potential to provide control over scene generation. In this proof of concept study, we demonstrate how generative AI can be utilized to quickly and efficiently generate experimentally-controlled scenes with search targets specific for computer-adaptive prompts. We utilized Bing Image Creator to generate a series of 40 images from 10 probabilistic prompts, given 5 unique scene types, control over the amount of visual clutter, balanced across conditions. New scenes were generated concurrently during the experiment for effective real-time unique image generation for each participant. Participants were given a target (presented as text) and subsequently searched the scenes for the given target and clicked on it when located, or off image if not located. The AI image generator is not perfectly consistent in presenting all objects listed in a prompt, consequently, if a target was missing from the scene it was coded as a target absent trial. We measured reaction time and the number of fixations (GazePoint 60Hz eye tracker) in each trial. Using a mixed effects model, we found that this method replicates traditional visual search findings: such that high clutter scenes yielded significantly longer reaction times ($b=1.045$, $SE=0.452$, $z=2.311$, $p=.021$) and more fixations ($b=5.684$, $SE=2.080$, $z=2.733$, $p=.006$) than low clutter scenes (set size effect), and that target present searches yielded significantly shorter reaction times ($b=-2.288$, $SE=0.365$, $z=-6.263$, $p<.001$) and fewer fixations ($b=-13.708$, $SE=1.681$, $z=-8.152$, $p<.001$) than target absent searches (target present/absent effect). These results demonstrate that a generative AI method can be utilized for advanced visual search tasks in naturalistic scenes, providing increased control and flexibility over traditional paradigms.

33.332 EXPLORING THE DIFFERENT ROLES OF FIXATIONS IN AN ACTIVE VISUAL SEARCH TASK

Tiffany C. Wu[1], John K. Tsotsos[1]; [1]York University

Common visual search paradigms conducted on 2D screens with passive observation do not capture the full breadth and reality of eye and head movements used in real-world search. One is not presented with an image in real-world search; one must determine which images to acquire and in what order using relevant eye, head, and body movements. To investigate viewpoint selection and the role of fixation in active observation, an active visual search task was conducted in a controlled real-world environment. The scene was a physical 3x4m space furnished with tables and wire cages. Stimuli were miniature everyday objects, scattered in various orientations on the tables and cages. Observers moved freely, untethered, to search for a target object, and their eye and head movements, reaction time, and accuracy, were synchronized and measured over 12 trials each. Resulting eye and head movement data naturally seemed divided into “environment”, “look-at”, and “target look-at” fixations. “Look-at” refers to fixations viewing tables or cages with stimuli in view, “target look-at” refers to fixations viewing the target object, and “environment” covers all other fixations. Interestingly, subjects became more efficient at searching with successive target present trials, particularly in the number of look-at fixations. Target look-at fixations were also significantly longer than other fixations. Finally, we discovered that environment fixations often occur between look-at’s while a subject is navigating to a different location to continue their search. This suggests a clear distinction in the role between look-at fixations and environment fixations - one for searching through stimuli, and one for searching and navigating through the environment to achieve the next viewpoint. These results emphasize the importance of conducting search and other visual tasks in the real world, in order to capture the nuances of eye and head movement and strategies not otherwise found from a 2D paradigm.

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).

33.333 DIG A LITTLE DEEPER: COMPARING 2D TO 3D SEARCH PERFORMANCE WHEN LOOKING FOR MULTIPLE BREAST CANCERS

Lyndon Rakusen[1] (rakusenl@arizona.edu), Brandon Eich[1], Stephen Adamo[1]; [1]University of Arizona

For over 60 years, radiologists have known they are more likely to miss a second abnormality after a first is detected in the same image. This Subsequent Search Miss (SSM) effect (also known as Satisfaction of Search in radiology) persists across many imaging modalities and has more recently been shown to cause misses in breast cancer detection with radiologists and novice observers. Because breast cancer accounts for 31% of all new cancer diagnoses and 10 to 30% of breast cancers are not reported, efforts to improve detection are critical. 3D breast imaging (tomosynthesis) has yielded better detection of cancer and fewer false alarms than a standard mammogram (a 2D representation of breast tissue), as radiologists can search in-depth without normal tissue obstructing their view. However, it is unclear how search performance for multiple cancers is affected when searching with tomosynthesis, particularly the extent to which SSM errors persist compared to mammography. In the current study, novice observers were asked to detect up to two masses within synthesized virtual tomosynthesis and mammography images. These images were created using VICTRE, a fully in-silico tool for generating realistic-looking breast images. By using VICTRE to generate mammography and tomosynthesis images, we could control the mass locations between image types, allowing us to compare SSM errors between modalities. Overall, observers: 1) were more accurate in cancer detection in tomosynthesis, 2) were more accurate in finding a second target in tomosynthesis, and 3) had a significant SSM effect in both, with no differences between the size of their SSM effect. Together, these results suggest that tomosynthesis should be used more frequently when attempting to detect breast cancer as it can improve detection when there are standalone and multiple masses. However, tomosynthesis may not alleviate the SSM effect compared to mammography.

Grant from the National Institutes of Health, National Cancer Institute (4R00CA267163-03)

33.334 CUEING DIRECTS ATTENTION TOWARDS VISUAL SEARCH ITEMS, RATHER THAN ENHANCING THEIR PROCESSING

Johan Hulleman[1], Aoqi Li[1], Jeremy M. Wolfe[2,3]; [1]University of Manchester, [2]Brigham and Women's Hospital, [3]Harvard Medical School

In two eye tracking experiments, 20 participants searched for T amongst L's on a 1/f^1.3 noise background (target prevalence 50%). The contrast of an item was defined as the difference between its grayscale [0-255] and average local background grayscale [75-180]. Absolute target contrast was 15, 45, 75, 105 in Experiment 1 and 15, 45, 75 in Experiment 2. Distractor contrast varied freely. We studied the benefits of cueing all item locations with yellow boxes drawn around each item. In both experiments, 24-item search displays were presented twice. Experiment 1 had 384 unique displays. Half of the displays were first presented cued and then uncued (Cue–NoCue), for the other half this order was reversed (NoCue–Cue). Experiment 2 had 288 unique displays. Half the displays were presented NoCue–Cue, the other half as NoCue–NoCue. Both experiments found the same result for present trials: cueing reduced the error rates for the lowest contrast targets (15), without increasing RTs. In Experiment 1, cueing did not change RTs or accuracy for higher contrast targets (45, 75, 105). In Experiment 2, cueing did not improve accuracy for higher contrasts (45, 75) either, but RTs slowed down. For the absent trials both experiments did not find any effect of cueing on either accuracy or RTs. Looking at eye movement data, we can characterize errors as "search" errors where the observer never fixates the target and "recognition" errors where the target is fixated but the eyes move away without registering it successfully. Cueing all potential targets reduced misses by reducing "search" errors in both experiments. Specifically, low contrast targets were more likely to be visited. We conclude that cueing directs attention to items, rather than enhancing their processing once attention arrives. Encouragingly, the accuracy gain from cueing does not necessarily come with an RT cost.

JH and AL were supported by UKRI grant ES/X000443/1. JMW was supported by NIH-NEI: EY017001, NSF: 2146617, and NIH-NCI: CA207490

33.335 DYNAMIC ADJUSTMENT OF FIXATION DURATIONS IN VISUAL SEARCH.

M Pilar Aivar[1], Laura Cepero[1], Miguel A Vadillo[1], Victoria Plaza[1]; [1]Universidad Autónoma de Madrid

The proposal of the functional viewing field (FVF) has highlighted the relevance of studying eye movements to better characterize visual search processes (Hulleman & Olivers, 2017; Wolfe, 2021). Still, it is often assumed that the FVF has both a constant size and a constant processing time (Shi, Zang & Geyer, 2017). Very few studies have investigated whether fixation durations vary during visual search depending on the repetition of the display or the kind of element that is fixated. Here, we present the results of three experiments in which we recorded eye movements while participants performed different types of visual search tasks. In all experiments some aspects of the visual search display were repeated over trials while other aspects were irrelevant to the task. In the first study we used a comparative visual search paradigm: participants had to find the only item that differed between both sides of the screen. Display configurations were repeated during the experiment. In the second study we employed a repeated search paradigm, presenting different targets but the same search display in all trials. In the third study we employed a contextual cueing paradigm, manipulating the number of relevant and irrelevant distractors. Half of the configurations were repeated, intermixed with newly generated configurations. Eye movements in these three tasks showed interesting patterns. In all experiments the number of fixations needed to find the target decreased significantly over repetitions. Interestingly, fixation durations differed depending on whether the fixation was on the target item or on other elements. Moreover, display repetition had an effect on fixation durations: search fixations increased duration while target fixations decreased duration over repetitions. These results suggest a dynamic adjustment of fixation duration which might be related to the size of the FVF and provide some insights regarding oculomotor scanning strategies during search.

Research supported by grants PID2020-118583GB-I00 and PID2021-125162NB-I00.

33.336 EXPLORING THE IMPACT OF REWARDS ON OCULOMOTOR BEHAVIOR IN VISUAL SEARCH TASKS

Orit Shdeour[1] (oritshdeour@gmail.com), Shlomit Yuval-Greenberg[1]; [1]Tel Aviv University

When searching for a specific target in a crowded visual environment, we continuously move our eyes across the image to locate the target as efficiently as possible. This exploratory process involves fixating on points of interest, processing foveal information, and utilizing peripheral input to guide subsequent fixations. While the spatial aspects of visual exploration are known to be influenced by both bottom-up and top-down processes, the temporal dynamics of eye movements during search remain less understood. Specifically, the oculomotor characteristics of search and their modulation by top-down mechanisms require further investigation. In a series of experiments, we characterized the modulation of oculomotor search by reward. In Experiment 1, participants performed a complex “Where’s Waldo?” search task. Two groups (N = 20 each) were tested: one group was offered a reward for speed, while the other was not. The reward group exhibited shorter fixation durations, more saccades per second, larger saccade amplitudes, and lower spatial entropy of gaze positions. These findings suggest faster, broader, and more efficient search behavior under reward conditions. In Experiment 2 (N = 18), using a within-subject design, participants were rewarded in some trials but not in others. Contrary to Experiment 1, no significant differences were observed in behavioral or eye-tracking metrics between conditions. This suggests that the reward effect on search performance is driven by global arousal rather than localized attentional processes. A third experiment, currently underway, examines the temporal dynamics of search within a controlled visual environment using simple stimuli. Preliminary findings indicate that the oculomotor characteristics of visual search are not solely driven by visual features but are substantially influenced by top-down processes and the observer's internal state. We conclude that arousal, associated with reward, modulates the temporal and spatial dynamics of saccades during visual search.

33.337 FORAGING ON A GRID: THE IMPORTANCE OF STIMULUS LAYOUT IN VISUAL SEARCH BEHAVIOR

Anna Hughes[1] (anna.hughes@essex.ac.uk), Manjiri Bhat[1], Alasdair Clarke[1]; [1]University of Essex

The visual foraging paradigm, where participants must search for multiple target exemplars in a display, allows us to study how human search behavior evolves in space and time. Previous work has shown that target properties, such as target complexity or salience, can affect foraging behavior in these tasks: we have also shown that spatial factors, such as target proximity, are important in accurately predicting which target a participant will select next. Here, we present a novel foraging experiment that explores how the layout of the foraging items influences the spatial strategies that participants use while foraging. We find the when items are placed on a cardinal grid, the foraging sequences and eye movement patterns generally follow the grid orientation. Rotating the grid to an oblique orientation leads to a preference for the oblique directions. However, the overall direction bias is stronger when the grid has a cardinal direction, suggesting that participants have an underlying preference for cardinal directions that goes beyond what is present in the stimulus. We develop a von Mises mixture model to account for these patterns of inter-item target selections. Finally, we incorporate this as a new component in our generative Bayesian model of foraging behavior (FoMo) and discuss our train-test split approach for assessing the accuracy of the model, and for determining the relative importance of different model parameters.

Undergraduate Just-In-Time 1

SUNDAY, MAY 18, 8:30 AM – 12:30 PM, BANYAN BREEZEWAY

33.338 EFFECTS OF EXPRESSION AND SEX ON THE PRIORS FOR NATURAL FACE COLORS

Camile A. Franke[1] (camilefranke@gmail.com), Jesse R. Macyczko[1], Michael A. Webster[1]; [1]University of Nevada, Reno

Variations in facial color have been associated with differences in both biological sex and facial expression. We examined priors for the chromaticity of different face categories by measuring the chromatic range that appeared natural. Stimuli were images from the Chicago Face Database and included 3 females and 3 males (white race) modeling neutral, angry, or happy expressions. The faces were cropped to remove external features and adjusted so that all had the same mean chromaticity (for which ground truth was not known). The mean chromaticity of each face was then further varied within a version of DKL color space by rotating the hue angle (+/- 32 deg in 4 deg steps) or scaling the saturation (0 to 3x the original chromatic contrast in 0.2 steps), forming 33 stimuli for each individual and expression. These were shown in random order on a display with a visual angle width of 10 deg, remaining on screen until participants gave their responses. Ten observers (7 female, 3 male) rated whether each face color appeared “natural” or “unnatural,” with the mean of the natural gamut estimated for each face and observer. Results were assessed with separate 2-way ANOVAs (sex by expression) for hue and saturation. Hue gamuts showed weak but significant effects for both factors: averages for male and (surprisingly) happy faces were more red. Male gamuts also had significantly higher saturation than female, with no effect of expression. Our results suggest that observers have small but significant differences in the priors for chromaticity for different categories of faces.

33.339 LESS VISUAL EXPLORATION IN CURVY VERSUS ANGULAR INDOOR SCENES

Xiaojing Zhu[1], Dirk B. Walther[1], Claudia Damiano[1]; [1]University of Toronto

Humans can quickly judge simple stimuli as curvy vs. angular, and tend to find curved contours, lines, and shapes more pleasant. This effect extends to architecture photographs, where spaces with curvilinear designs are perceived as more aesthetically pleasing than rectilinear ones, suggesting that curvature information can be extracted from 2D images of 3D spaces. However, image-computable measures of curvature do not correspond to subjective ratings of curvature for complex real-world scenes. We conducted an eye-tracking study using a 2 (space: curved vs. angular) × 2 (furniture: curved vs. angular) design to determine which visual information people use when making curvature judgments of real-world scenes. Each participant (N=25) viewed 100 indoor images while their eye movements were tracked, and they rated the curvature of scenes on a 5-point Likert scale. We found that participants rated scenes with curvy backgrounds ($\beta = 1.45$, $p < 0.001$) and foregrounds ($\beta = 1.25$, $p < 0.001$) as more curved than angular ones. A wider distribution of fixations, larger saccade amplitudes, and higher saccade velocities were each linked to a small but significant decrease in perceived scene curvature (all $p < 0.05$). Moreover, curvy backgrounds were associated with lower average saccade amplitude ($\beta = -0.46$), saccade velocity ($\beta = -6.62$) and fixation spread ($\beta = -13.92$), while curvy foregrounds were also linked to smaller amplitude ($\beta = -0.23$), velocity ($\beta = -3.15$) and fixation spread ($\beta = -6.54$; all $p < 0.005$). In summary, our findings indicate that the presence of curviness in both the background and foreground is linked with higher levels of curvature perception and reduced visual exploration, with the background information playing a more important role in curvature judgments. This is the first step in understanding how humans judge curvature from 3D spaces.

33.340 PERCEIVING BEAUTY AND THREAT: THE ROLE OF SYMMETRY AND COMPLEXITY IN PREDICTABILITY, MEMORABILITY, AND VISUAL AESTHETICS

Jia Gu[1] (jjia.gu@mail.utoronto.ca), Dirk B. Walther[1], Claudia Damiano[1]; [1]University of Toronto

Many factors shape how humans judge something as beautiful or threatening. Research has extensively explored the relationship between symmetry and perceived beauty—demonstrating a consistent preference for symmetrical faces—while findings in other visual contexts, like art and landscapes, are less pronounced. Additionally, the relationship between symmetry and threat perception remains unclear. Given the link between threat judgments and felt expectancy, symmetry may influence threat perception by affecting the predictability of visual environments. In understanding the role of visual features on perception of beauty and threat, this study investigated whether the level of symmetry and complexity correlates with pleasure and threat ratings of natural scenes. To explore the underlying mechanisms of such influence on human perception, we included tasks to measure expectancy and memorability of each scene. We hypothesized that more symmetrical images would be more predictable, leading to reduced threat perception. Participants viewed 132 natural scenes with either the left or right half covered for 4 seconds and reported how well they could anticipate the covered half. Next, participants completed an old-new recognition test. Finally, they rated the pleasure and threat of each scene from the first task. We found a negative correlation between symmetry and complexity, $r(130) = -.25$, $p = .003$. We also observed that participants predicted images with more symmetric arrangements better as we hypothesized, $r(130) = .48$, $p < .001$. However, the analysis revealed that pleasure and threat ratings were not significantly associated with symmetry, memorability, or expectancy. Instead, visual complexity of natural scenes was correlated with threat ($r(130) = .30$, $p < .001$) and beauty ($r(130) = .64$, $p < .001$) ratings. These findings suggest that while visual complexity is related to symmetry, only complexity affects beauty and threat perception of natural scenes.

This work was supported by the University of Toronto Undergraduate Research Fund to J. G.

33.341 USING DYNAMIC KEY DEPTH TO IDENTIFY THE EMERGENCE OF DIAGNOSTIC FEATURES

Mathias Salvas-Hébert[1], Guillaume Lalonde-Beaudoin[2], Ian Charest[1], Daniel Fiset[2], Caroline Blais[2], Fréréric Gosselin[1]; [1]Département de psychologie, Université de Montréal, [2]Département de psychoéducation et de psychologie, Université du Québec en Outaouais

Psychophysics characterizes the relationship between experimental conditions and behavioral responses, offering insights into perceptual and decision-making processes. Among these responses, key presses are the most widely used. Traditionally discrete, key presses can now be tracked continuously using analog keyboards, providing fine-grained motor data. In this study, we leveraged Wooting analog keyboards to capture real-time responses with millisecond precision. Using Python and the tachypy (https://github.com/Charestlab/tachypy) and pyWooting libraries, we implemented a face classification task where participants judged emotional expressions (joy or fear) in stimuli partially revealed through sparse Gaussian apertures (“bubbles”). Five participants completed 592 trials, with the number of apertures dynamically adjusted to maintain 75% accuracy. Continuous response key depth data were recorded throughout each trial, allowing analysis of motor dynamics beyond traditional binary responses. As expected, facial regions associated with accurate classifications aligned with established findings. Interestingly, classification images derived from 10 ms bins of key positions (using the most depressed key within each bin as the response) revealed that the left eye region, from the observer’s point of view, drove responses as early as 500 ms — about 50 ms earlier than the right eye region.

33.342 READING THE BODY: NEURAL REPRESENTATIONS OF EMOTION AND SOCIAL MEANING IN NATURALISTIC HUMAN MOTION

Kiwa Tanaka[1], Chia-huei Tseng[1], Miao Cheng[1], Yuki Murai[2], Sai Sun[1]; [1]Tohoku University, [2]National Institute of Information and Communications Technology

Understanding how the human brain encodes body motion (BM) is essential for advancing theories in cognitive-social neuroscience. BM conveys rich, nonverbal information about inferred intentions (e.g., helping or harming), context-specific actions (e.g., greeting, expressing gratitude), or social emotions (e.g., anger directed toward others) even from a single individual’s movements that implicitly involve others. Using naturalistic video stimuli of professional performers with blurred faces, we focused on how emotion and sociality are conveyed purely through BM. In Experiment 1, we selected a set of 80 from 720 naturalistic videos in a performer-generated database (Tseng et al, 2025 VSS), based on 96 participants’ behavioral ratings along three dimensions: valence (positive vs. negative), arousal (high vs. low), and sociality (self-directed vs. other-directed). The selected stimuli were of equal arousal level with clear valence and sociality tendency. In Experiment 2, 28 participants viewed these videos during fMRI scanning, followed by a post-scan ratings task, allowing us to link brain responses to perceived emotional and social dimensions and validate the initial categorization. Neural results revealed that negatively valenced movements elicited stronger activation in the bilateral posterior insula and lingual gyrus, regions associated with interoception, empathy, and body-based semantic processing. In contrast, other-directed movements activated the bilateral superior temporal sulcus, a region involved in social perception and interpreting social interactions conveyed through BM. These findings provide novel neural evidence that the brain distinguishes between emotional, semantic, and social content conveyed through BM. By leveraging dynamic stimuli, our study advances understanding of how the brain encodes the multidimensional nature of body language. Future work will explore visual motion dynamics and individual differences in BM processing, with implications for early diagnosis and targeted interventions for individuals with social cognitive deficits.

33.343 ARE WE BETTER AT DIFFERENTIATING OUR GENETICALLY CLOSER RELATIVES? DIFFERENCES IN PERCEPTUAL SENSITIVITY TO HUMAN, CHIMPANZEE, AND MACAQUE FACES.

*Anuk Dias[*1] (dias0019@umn.edu), Vandita Gupta[*2], Abigail West[1], Charisse Pickron[2]; [1]Department of Psychology, University of Minnesota, [2]Institute of Child Development, University of Minnesota*

Human adults are better at differentiating among human faces relative to non-human primates faces. This robust finding has been demonstrated by the inversion effect eliciting larger, more negative neural responses to inverted human faces. However, there is a research gap in comparing the human perception of different species of non-human primate faces. Differences in facial recognition could arise from evolutionary impulses or featural differences (presence of visible sclera or spacing). In this on-going EEG study, we examined differences in time-locked face-sensitive neural responses (P1 and N170 components) to humans, chimpanzees, and macaques faces across 38 adult participants using an ERP inversion effect task. Participants viewed 600 randomized grayscale, front-facing, luminance-matched upright (n=300) and inverted (n=300) images. ERP component amplitude and latency were separately analyzed in a 2x3x2 repeated measures ANOVA with face orientation (upright, inverted), face species (human, chimpanzee, macaque), and hemisphere (left, right) as within-subjects variables. For the P1, a statistically significant main effect of species was found ($F=23.68$, $p<0.001$, $\eta p2= 0.39$) such that chimpanzee faces elicited the highest amplitude ($p<0.001$) relative to both human and macaque faces. For the N170, a statistically significant interaction between species and orientation ($F=13.13$, $p<0.001$, $\eta p2=0.27$) was found. A pairwise comparison revealed a larger, more negative, N170 amplitude for human inverted faces compared to upright faces (p<001). No differences were found for either chimpanzees or macaque faces. These preliminary findings indicate a replication of previously reported face inversion effect. This effect was not present for chimpanzee or macaque faces, suggesting differential engagement of face-sensitive processing mechanisms. Chimpanzees had an unexpectedly higher P1 amplitude compared to humans or macaques. This may indicate that chimpanzee faces are perceived distinctly from animal and human face categories.

33.344 IS MULTISENSORY INTEGRATION NECESSARY TO RELIEVE SACCADIC SUPPRESSION OF DISPLACEMENT THROUGH AUDITORY BLANKING?

Gavin Woodward[1], Satoshi Shioiri[2], Chia-Huei Tseng[2], Hiu Mei Chow[1]; [1]St. Thomas University, Fredericton, Canada, [2]Research Institute of Electrical Communication, Tohoku University, Sendai, Japan

Observers often fail to detect the displacement of objects during saccades—a phenomenon termed saccadic suppression of displacement (SSD). Interestingly, SSD is reduced when a brief gap occurs in a sound presented alongside the saccade target, resulting in improved displacement discrimination (auditory blanking effect; Chow et al., 2025 VSS). The auditory blanking effect suggests that sounds are weighed in visual stability judgements, but how is not yet known. We evaluated one possible mechanism: that the sound is bound with the visual target, such that auditory blanking indicates a change in the visual target. Given that such binding is most likely to occur when multimodal stimuli are aligned in time, we hypothesize that reducing temporal synchrony between the sound and visual target should reduce the auditory blanking effect. Twenty participants made saccades (18°) toward visual targets that were presented alongside a sound, either synchronously or asynchronously (with a 150 ms lead for the sound). Once the participants' saccade was detected, the visual target was displaced to the left or right (0.33°) and participants were tasked to report the direction of displacement. Critically, to test the auditory blanking effect, the sound was paused for 100 ms upon saccade detection (auditory gap) in half the trials and was continuous (no gap) in the other half. Results showed that the auditory gap improved participants' displacement discrimination ($d' = 0.47$), compared to no gap ($d' = 0.31$; $p = .038$, $\eta 2 = .031$), successfully replicating the auditory blanking effect. However, manipulating the temporal synchrony of the visual target and sound did not influence the size of the auditory blanking effect ($p = .266$, $\eta 2 = .007$). The latter finding does not support the notion that multisensory integration is necessary for the auditory blanking effect to occur, suggesting the need for further investigation into alternative explanations.

This work was carried out under the Discovery Grant Program of the Natural Sciences and Engineering Research Council of Canada (RGPIN-2024-06028) and the Cooperative Research Project Program of the Research Institute of Electrical Communication, Tohoku University (R05/A17).

33.345 THE PERCEIVED VANISHING POINT IN RENAISSANCE PAINTINGS

Jacob Tack[1], Jaya Joshi[1], Emily Fitzgerald[1], Ederlyne Mae Alcaraz[1], Ethan Brickey[1], Anirudh Praveen[1], Lars Strother[1], Gideon Caplovitz[1]; [1]University of Nevada Reno

The development and use of accurate linear perspective was one of the revolutionary innovations of the early Renaissance period. One point perspective employs a single vanishing point and there are several examples of Renaissance period paintings in which this vanishing point is precisely located. Despite the skill and effort required to accurately represent the vanishing point in one point perspective an unanswered question is whether or not observers perceive the vanishing point where it actually is? In this experiment we sought to answer this question by having non-art experts report where they thought the vanishing point appeared to be located in examples of one-point perspective Renaissance paintings. Our findings are quite remarkable! First, in certain paintings, particularly those with relatively explicit vanishing points located towards the center of the image, observers are quite accurate in recognizing the location of the vanishing point. Second: in many images, observers have little to no idea where the vanishing point is located, demonstrated by high inter-observer variability and inaccurate response. Third and most excitingly, in a small subset of paintings, including some of the most famous, the perceived vanishing point is systematically biased away from where it is actually located with a high degree of inter-observer reliability. We hypothesize that such biases arise due to configural aspects of the paintings that lead to perceptual phenomena/illusions familiar to the perceptual scientist.

33.346 HOW PERCEIVED STRESS AND VALENCE INTERACT WITH SCENE PERCEPTION AND MEMORY

Maeve McClure[1], Elissa Aminoff[1]; [1]Fordham University

Episodic memories can be impacted by a variety of factors, one being stress. Stress can facilitate memory systems in the brain, specifically increasing memory for unpleasant scenes. Current research regarding stress and memory has largely been conducted without considering individual responses to the valence of stimuli as well as focusing on objective stress (e.g. cortisol levels). This study aimed to address the relationship between perceived stress and memory, as well as perception of valence. Participants briefly viewed 120 images of valenced scene stimuli, and made simple judgments about the scenes. This was followed by a recognition memory task in which participants were asked if the scene was old or new. Following the scene stimuli encoding and recognition task, participants rated the valence of stimuli and completed the Perceived Stress Scale. There was a significant effect of valence at encoding, with accuracy for negative images being significantly worse than positive and neutral images, and with reaction time for negative images being significantly higher than for positive and neutral images. There was also a significant effect of valence on memory, with participants recognizing significantly fewer neutral images than positive and negative images. The critical question was how does perceived stress modulate valenced scene processing? A significant negative correlation was found between perceived stress and accuracy for negative images at encoding, and a statistically significant positive correlation was found between perceived stress and reaction time for negative images at encoding. This indicates that perceived stress modulated how individuals viewed negative images. Together, using ratings from individuals we found an effect of valence at encoding and memory, and this was modulated by perceived stress.

33.347 ACTIVE FORGETTING IS TRIGGERED BY THE IMPLIED POSITION OF THE SUN

Chirag H Deepak[1] (chirag07d@gmail.com), Joan Danielle K. Ongchoco[1]; [1]The University of British Columbia

Visual processing organizes continuous visual input into discrete events. The boundaries between these events facilitate adaptive forgetting: when the statistics of our immediate environments have dramatically changed, it may be helpful to 'clear' memory of previous information that no longer applies to the new event. Event boundaries have been explored in many contexts—from scene cuts in film, chapters in narratives, and even walking through doorways. But one of the clearest and most natural event boundaries we cross is simply going from one day to the next. The time of the day can be marked by external clocks and calendars, but it can also be more implicitly extracted from the position of the sun. Here we asked whether these subtler visual cues are spontaneously encoded, and in turn, then trigger active forgetting. Observers watched virtual animations in which they were situated in a waiting room with three windows. They were shown a list of pseudo-words, after which the animation played simulating the actual movement of the sun over the course of the day. The position of the sun was implied only by the shadows reflected by objects in the room. Observers either saw a sunrise-to-noon or noon-to-sunset transition. Animations were matched for duration and visual complexity, differing only in the direction of lighting changes. Recognition memory was tested immediately after. Results revealed worse memory in the noon-to-sunset compared to the sunrise-to-noon condition. No such memory differences were found in a control experiment with matched internal room lighting changes (i.e., brightening versus dimming light strips), with the sun position remaining constant throughout. This confirms that the observed memory effects were driven by sun position rather than general illumination changes. Altogether, these findings suggest that the visual system uses environmental priors, such as the sun's position, to regulate memory dynamically.

33.348 READING EFFICIENCY IN AMBLYOPIA, ASSESSED WITH RANDOM TEMPORAL SAMPLING

Dasha Vanichkina[1], Nicole Dranitsaris[1], Martin Arguin[2,3], Alexandre Reynaud[1,4]; [1]McGill University, [2]Université de Montréal, [3]Centre de recherche, Institut Universitaire de Gériatrie de Montréal, [4]Research Institute of the McGill University Health Centre

Amblyopia, or "lazy eye", is a condition where incoming visual information from one eye is processed poorly in the brain and binocular integration is disrupted. It is known that reading is altered in amblyopia; individuals affected by this condition tend to have a slower reading speed and present different eye movement patterns compared to controls. We believe that differences in visual oscillatory mechanisms of amblyopic participants may account for this dissimilarity. Thus, we used the novel method of random temporal sampling to evaluate this theory. In the present study, we aimed to investigate whether there are differences in the temporal features of the reading efficiency of controls and amblyopic subjects. Six amblyopic and six control participants were tested monocularly and binocularly. Participants were asked to read three to six letter words presented for 400 ms and sort them into one of four categories. During exposure, the visibility of targets varied randomly through time by manipulating the signal (target word) to noise (white noise field) ratio. Classification images, which reflected processing efficiency, were calculated based on response accuracy, which was maintained at 50%. As anticipated, controls read more efficiently overall and were most efficient in the binocular condition. Amblyopic participants processed most efficiently with their non-amblyopic eye and struggled the most in the binocular condition. Lastly, controls showed an overall better level of binocular summation than amblyopic participants, meanwhile amblyopic subjects seemed unable to integrate binocular visual information to aid them in word identification. Overall, our results demonstrated that the temporal features of processing efficiency of controls and amblyopic subjects differ in meaningful ways during reading. These results have implications for the development of treatments for amblyopia that specifically target high-level processes.

This research was funded by the Vision Sciences Research Network mobility grant to DV; an NSERC discovery grant RGPIN-2024-06447 and a startup fund from the RI-MUHC to AR; and operating NSERC and FRQST grants to MA.

33.349 TESTING THE LINK BETWEEN VISUAL IMAGERY AND THE TACTILE BOUBA-KIKI EFFECT

Sid d'Entremont[1], Hiu Mei Chow[2]; [1]St. Thomas University, Fredericton, Canada, [2]St. Thomas University, Fredericton, Canada

The tactile Bouba-Kiki (BK) effect refers to a non-arbitrary mental association between tactile sensations (e.g., shape with rounded edges) and other stimuli such as nonsense words (e.g., bouba). It has been reported that individuals with congenital blindness and young children exhibit weaker tactile BK effects. One possible explanation is that these populations may also have weaker visual imagery for abstract shapes, suggesting a link between visual imagery and the tactile BK effect. To test this link directly, we evaluated whether the strength of the tactile BK effect was related to self-reported visual imagery abilities. To measure participants' visual imagery, we used the Vividness of Visual Imagery Questionnaire (Marks, 1973). Questions involve visualizing familiar images and rating their mental vividness. To measure the strength of the tactile BK effect, we asked participants (n = 30) to perform an audio-tactile task in which they blindly felt some spiky and some round 3D-printed shapes, listened to recordings of nonsense words, and rated the congruence of the word-shape pairs on a 1-7 scale. Results showed that nonsense words containing back vowels (e.g., /a/, /o/, /u/) and bilabial consonants (e.g., /b/, /p/, /m/) were associated with felt round shapes, whereas nonsense words containing voiceless stop consonant (e.g., /t/, /k/) were associated with felt spiky shapes, in alignment with our hypotheses. However, front vowels were not associated with spiky shapes, contrary to typical findings, suggesting that audio-tactile associations might differ from audio-visual ones. Split-half analyses revealed that weaker visualizers (VVIQ < group median, i.e., 57/80) did not exhibit different patterns compared to stronger visualizers (VVIQ ≥ group median), suggesting that voluntary visual imagery might not mediate these associations. A larger sample of weak visualizers (VVIQ < 32/80) may better conclude the involvement of visual imagery in the tactile BK effect.

This work was carried out under the Discovery Grant Program of the Natural Sciences and Engineering Research Council of Canada (RGPIN-2024-06028).

33.350 HOW PERCEIVED EMOTIONS OF OTHERS **DISRUPT OUR SENSE OF THEIR PERSISTING "SELVES":** EVIDENCE FROM THE TUNNEL EFFECT

Jocelyn S. Zhang[1], Joan Danielle K. Ongchoco[1]; [1]The University of British Columbia

Emotions have the capacity to make people feel 'not quite like themselves'. This common expression implies a shift in identity—not in terms of mere facial identity, but a deeper sort, involving a sense of who a person is over time. Questions on identity are long-standing philosophical puzzles that have been difficult to explore empirically. Here, by leveraging a paradigm in visual perception, we ask how perceived emotional transitions (e.g., seeing a neutral face become sad) might interact with identity persistence (i.e., how we represent an individual as persisting over time). We adapted a classic object persistence paradigm, the tunnel effect, in which an object that passes behind an occluder is perceived as the same object, despite changes to superficial features (e.g., shape, color). Observers saw faces pass behind an occluder. These faces were imposed with a target letter, and observers simply reported whether the letter stayed the same or different across the occluder. Critically, we varied the face's emotions, where sometimes the face maintained its emotion through the occluder (e.g., starting sad and staying sad), or not (e.g., starting neutral and becoming sad). When emotions were maintained across the occluder, response times were facilitated in the same-target condition than in the different-target condition—replicating the basic tunnel effect. This effect, however, disappeared when the face changed emotions across the occluder—despite the face maintaining its superficial identity, with a reliable interaction across conditions. This disruption in identity persistence mimicked the same pattern of results when the actual identity switched (e.g., when a different person's face emerged from the occluder), while emotion remained constant. These findings first demonstrate that a 'deeper' sense of identity is spontaneously tracked in visual processing—and that this may be more contingent on perceived changes in emotion than we might have previously thought.

33.351 FEEDFORWARD TRAVELING WAVES OF NEURAL ACTIVITY TRACK SENSORY-TO-MEMORY TRANSFORMATIONS DURING THE RECALL OF WORKING MEMORY CONTENT

Cassidy Wolfe[1] (cassidywolfe63@gmail.com), Edward Ester[1]; [1]University of Nevada, Reno

Spatially organized neural oscillations –or traveling waves– are a key mechanism for guiding the propagation of neural activity and computational processes across the brain. In a recent EEG working memory (WM) study, we documented a feedforward traveling wave propagating from posterior occipitoparietal electrode sites to frontocentral electrode sites that emerged right before the initiation of a working-memory-guided behavioral response and whose the peak latency predicted intra- and inter-individual differences in response times. Our working hypothesis is that this wave indexes the transfer of mnemonic content from storage sites in occipitoparietal cortex to motor sites responsible for producing task-relevant behaviors, but an alternative (but non-exclusive) hypothesis is that this wave is instead related to the planning of an upcoming behavior. Participants performed two versions of a visuomotor WM task in which they memorized the orientations of two bars, then recalled the orientation of a retrospectively probed bar as precisely as possible. During the "predictable recall" task, participants recalled the orientation of the retrospectively probed bar by adjusting the orientation of a recall stimulus that always started with a vertical orientation. Thus, in this condition, participants could anticipate how much (or for how long) they would have to rotate the recall stimulus before beginning their response. Conversely, during the "random recall" task, on each trial the recall stimulus was assigned a random starting orientation knowable to participants only after they began their response, obviating advance planning. The onset, timing, and peak latency of the feedforward traveling wave during the predictable recall task was identical to values observed in prior studies, however, the feedforward

traveling wave was altogether eliminated during the random recall task. From this, we speculate that the feedforward traveling wave we identified in our prior research indexes a sensory-to-motor transformation that guides the precise recall of stimuli under predictable conditions.

33.352 BACKGROUNDS AFFECT PERCEIVED LIGHTNESS THROUGH MULTIPLICATIVE GAIN

Heshu Yin[1], Joseph R Busch[1], Vijay Singh[1]; [1]Haverford College

The perceived lightness of an object depends on its surrounding context. We measured lightness discrimination thresholds for judging the lightness of two gray patches when the patches were viewed against backgrounds of varying lightness. In each trial of a 2AFC task, human observers (N=3 observers) viewed two square patches (1.68 degree visual angle) presented on larger square backgrounds (5.57 degree visual angle) for 500ms. The patches and the backgrounds were presented simultaneously side by side (centers 5.57 degree apart) on a monitor. One patch had a fixed standard lightness, while the lightness of the comparison square patch was varied. Observers reported the patch that they perceive to be lighter with a controller. Between trials, the lightness of the comparison patch and the lightness of its background were changed. Percent comparison chosen data was collected for 11 comparison lightness values for each standard lightness (30 trials each). The resulting psychometric function was fit with a cumulative normal function to calculate the point of subjective equality (PSE) and 76% discrimination threshold. Psychometric data was collected for three standard lightness levels and three comparison background lightness levels. We develop a signal detection theory that includes the effect of background lightness on the perceived lightness of an object. The model predicts that an additive effect would shift the psychometric function, changing the PSE without altering the slope. On the other hand, a multiplicative effect would change both the PSE and the slope of the psychometric function. Our experimental data shows that for a fixed standard lightness, the PSE and thresholds changed proportionally with the lightness of the comparison background, supporting a multiplicative gain mechanism.

33.353 IS THE FRONTO-PARIETAL PHYSICS NETWORK VISUAL OR AMODAL?

Thomas W. Brewitt[1] (twbrew@mit.edu), Vivian C. Paulun[1], RT Pramod[1], Nancy Kanwisher[1]; [1]MIT

The Physics Network (PN) is a set of brain regions in the fronto-parietal cortex engaged when people make physical rather than descriptive judgments on dynamic visual stimuli, and representing physical properties and relations of objects ("Things") and non-solid substances ("Stuff"). Prior studies on this network have focused exclusively on the processing of visual information, whereas we also make physical inferences based on sounds, such as the 'boing' of a bouncing ball, the 'splash' of water, or the 'clink' of wine glasses. Here, we test whether the PN responds to audio recordings of physical events, or whether it operates primarily on visual input. We compiled short (2 sec) naturalistic audio and video clips of three different types: 1) "Things", e.g., rolling dice, 2) "Stuff", e.g., water sloshing, and 3) social scenes, e.g. people laughing. These stimuli were presented in a 2 (modality) x 3 (condition) blocked fMRI design while subjects (N=4) performed a 1-back task. The independently localized PN showed a significantly stronger response to videos of physical events (both Stuff and Things) than social events (p<0.05), replicating and extending previous findings. Crucially, we found that the response to audio clips of physical events was significantly lower than to video clips showing similar physical events (p<0.005), and indeed no higher than the fixation baseline. There was no difference between the auditory "Things" and "Stuff" conditions. This pattern of responses was robust and present in each participant individually. Our findings show that the PN doesn't respond to the sounds of physical events, suggesting that the PN is not a general system for amodal intuitive physics, but that it is specifically engaged in visual physical scene understanding.

33.354 OSCILLATORY DYNAMICS OF VISUAL PERCEPTION AND WORKING MEMORY UNDER COMPETITION

Yifei Wu[1], Khayla Santiago[1], Chunyue Teng[1]; [1]Lawrence University

Attention fluctuates rhythmically over time, influencing both visual perception and visual working memory (WM). Previous research has demonstrated periodicity in attentional sampling across these domains, yet it remains unclear how rhythmic attention coordinates the concurrent processing of external and internal visual representations. The current study investigates whether rhythmic attentional sampling is modulated by competition between these processes. Participants performed a dual-task paradigm to maintain a visual stimulus in WM while performing a concurrent perceptual judgment. In Experiment 1, both tasks involved orientation judgments, maximizing competition between internal and external processing. In Experiment 2, the perceptual task required luminance judgments, while the WM task involved orientation, reducing direct competition. Stimulus onset asynchrony (SOA) between the perceptual stimulus onset and the probe onset varied systematically across 50 intervals (0.5–1.5 seconds, in 0.02s steps). Time-frequency decomposition was applied to response time and accuracy using a Fast Fourier Transform. Spectral power and phase angles were analyzed in the theta and alpha frequency ranges (3–15 Hz). Non-parametric permutation testing assessed significance in increased spectral power across frequencies. In Experiment 1 (High Competition), performance fluctuations in both tasks were observed in the theta and low-alpha range. In Experiment 2 (Low Competition), rhythmic fluctuations in WM performance were identified at ~6 Hz (theta range). Cross-task comparisons showed greater phase offset between WM and perceptual performance in Experiment 1 than in Experiment 2, suggesting increased interference. These findings suggest that visual perception and WM are governed by a shared oscillatory control process. Moreover, the phase offset between internal and external sampling appears to increase under high competition, potentially serving as a mechanism for resolving representational interference. These results support the idea that rhythmic attention dynamically regulates access to perceptual and mnemonic representations based on task demands.

33.355 INVESTIGATING THE LINK BETWEEN TASK SWITCHING AND RESPONSE INHIBITION IN A MULTISENSORY ANTISACCADE TASK

Ashika Kamboj[1] (ashika03@student.ubc.ca), Miriam Spering[1], Anthony Herdman[1], Skadi Gerkensmeier[2], Jessica Chalissery[1]; [1]University of British Columbia, [2]University of Luebeck

Controlling behavior in our complex, multisensory environment requires rapid switching between sensory inputs and tasks. The antisaccade task, a well-established paradigm for investigating response inhibition, can be adapted to study task switching by cueing participants to either perform a prosaccade or an antisaccade. Here we introduce a novel antisaccade paradigm that uses uni- and bimodal visual and auditory stimuli to elicit task switching in response to changing stimulus configurations rather than cued motor responses. Participants (n=22) completed an antisaccade task with either visual, auditory or combined visual-auditory trials with stimuli either spatially congruent (same side of fixation) or incongruent (opposite sides of fixation). In separate visual- and auditory-cued antisaccade blocks, stimulus configurations (unimodal, congruent, incongruent) were randomly interleaved. Trials were categorized as switch trials if the configuration changed from the previous trial and no-switch trials if it remained the same. A subset of participants (n=5) completed a blocked version of the task. Antisaccade latency did not differ between switch and no-switch trials, revealing no switch cost. A within-subject comparison (n=5) of the blocked and interleaved task latencies uncovered an alternation cost, with the average latency in the blocked task 36 ms shorter than that of switch trials in the interleaved task ($p<.001$). These costs were significantly larger in the auditory task compared to the visual task ($p=.002$). Our findings indicate that switching between sensory modalities within an antisaccade task may not engage task-switching mechanisms in the same way as classic task-switching protocols. However, the presence of an alternation cost suggests that switching between different trial types in the interleaved task increases cognitive demands, as participants are unable to rely on a consistent strategy.

33.356 CONGRUENT CATEGORY LABELS FACILITATE RAPID SCENE CATEGORIZATION

Carina Wong[1], Sage Aronson[1], Michelle Greene[1]; [1]Barnard College

Reading and visual categorization are both considered automatic cognitive processes, but their relative time courses are unknown. Pictures and their respective category word labels can each provide category information. Still, we do not yet know whether these interact to facilitate or interfere with processing speed. We used a paradigm inspired by the Stroop effect to investigate the extent and time course of the interaction between scenes and category labels that are either congruent or incongruent. We recorded 128-channel EEG while participants viewed stimuli in each of four conditions: words alone (scene category names printed over a phase-randomized scene), pictures alone, and word-picture composite stimuli in which both cues corresponded to the same scene category (congruent, e.g., the word “church” on top of an image of a church), and both cues corresponded to different categories (incongruent, e.g., the word “church” on top of an image of a park). The participants were tasked with categorizing the target word or picture as “indoor” or “outdoor”. We used a linear support vector machine to assess the time course of scene categorization using whole-brain patterns and five-fold cross-validation. As expected, we found earlier and more robust decoding for pictures than words alone. Interestingly, although pictures alone and congruent composites had similar decoding onsets, we observed stronger and more sustained decoding accuracy for the composites, suggesting that congruent scene category words can affect the robustness of early scene category representations. Finally, we noticed little difference in decoding accuracy for congruent versus incongruent composites, suggesting that although congruent words can facilitate scene representations, pictorial content alone is sufficient to create a scene representation. Altogether, these results suggest that there is facilitation in neural processing when two cues contribute to corresponding to the same information, but little to no interference when the cues conflict.

NSF CAREER 2240815 to MRG

33.357 DO THESE RATIOS MATCH? EXPLORING THE ROLE OF PROPORTIONAL REASONING IN SYMBOLIC MATH FLUENCY

Margaret Shideler[1], Yaxin Liu[1], Adam Green[1]; [1]Georgetown University

Extracting and manipulating relational structure from visual information is central to human cognition. In mathematical contexts, visuospatial processes such as magnitude comparison, pattern recognition, and proportional reasoning play an important role in supporting symbolic mathematical competence. Of particular importance to this relationship is nonsymbolic proportional reasoning (NPR), or the ability to extract proportional relationships from non-numerical stimuli. However, the role of analogical reasoning—particularly when applied to proportional relationships—remains underexplored. This study investigates whether the ability to reason analogically about proportions predicts symbolic math performance, and how it relates to performance on traditional NPR tasks. Ten adult participants recruited from Prolific completed two traditional NPR tasks (area-based probability comparisons and clock reading) as well as a novel visual analogy task designed to integrate NPR and analogical reasoning. In the visual analogy task, participants compared two pairs of shapes with an area ratio of 1:1, 3:2, 2:1, 5:2, or 3:1. On each trial, participants judged whether the two pairs form a true or false analogy: a true analogy occurs when both pairs share the same ratio, while a false analogy occurs when they do not. Fraction fluency was assessed using a fraction sum estimation task. Preliminary findings indicated that performance on the visual analogies task showed similar trends as on traditional NPR tasks: in both the visual analogy task and the probability task, accuracy decreased as the ratios depicted in the stimuli became more similar. Moreover, initial analyses suggest that the strength of relationship between fraction fluency and NPR performance varies by task type (probability, clock, or visual analogy) and difficulty level. These findings highlight the broad role of proportional reasoning in supporting symbolic math fluency, as well as a potential role of analogical reasoning above and beyond NPR.

33.358 TOP-DOWN, HEMISPHERE-SPECIFIC TRAVELING WAVES OF CORTICAL ACTIVITY TRACK THE UPDATING OF WORKING MEMORY CONTENT.

Gabriel Hunt[1], Dipin Kapila[1], Edward Ester[1]; [1]University of Nevada, Reno

Contemporary neuroimaging research supports a sensory recruitment model of working memory (WM), where top-down signals originating in frontal and parietal cortex coordinate the storage of mnemonic information in sensory cortical areas. Communication between frontoparietal and sensory areas during WM is established and maintained via neural oscillations. Mounting evidence suggests that neural oscillations frequently take the form of traveling waves, or spatially coherent oscillations that propagate across the cortex. Whether and to what extent traveling waves contribute to the encoding, maintenance, and updating of WM content is unknown. Here, we explored this issue by tracking hemisphere-specific traveling waves while participants encoded, updated, and held information in WM. We recorded EEG while participants remembered the orientations of two colored bars – one per visual hemifield – over a short delay. At the end of each trial, a retrospective color probe instructed participants to recall the orientation of one bar as precisely as possible. During each trial, a third bar ("update stimulus") was presented in the left or right visual hemifield midway through the delay period. On 50% of trials, a cue instructed participants to remember this new bar rather than the one it replaced ("replace" trials); on the remaining 50% of trials a cue instructed participants to ignore the new bar and continue remembering the original bar presented in the same location ("hold trials"). Analyses of cortical traveling waves time-locked to the onset of the update stimulus revealed a low-frequency (4-6 Hz), feedback (i.e., frontal-to-occipital) traveling wave over electrode sites contralateral to the location of the update stimulus during replace but not hold trials. This wave was distinct from a second low-frequency, feedforward (i.e., occipital-to-frontal) traveling wave evoked by a subsequent recall probe. These results provide preliminary evidence linking hemisphere-specific traveling waves of cortical activity to the encoding of new WM content.

33.359 FEEDFORWARD TRAVELING WAVES OF NEURAL ACTIVITY INDEX THE READ-OUT BUT NOT THE SELECTION OF TASK-RELEVANT WORKING MEMORY CONTENT.

Dayana Valdez[1], Edward Ester[2]; [1]University of Nevada, Reno

Spatially organized neural oscillations – or traveling waves – are a key mechanism for guiding the propagation of neural activity and computational processes across the brain. In a recent EEG working memory (WM) study, we documented a feedforward traveling wave propagating from posterior occipitoparietal electrode sites to frontocentral electrode sites that emerged right before the initiation of a working-memory-guided behavioral response and whose the peak latency predicted intra- and inter-individual differences in response times. Our working hypothesis is that this wave indexes the transfer of mnemonic content from storage sites in occipitoparietal cortex to motor sites responsible for producing task-relevant behaviors, but an alternative hypothesis is that this wave is instead related to the selection of task-relevant memory content. Here, we tested the latter possibility. Participants performed two versions of a visuomotor WM task in which they memorized the orientations of two bars, then recalled the orientation of a retrospectively probed bar as precisely as possible. During the pro-cue task, participants recalled the identity of the bar matching the color of a retrospective probe, while during the anti-cue task, participants recalled the identity of the bar that did not match the color of the retrospective probe. Participants responses were ~170 ms slower during the anti-cue compared to the pro-cue task, and the onset latency of lateralized alpha-band EEG activity associated with the selection of task-relevant memory content was ~120 ms slower during the anti-cue compared to the pro-cue task. However, the onset and peak latencies of the feedforward traveling wave that we identified in our prior work was identical during the pro- and anti-cue tasks. From this, we argue that feedforward traveling waves observed before the onset of a WM-guided behavior reflect more than the mere selection of task-relevant WM content.

33.360 NEURAL SIGNALS REFLECT DIFFERENCES IN ILLUSORY FACE PROCESSING AS OPPOSED TO FACES AND OBJECTS

Maleeha Chowdhury[1], Arianna Thoksakis[1], Edward Ester[1]; [1]University of Reno, Nevada

Pareidolia is the phenomenon in which one perceives meaningful patterns in ambiguous stimuli. A specific example of this is face pareidolia, which involves the perception of faces in stimuli in which none exist. This act of false perception is a natural error of the face processing system and thus operates as a gateway to study the neural mechanisms of face processing. A significant component of face perception research is the N170 event-related potential (ERP), a face-selective scalp potential. While previous studies have examined the involvement of the N170 in the neural bases of pareidolia, evidence has been inconsistent and limited to small samples of neurodivergent groups. To contrast such limitations in our study, we asked human observers (N = 14) to passively view stimuli of illusory faces, matched nonface objects, and faces while monitoring for face-selective electroencephalography (EEG) signals over 9,216 trials. This large number of trials provides the important advantage of providing insight on an individual basis due to the large number of measurements per observer. The N170 effect was present in all stimuli, but was only significant in face trials. However, this result was inconsistent at the individual level. We observed a heterogeneity of the N170 effect, with differing amplitudes for illusory face stimuli across individuals. Together, these results suggest differing neural amplitudes for face, object, and illusory face stimuli. We are now following up on these results with a response-limited paradigm that will allow us to track the latency and amplitude of N170 signals in relation to response time. These results will be presented at the VSS 2025 meeting.

This work was funded by the National Science Foundation

SUNDAY MORNING POSTERS IN PAVILION

Action: Grasping, reaching, pointing, affordances

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.401 VISUOSPATIAL PERFORMANCE DURING A POINTING TASK IN DEMENTIA WITH LEWY BODIES.

Annalisa Bosco[1] (annalisa.bosco2@unibo.it), Caterina Foglino[1,2], Lucia Guidi[3], Davide Braghittoni[3], Greta Venturi[3], Luisa Sambati[3], Patrizia Fattori[1], Raffaele Lodi[1,3], Caterina Tonon[1,3], Micaela Mitolo[4]; [1]University of Bologna, [2]Italian Institute of Technology, [3]IRCCS Istituto delle Scienze Neurologiche di Bologna, [4]University of Parma

Dementia with Lewy Bodies (DLB) is a neurodegenerative disorder marked by alpha-synuclein aggregates in the brain, with early visuospatial deficits distinguishing it from other dementias (Mc Keith et al. 2017). Understanding visuospatial dysfunction in DLB is crucial for diagnostic accuracy, as it often precedes or accompanies hallmark symptoms like visual hallucinations and motor disturbances. This study aims to elucidate such impairments by analyzing measures including movement accuracy, movement and task duration, reaction times, associated ocular patterns and brain volumetry in a pointing to visual targets task. A total of 16 patients with DLB (6F; mean age = 75.416.38) and 16 age-matched healthy controls (HC) (9F; mean age = 67.87 9.63) underwent a pointing task that required them to touch one of four visual target positions displayed on a touchscreen. By application of a linear mixed-effects model, we evaluated the effect of the group (patients vs HC) and the target position on the pointing accuracy, movement duration, task duration, reaction time, saccade duration, saccade amplitude and saccade accuracy. The same model was applied to evaluate the effect of group on the volumetry of the left and right superior parietal lobule (SPL) of the brain, as the SPL plays a critical role in visuospatial processing. A significant group effect was observed across the multiple motor and ocular parameters (F values ranging from 1.16 to 247.84, all $p < 0.01$), with patients consistently performing worse than controls independent from target position. Moreover, a significant volumetric loss in the left and right SPL was observed in the DLB patients compared to HC (F values ranging from 9.98 to 53.06, $p < 0.05$). These findings highlight the pervasive visuospatial and motor impairments in DLB, highlighting the importance of assessing both behavioral and neural metrics for improved understanding and differentiation of DLB from other dementia subtypes.

This study has been supported by the Italian Ministry of Health (#GR-2019-12369242) and by #NEXTGENERATIONEU (NGEU), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006).

33.402 DISTINCT MECHANISMS OF DISTRACTOR SUPPRESSION EVIDENCED BY REACHING MOVEMENTS

Shinhae Ahn[1] (a.shinhae@wustl.edu), Richard A. Abrams[1]; [1]Washington University in St. Louis

Suppressing irrelevant distractors is crucial for achieving goal-directed behavior in dynamic and unpredictable environments. Previous studies have shown that when salient distractors more frequently appear at a specific location, participants learn to suppress attention to that location. This learned suppression reduces distractor interference, even when distractors remain perceptually salient. Statistical learning-based distractor suppression differs from feature-based suppression, which relies on top-down attentional control to filter out distractors based on task-relevant target features. Although reaching movements have provided insight into many aspects of attentional control, little is known about how such movements might augment our understanding of the distinct types of suppression. This study examined the mechanisms of statistical learning-based and feature-based distractor suppression through reaching movement trajectories of a mouse cursor during target selection. Distractor suppression type was manipulated across two experiments: (1) a singleton-detection task, where the target was defined as a unique shape among homogeneous distractors, and a color singleton distractor appeared more frequently at one location (high-probability) than others (low-probability), enabling statistical-based suppression; and (2) a feature-search task, where the target was a specific shape among heterogeneous distractors, facilitating feature-based suppression. In the singleton-detection task, reaching movement trajectories exhibited no significant differences between high-probability distractor locations and distractor-absent trials, reflecting effective statistical learning, whereas attraction biases toward the distractors were observed at low-probability locations. In contrast, in the feature-search task, the movement trajectories demonstrated repulsion biases away from the suppressed distractors. Overall, the reaching movement trajectories revealed distinct patterns caused by different distractor suppression mechanisms. These findings highlight the utility of reaching movements in uncovering the attentional control processes underlying statistical learning-based and feature-based distractor suppression.

33.403 SERIAL DEPENDENCE IN GOAL-DIRECTED HAND MOVEMENTS

Cristian Muñoz-Puelman[1], Alexander Goettker[2,3], Cristina de la Malla; [1]Vision and Control of Action Group, Department of Cognition, Development, and Psychology of Education, Institute of Neurosciences, Universitat de Barcelona, Barcelona, Catalonia, Spain, [2]Justus Liebig Universität Giessen, Germany, [3]Center for Mind, Brain and Behavior, University of Marburg and Justus Liebig University, Marburg, Germany

Attractive serial dependence effects have been shown across various perceptual tasks, and for actions like eye and head movements. We know less about how these biases may influence performance in dynamic tasks such as interception, where accurate responses depend on real-time information about moving targets. To investigate this, we asked participants (N=14) to perform an interception task where each trial consisted of two sequentially presented targets. The first target, the prior, moved either leftwards or rightwards at a velocity of 5 or 15 cm/s. The second target, the probe, always moved at 10 cm/s, either in the same or in the opposite direction as the prior. To dissociate the role of active motor engagement with the prior from its purely perceptual influence on probe interception, participants completed two conditions: (1) intercepting both prior and probe targets, and (2) observing the priors and intercepting only the probes. Consistent with previous findings, probes tended to be intercepted further ahead following a fast prior and further behind following a slow prior. However, the magnitude of this effect varied significantly across conditions: When participants intercepted both targets, the prior's velocity influenced interception errors when the prior and probe moved in the same direction, but less so when they moved in opposite directions. In the observation condition, the effect had a similar magnitude regardless of whether the prior and probe moved in the same or opposite directions. These results suggest that repetitive movement execution could play a role in modulating the influence of prior target velocity and indicate potential differences in how velocity information is encoded and used in different cases: as a vector when intercepting all targets, and as a scalar when intercepting only the probes. Such flexibility in velocity processing may reflect adaptive strategies to balance perceptual and motor requirements in dynamic tasks.

Work supported by grants PID2023-150883NB-I00 and CNS2022-135808 funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR to CM. CMP was supported by an ANID grant 72230035. AG was supported by the DFG; project number 222641018-SFB/TRR 135 Project A1.

33.404 TRAINING CONTEXT DETERMINES IMPLICIT AND EXPLICIT CONTRIBUTIONS TO MOTOR ADAPTATION

Pamela Villavicencio[1], Jonathan S. Tsay[2], Cristina de la Malla[1]; [1]Universitat de Barcelona, [2]Carnegie Mellon University

Motor adaptation relies on both explicit and implicit learning processes. However, how these processes are impacted by the training context is not well understood. To address this, we compared two groups of participants (N=40) that adapted to a visuomotor gain perturbation (0.7) in a centre-out reaching task with different training target configurations. The Angular Group trained with targets of equal amplitude but in different directions, while the Extent Group trained with targets of different amplitudes in a single direction. Both groups shared a common training target. After training with continuous perturbed feedback, we assessed generalization by interleaving no feedback probe trials in which participants reached either to novel targets or to the shared target. In the final phase, we measured aftereffects by removing visual feedback entirely and having participants reach to the same target positions tested during the generalization phase. During the training and generalization phase, motor performance reflected both implicit and explicit learning processes. In the aftereffects phase, participants were instructed to forgo any explicit strategies, ensuring that performance indexed only implicit adaptation. Our primary dependent variable was the change in movement amplitude across these phases. We found that both groups learned to compensate for the visuomotor perturbation rapidly and fully, showing no measurable differences. However, the groups differed strikingly in how they learned: At the common training target, the Extent Group showed smaller aftereffects than the Angular Group, indicating a stronger reliance on explicit strategies. At a novel amplitude in the trained direction, the Extent group outperformed the Angular group through the use of a successful explicit strategy. At novel angular targets, the Angular group outperformed the Extent group, relying on greater implicit adaptation. Together, these findings underscore the role of training context in shaping how we learn.

CM was supported by PID2023-150883NB-I00, CNS2022-135808, MCIN/AEI/10.13039/501100011033 and the European Union NextGeneration EU/PRTR . PV by grant PRE2021-097890, MICIU/AEI/10.13039/501100011033 and FSE+. JST by Carnegie Mellon University Department of Psychology Seed Funding.

33.405 TRANSFER BETWEEN SACCADE AND REACH ADAPTATION USING CONCURRENT SINE-WAVE PERTURBATIONS

Zinong Li[1,2,3] (zl2242@nyu.edu), Michael Landy[2,3,4], Robert Volcic[1,5,6]; [1]New York University Abu Dhabi, [2]New York University, [3]NYU Department of Psychology, [4]NYU Center for Neural Science, [5]NYUAD Center for Artificial Intelligence and Robotics, [6]NYUAD Center for Brain and Health

INTRODUCTION. Saccades and reaches are tightly coordinated movements, yet the extent to which they share planning resources is unclear. METHODS. To investigate cross-effector interaction, we employed a motor-adaptation paradigm that applied concurrent perturbations to reaches and saccades. The perturbation amplitudes followed a sine wave time course across trials, with different temporal frequencies and orthogonal perturbation directions across effectors. Participants reached for a visual target with view of the hand and arm occluded and made a saccade to the target after reach completion. Saccade adaptation was triggered by displacing the target during the saccade, while reach adaptation was induced by altering the feedback of the reach endpoint relative to the original target location. By extracting reach and saccade errors as adaptation responses, we identified both spatial and temporal signatures of adaptation, confirming that each effector's adaptation was successfully triggered. We fitted a sine wave to the sequence of adaptation responses using a maximum a posteriori criterion. The sine-wave frequency was fixed to match the perturbation frequency, and the posterior distribution of sine-wave amplitudes was computed using Markov Chain Monte Carlo sampling, providing credible intervals to assess the presence and magnitude of adaptation and transfer effects. RESULTS. Both reaches and saccades adapted to their respective perturbations. In addition, we found that reach responses also mirrored the temporal and spatial signature of the saccade perturbation: a transfer of saccade adaptation to reach planning. In contrast, reverse transfer from reach adaptation to saccade planning was not observed. CONCLUSIONS. Our approach provides a sensitive method to quantify adaptation effects that can help resolve previous conflicting results. The observed asymmetry may stem from an asymmetry of the control of saccades vs. reaches in natural behavior such that saccade planning influences reach planning, but not the converse.

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. We acknowledge the support of NIH grant EY08266 (MSL).

33.406 THE IMPACT OF BINOCULAR DEPTH CUES AND MOVEMENT DIRECTION ON GRASPING AND PLACEMENT BEHAVIOURS

Nathaniel Goldstein[1] (nathany@my.yorku.ca), Laurie M. Wilcox[1], Erez Freud[1]; [1]York University, Toronto, Canada

We typically rely on binocular vision to understand the spatial layout and physical properties of objects in the visual world. This information is essential both for perception and action-based behaviours and there is evidence that when adults have access to binocular depth cues, they make faster and more accurate grasps towards objects. However, most studies examining this topic have used tasks in which participants reach for objects at a distance and place them closer towards them. Conversely, real-world movements occur in different directions and distances relative to the observer. Here we evaluate the contribution of binocular depth information (e.g., disparity, vergence) to a range of grasp characteristics for bidirectional movements (i.e., towards vs. away from the body). In a within-subjects design, participants grasped 3D discs of varying sizes (3.5 - 5.5 cm diameter) at two distances (18 and 36 cm from observer). Viewing was binocular or monocular with the non-dominant eye patched. On each trial, participants grasped an object and placed it on a peg positioned closer or further away from their body as quickly and accurately as possible. Our results show that binocular depth information improves both components of the task (grasping the object and positioning the object at the new location). In particular, under the binocular condition the velocity of the grasping movement was higher and the time to reposition the objects shorter. These effects were consistent for both movement directions. We observed minimal interactions between depth, distance, and size, suggesting mostly independent effects of these variables on visuomotor behaviour. Additional analyses of the grasp trajectories, including machine learning approaches, will provide insight into the contribution of 3D cues to multi-dimensional visually guided movements.

Natural Sciences and Engineering Research Council of Canada (NSERC); CFREF Vision: Science to Applications (VISTA)

33.407 REDUCED SENSITIVITY TO SIZE BUT PRESERVED PERCEPTION-ACTION DISSOCIATION IN OLDER ADULTS.

Felicia Tassone[1] (feliciat@yorku.ca), Zoha Ahmad[1], Tzvi Ganel[2], Erez Freud[1]; [1]York University, [2]Ben-Gurion University of the Negev

According to the two visual pathways hypothesis, vision-for-perception and vision-for-action are mediated by different neural pathways. While this dissociation has been well-studied in younger adults using psychophysics and neuroimaging, it is unclear whether healthy aging modulates the dissociation across the lifespan. To bridge this gap, we examined the effect of aging on visuomotor and perceptual tasks in two experiments, both completed by older (n = 26, range: 60–95 years) and younger adults (n = 26, range: 18–25 years). Experiment 1 aimed to test for potential differences among perceptual and visuomotor sensitivity to object size using “Efron blocks”, which vary in width and length but have an identical surface area. In the two conditions, participants were asked to perceptually estimate the width of the blocks using their fingers or to grasp the objects across their width. Experiment 2 utilized the Ponzo illusion to explore potential differences in the dissociation between perception and visuomotor control between younger and older adults. Participants completed the perceptual and grasping conditions with objects placed on the “close” and “far” surfaces of the illusion. The results of Experiment 1 showed an age-related decrement in size sensitivity, which was more strongly evident in the grasping task. In Experiment 2, both groups demonstrated a dissociation between perception and action, such that greater perceptual estimations were observed when objects were placed on the “far” surface of the illusion, while this effect was not observed during the grasping task. Unexpectedly, older adults even exhibited a reversed effect during grasping, with larger apertures for objects placed on the “close” surface. These findings suggest that while visuomotor abilities decline with age, the dissociation between perception and action is largely preserved.

Natural Sciences and Engineering Research Council of Canada (NSERC); CFREF's "Vision: Science to Applications" (VISTA)

33.408 NATURALISTIC VISUOMOTOR BEHAVIOURS REVEAL REDUCED HANDEDNESS LATERALIZATION IN AUTISM.

Emily Fewster[1] (emily.fewsterr@gmail.com), Bat-Sheva Hadad[2], Erez Freud[1]; [1]Department of Psychology, Centre for Vision Research, York University, [2]Department of Special Education, University of Haifa

Autistic individuals often exhibit differences in perceptual and visuomotor functioning. One account for these alterations suggests that autism involves reduced specialization of cortical systems. In the current study, we investigate the extent to which handedness, one of the most robust markers of cerebral lateralization, is modulated in autistic individuals. While previous studies already demonstrated more cases of left-handedness and reduced right-handed specialization in autism, those studies primarily relied on self-reports and questionnaires. Here, we employed a naturalistic task to explore this topic. Autistic and non-autistic right-handed participants (27 in each group) recreated five Lego models using blocks placed on a standardized tabletop. This design allowed us to capture dynamic, real-world visuomotor behaviours and explore how autistic individuals use their right and left hands to explore and act on their reaching space. Our results revealed key group differences. First and most importantly, autistic participants displayed a lower proportion of right-hand grasps, suggesting reduced lateralization. Second, we observed differences in the use of 3D space, with autistic participants showing a stronger preference for blocks closer to their hands, suggesting larger safety margins in visuomotor interactions with the surrounding 3D environment. Finally, autistic participants completed the task slower than non-autistic participants, indicating differences in motor efficiency. Taken together, these findings demonstrate reduced specialization of hand use in autism, which may contribute to challenges and differences in visuomotor control.

33.409 ADAPTATION TO HAPTIC TARGET SIZE HIGHLIGHTS THE HIERARCHICAL NATURE OF GRASP PLANNING

Robert Whitwell[1], Alice Tan[1], Ana Victoria de Meira[1], Michelle Wong[1]; [1]The University of Western Ontario

We used a grasp-adaptation paradigm to test two models of grasp planning: a ‘classic’ model in which the target is coded separately as size and position, and an integrated model that postulates the target is coded as a set of egocentric ‘grasp points’. Participants (N=144) reached, without visual feedback, for virtual targets and grasped real (haptic) ones whose sizes were either the same, larger, or smaller. Two adaptation sequences were administered, each comprised of a series of baseline, adaptation, and then washout trials. On baseline and washout trials, the virtual and haptic objects were identical. On adaptation trials, the target's haptic size was either always larger or always smaller than its virtual size. Consistent with prior work, we observed strong aftereffects in the first (control) sequence. We administered a second sequence to test if the aftereffect would generalize to a novel target orientation and/or position while preserving target size. Notably, the classic model implies that the aftereffect should generalize under such circumstances. In contrast, the integrated model implies the aftereffect should fail to generalize, because a novel target position and/or orientation requires a novel set of grasp points. Surprisingly, generalization depended on the direction of the grasp aperture’s induced adjustment and, therefore, neither model can fully accommodate our findings. Specifically, the aftereffect induced by the larger haptic object generalized to novel target position and/or orientation, whereas the aftereffect induced by the smaller haptic object did not. In the latter condition, novel target orientations resulted in the strongest attenuation of aftereffects. Considering the asymmetric consequences of under- vs. over-sizing grasp aperture, our findings suggest grasp planning prioritizes haptic size when compensating for an under-sized grasp aperture. When under-sizing is not applicable, however, haptic integration is subsumed under de novo coding of the orientation of the hand opposition space.

This work was supported by an NSERC Discovery Grant to RLW (RGPIN-2022-05050).

33.410 PERCEIVING THE AFFORDANCE OF INTERCEPTABILITY FOR OTHERS

Frank Zaal[1] (f.t.j.m.zaal@umcg.nl), Samruddhi Damle[1], Reinoud Bootsma[2]; [1]University Medical Center Groningen, The Netherlands, [2]Aix-Marseille Université, CNRS, France

Knowing whether certain actions are (still) possible for others is often relevant. Using a lateral-interception task, we studied the affordance of interceptability for others. Balls moving down from the top of a 55" screen, following linear trajectories, could be intercepted with a "paddle" that could move horizontally along an interception axis near the bottom of the screen. We varied the balls' horizontal departure positions, arrival positions, and flight times. Two groups ("actors" and "observers") of 12 participants each completed three conditions: a "training", an "action", and a "judging" condition. In the "action" condition, the participants were instructed to try to intercept the virtual balls even when they felt that this would be impossible. In the "judging" condition, the actors also tried to make an interception but when they perceived that the interception would not be possible (anymore), they indicated this by calling "no" and were free to abandon their interception attempt. The actors successfully intercepted the ball in 80% of their "action" trials and indicated that an interception was perceived to be impossible in 12% of a set of similar "judging" trials. In their "judging" condition, the observers were shown the played-back kinematics of ball and paddle for a selection of the trials from the actors' "judging" condition, with an even distribution of interceptable and uninterceptable balls. The observers were instructed to indicate with a button press whether or not a ball would be interceptable for the original player (i.e., the actors). They turned out to be correct on 72% of the trials. We conclude that people are able to perceive the affordance of interceptability both for themselves and for others.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie SkłodowskaCurie grant agreement No 956003.

33.411 DISSOCIATING THE ROLES OF SPATIAL POSITION WITHIN THE VISUAL FIELD AND THE IMPLIED DESTINATION OF A MOVING OBJECT IN THE PERCEPTUAL-MOTOR CONGRUENCE EFFECT KNOWN AS THE SIMON EFFECT

Kade T. Tanke[1] (kadetanke@hotmail.com), Michael L. Paavola[1], J. Toby Mordkoff[1], Cathleen M. Moore[1]; [1]University of Iowa

When responding to a non-spatial attribute of a stimulus such as its color, the spatial position of the stimulus affects the response, despite being task irrelevant. Specifically, responses tend to be faster and more accurate when the stimulus is on the same side of the responding effector (congruent) than when they are on different sides (incongruent). This effect, known as the Simon Effect, reflects processing interactions between spatial codes associated with the response and spatial codes associated with the stimulus, and has been used to study the relationship between motor affordances and visual processing. We asked whether the relevant spatial code of the stimulus is position in the visual field or whether in the case of a moving stimulus, its apparent destination, which is especially relevant to visuomotor coordination. Using virtual reality, we presented spheres that originated on the left or the right side of a virtual environment and moved toward the participant in a trajectory that headed either to the hand on the same side (straight motion) or to the hand on the opposite side (crossed motion). The spheres began as gray and changed to either blue or orange after 200ms. The task was to report the color as quickly and accurately as possible by making a left or right manual trigger response. Trials on which responses were made after the sphere crossed the midline were excluded from analyses so that regardless of the trajectory condition, stimuli appeared only on the side of origin. Responses were faster when the apparent destination of the moving sphere was the responding hand than when it was the opposite hand. This was true regardless of the side of origin. These results are consistent with the Simon Effect reflecting interactions associated with perceptual-motor coordination.

33.412 SPATIAL AFFORDANCES TRIGGER SPONTANEOUS VISUAL PERSPECTIVE-TAKING EVEN IN THE ABSENCE OF OTHER AGENTS

Robert Walter-Terrill[1], Brian Scholl[1]; [1]Yale University

A central goal of vision is to recover information about local environments that is not tied to a single perspective ("What does it look like from here?"), but is generalizable ("What's out there?"). Most work on such themes involves either deliberate imagery ("What would it look like from over there?") or social cognition -- as when taking another agent's perspective ("What would it look like from her shoes?"). Here, in contrast, we show that spontaneous visual perspective-taking occurs even in the absence of other agents -- triggered just by the spatial affordances of the environment itself. Observers viewed two buttons on a table and pressed keys with their right or left hands to indicate when those buttons haphazardly changed to particular colors. The two buttons were vertically aligned from the observers' viewpoint (one closer, one further away), such that adopting a perspective from the left or right side rendered responses spatially congruent or incongruent (as in the Simon effect). Critically, this different perspective was afforded by an environmental regularity in the scene that influenced whether the buttons were reachable. For example: (1) One side of the table had a chair, with the other side flush against a wall. (2) One side of the table had a normally-oriented chair, while the other had a backwards-facing chair. (3) One side of the table had no obstruction, while the other had a translucent screen. Or (4) one side of the table was on solid ground, while the other stood over a sheer cliff face. Each case yielded robust spontaneous visual perspective-taking: despite the task-irrelevance of these manipulations, observers responded faster when the afforded perspectives yielded congruent spatial button/response mappings. These results show how spatial affordances spontaneously promote a type of generalizability during scene perception, and how visual perspective-taking does not require other agents.

This project was supported by the Wu Tsai Institute, Yale University.

Action: Perception and recognition

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.413 THE EFFECT OF ACTION ON VISUAL FALSE PERCEPTS

Juan Carlo Cabato[1] (carlo.cabato@psychiat.med.uni-giessen.de), Yannik Heidelbach[1], Gizem Yildiz[1], Katja Fiehler[1], Bianca van Kemenade[1]; [1]Justus Liebig University Giessen, Germany

Hallucinations, a primary symptom of schizophrenia, are commonly defined as the perception of sensory information despite the lack of corresponding sensory input. Hallucination-like experiences can be studied in signal detection tasks, in which participants have to detect signals in noise. Indeed, patients with schizophrenia report more false alarms than healthy controls, suggesting a general perceptual deficit. Another proposed mechanism underlying hallucinations is the disruption of efference copy signaling. The efference copy is a copy of the motor command, which is used to anticipate the sensory consequences of self-generated actions. These predictions usually lead to perceptual and neural suppression of the sensory action outcomes. Patients with schizophrenia show less suppression of self-generated stimuli than healthy controls. As such, it has been suggested that dysfunctional efference copy mechanisms are the source of hallucinations in schizophrenia. It is still unclear whether psychotic symptoms can be attributed to aberrant efference copy mechanisms, or to dysfunctional perceptual mechanisms in general. To investigate this question, this study employed a signal-detection task with gratings obscured in dynamic noise under both active (self-generated) and passive (externally generated) conditions. Furthermore, participants filled out questionnaires testing for hallucination-proneness. Results showed that false alarms were more frequent in the active condition relative to the passive condition, driven by an increased response bias in the former while perceptual sensitivity (d') did not significantly differ. Interestingly, the number of false alarms correlated positively with schizotypy scores. Overall, these results show that people are more likely to believe they perceived a stimulus in noise when they generated the stimulus themselves, possibly due to an increased sense of agency. Nevertheless, people scoring high on schizotypy showed a general tendency to report more false alarms, suggesting that schizotypy may be related to dysfunctional perceptual mechanisms that are not specific to action.

Deutsche Forschungsgemeinschaft (DFG), SFB/TRR135 Project A10

33.414 CODING VISUO-SPATIAL INFORMATION IN A VIBROTACTILE BELT: PERCEIVED EGOCENTRIC DIRECTION FROM PATTERNS OF VIBRATION

Sina Feldmann[1], Qiwu Zhang[2], Chiang-Heng Chien[2], Brian Free[1], Benjamin B. Kimia[2], William H. Warren[1]; [1]Department of Cognitive & Psychological Sciences, Brown University, [2]School of Engineering, Brown University

How might we design more useful assistive devices for individuals with impaired vision? Although traditional tools such as the long cane are simple and effective for short-range guidance (2-3 steps), there are few practical aids for mid-range navigation and collision avoidance over intermediate distances. We are investigating vibrotactile belts as a potential sensory substitution interface, and are currently comparing alternative methods of encoding visuo-spatial information in patterns of vibration. In the present experiment, we tested perceived egocentric direction from single-tactor and distributed (Gaussian) patterns of vibration. Participants (N=16) wore a vibrotactile belt containing 16 pager motors spaced 22.5˚ apart, with an Arduino controller. Eight directions at 45-degree intervals were stimulated for 2 seconds at a fixed intensity (225 Hz), and participants indicated the perceived direction of vibration by clicking on a circle surrounding an icon of a person on a computer screen. Three vibration patterns were compared: (a) a single motor, (b) 3 adjacent motors in a narrow Gaussian distribution (spanning 45˚), (c) 3 or (d) 5 motors in a wide Gaussian distribution (spanning 90˚). Each stimulus was repeated ten times, randomized within each vibration block. The results show that the mean reported direction was linear and highly accurate (mean constant error = 0.35°). The mean variable error depended on direction, with the smallest within-subject SD in the anterior and posterior directions (mean SD = 6.22°) and the largest in the lateral directions (mean SD = 13.77°) (see also Cholewiak & Schwab, 2004). Furthermore, responses were unaffected by the vibration pattern. We conclude that single-tactor vibrations are sufficient to specify egocentric direction quite accurately, simplifying the encoding and reducing the controller computation. Next, we plan to evaluate different encoding strategies to guide walking to a series of spatial targets.

This research is supported by NIH R01EY029745.

33.415 ACTIVE MANIPULATION PROMOTES PREDICTIVE GAZE STRATEGIES DURING VIRTUAL OBJECT EXPLORATION

Cassandra W. Bacher[1], Karsten Babin[1], Marieke Mur[1], Melvyn A. Goodale[1], Jody C. Culham[1]; [1]Unversity of Western Ontario, London, Ontario, Canada

Actively manipulating novel objects rather than passively observing them can facilitate object recognition, but the mechanisms behind this effect still remain largely unknown. One potential explanation is that active learning facilitates an iterative process of generating and testing hypotheses about the effects of actions on the objects. We investigated how the gaze strategies utilized during active learning differ from those utilized during passive learning. Of particular interest was the question of whether participants would display any preference for the "leading edge" (the edge of a rotating object where previously unseen features are becoming visible) or the "trailing edge" (the edge of a rotating object where previously visible features are disappearing from view) of an object and whether such a preference would differ between learning conditions. Participants (n = 26) learned novel virtual objects either actively or passively while eye tracking data was

collected, then completed an old/new discrimination task. Results indicated that during active learning, participants spent significantly more time looking at the leading than trailing edge (53%:47%). In contrast, participants spent equal amounts of time looking at the leading and trailing edges during passive learning (50%:50%). These results suggest that people may indeed be more likely to generate and test hypotheses about objects that are being learned actively rather than passively. However, no effect of active versus passive learning on recognition accuracy or speed was observed, likely because the objects we used had less distinctive parts than those used in prior studies did (Harman et al., 1999; Curr Biol). Thus, further investigation is needed to determine what role such predictive behavior plays in the "active learning effect" observed in past studies.

Natural Sciences and Engineering Research Council of Canada (04271-2022-RGPIN)

33.416 VISUAL FEATURES FOR EGOCENTRIC ACTION RECOGNITION

Filip Rybansky[1], Sadegh Rahmani[2], Andrew Gilbert[2], Frank Guerin[2], Anya Hurlbert[1], Quoc Vuong[1]; [1]Newcastle University, [2]University of Surrey

People quickly recognize daily actions (e.g., washing dishes) in visual stimuli. Evidence suggests that Minimal Recognizable Configurations (MIRCs) contain spatial and temporal features for reliable recognition (Ben-Yosef et al., 2020) and could be used to identify important features to improve computer vision. To further investigate the contribution of MIRCs to action recognition, we progressively reduced the available spatial information in egocentric action videos. We selected 18 videos from the Epic-Kitchens-100 dataset (Damen et al., 2022) which were correctly categorised (Easy) by our computer vision network (Ahmadian et al., 2023), and 18 incorrectly categorised (Hard). Participants (N=3800) viewed Easy and Hard videos online and identified the action. Videos recognized by ≥50% of the participants were cropped to produce four reduced quadrants for the next data collection round. A spatial MIRC was identified if all subsequent quadrants (sub-MIRCs) were not recognized. Videos were reduced until we obtained at least one spatial MIRC for each video. We tested 7604 video quadrants, identifying an average of 15.17 spatial MIRCs per video. In terms of the recognition gap, which is the difference in accuracy between MIRC and sub-MIRC quadrants, we extended Ben-Yosef et al.'s findings to complex egocentric action videos. Recognition gaps were significantly greater for Hard (Md=0.40, IQR=0.25) than Easy (Md=0.35, IQR=0.20) videos, suggesting that feature importance is concentrated into a smaller subset of features in Hard than Easy videos. SHAP importance analysis of surface-area features for recognizability also supported this claim and indicated that visibility of the active hand and active object are important for action recognition. Finally, Graph-Based Visual Saliency (Harel et al., 2006) was significantly greater in MIRCs (Md=0.32, IQR=0.19) vs. randomly located size-matched quadrants (Md=0.26, IQR=0.19), and in Hard (Md=0.33, IQR=0.20) vs. Easy (Md=0.31, IQR=0.19) MIRCs. Our results suggest that MIRCs reveal the important features for egocentric action recognition.

Leverhulme Trust

33.417 VARIANCE PERCEPTION AFFECTS THE SENSE OF AGENCY, BUT NOT VICE VERSA

Yunsoo Yeo[1], Sang Chul Chong[1]; [1]Yonsei University

When we interact with objects, we may experience a sense of initiating and controlling certain events. This feeling is referred to as the sense of agency. Previous research has shown that the sense of agency is not determined by a single mechanism, but rather depends on various sources of information. Also, recent studies suggest that it influences our other cognitive processes such as visual attention and memory. In this study, we focused on variance perception as another potential factor that might relate to the sense of agency. Item variance reflects uncertainty and volatility in the environment, both of which are closely linked to the sense of agency. Across three experiments, we tested whether the sense of agency can affect (Experiments 1 and 2) and be affected by (Experiments 2 and 3) perceived variance. In each experiment, participants were asked to move 40 differently colored circles with or without a 1-second delay, which was introduced to reduce the sense of agency. They then reported either perceived variance of the stimuli (Experiments 1), their sense of agency over the stimuli (Experiment 3), or both (Experiment 2). We found that while the sense of agency did not influence perceived variance, item variance had a negative impact on the sense of agency. However, this effect was observed only when participants reported both perceived variance and the sense of agency on each trial (Experiment 2). The effect disappeared when they only reported the sense of agency (Experiment 3), where the variance of the stimuli became task-irrelevant. Altogether, the present study suggests that variance perception can influence the sense of agency in a unidirectional manner, but this effect is modulated by the task-relevance of variance information.

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(RS-2022-NR070542).

33.418 ACTION BOOSTS VISION, ESPECIALLY IN TOUGH TASKS

Gizem Y. Yildiz[1] (gizem.yildiz@psychiat.med.uni-giessen.de), Elena Witzel[1,2], Katja Fiehler[2], Bianca M. van Kemenade[1]; [1]Center for Psychiatry, Justus Liebig University, Giessen, Germany, [2]Experimental Psychology, Justus Liebig University, Giessen, Germany

Every action we perform triggers the brain to generate predictions using copies of motor commands, known as efference copies. These action-predictions are then compared with incoming sensory information, shaping our perceptions. Action-predictions usually lead to sensory attenuation. However, in the auditory domain, it has been shown that while attenuation occurs for clearly audible action outcomes, this effect reversed to enhancement for near-threshold stimuli. It remains unclear whether similar effects occur in the visual domain. To investigate whether action-predictions influence visual perception differently based on stimulus visibility, we asked participants to perform an orientation discrimination task on tilted gratings presented at supra-threshold (high-visibility) and sub-threshold (low-visibility) levels under active and passive viewing conditions. In the active viewing conditions, participants had to press the button indicated by an auditory cue in order to elicit the presentation of the gratings. In the passive viewing conditions, the gratings were presented automatically following the auditory cue. Experiment 1 demonstrated that both the discrimination thresholds and the slopes were affected by the visibility of the grating. The discrimination thresholds were lower, and the slopes were higher in the high-visibility trials compared to the low-visibility trials. Although no differences were observed between the active and passive viewing conditions, post-hoc analysis indicated a trend toward significance in the accuracy difference between the two conditions, specifically in challenging comparison trials within high-visibility blocks. In Experiment 2, we manipulated task difficulty to create blocks of challenging and easy comparisons within high- and low-visibility conditions under active and passive viewing. The results showed that participants were significantly more accurate in the active trials than in the passive trials of high-visibility blocks, especially when the task was more difficult. We conclude that action-predictions enhance perception primarily in situations where visual stimuli are less noisy, but the task is more challenging.

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–SFB/TRR 135, Project A10

33.419 LATERAL OCCIPITO-TEMPORAL AND PARIETAL AREAS ENCODE DIFFERENT FEATURES DURING NATURAL ACTION PERCEPTION

Diana C Dima[1,2], Jody C Culham[1], Yalda Mohsenzadeh[1,2]; [1]Western University, [2]Vector Institute for Artificial Intelligence

In everyday life, we rapidly recognize and react to many actions performed by others despite variations in the setting, agents, and objects involved. How are everyday actions processed in the ventral “vision-for-perception” and dorsal “vision-for-action” streams? Human neuroimaging research has found representations of perceptual features and action goals in the lateral occipito-temporal cortex (LOTC), while primate research suggests that parietal areas encode action classes with shared motor goals. Here, we collected fMRI data from 20 participants while they viewed two-second videos of everyday actions and read sentences describing the same actions. Representational similarity analysis on neural patterns revealed LOTC representations of effectors (e.g., leg or arm), action class (e.g., locomotion or manipulation), and action target (e.g., an object, the self, or another person). In contrast, univariate analyses on activation levels revealed differences among action classes in parietal cortex, especially the superior parietal lobule (SPL). Cross-decoding of multivariate representations across videos and sentences revealed that only LOTC, but not parietal cortex, encoded modality-invariant responses to actions. Together, our results reveal different representations of natural actions in the ventral and dorsal streams.

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.

Eye Movements: Saccades, remapping

SUNDAY, MAY 18, 8:30 AM – 12:30 PM, PAVILION

33.420 SUPPRESSED LUMINANCE STIMULI ALTER THE REPRESENTATION OF THE SACCADE TARGET

Golnaz Forouzandehfar[1] (gforouz@vols.utk.edu), A . Caglar Tas; [1]University of Tennessee, Knoxville

Saccadic suppression refers to the decreased sensitivity to visual information during saccades. Previous studies showed that high-luminance contrast stimuli that are presented during the saccade typically are not detected (e.g., Knöll et al., 2011). Although saccadic suppression reduces conscious awareness of transient visual events, suppressed stimuli may still influence perceptual representations. The present study investigates whether suppressed luminance stimuli affect perception of the saccade target. Participants (N=4) executed a saccade to a colored target on a gray background. Brief luminance flashes (6% or 24% contrast levels with the background) were presented equidistant from the fixation and the target at various temporal offsets relative to saccade onset. Participants reported whether they perceived the luminance target and the color of the saccade target on a color spectrum with the same hue but varied luminance. The results showed that participants failed to perceive the luminance target, replicating previous work. Importantly, they chose brighter colors compared to when no luminance flash was presented. Comparison with the no-saccade control condition also showed perception of a brighter target when a saccade was executed. These results highlight the role of suppressed stimuli in shaping perceptual representations. When the luminance target was presented during the saccade, it biased remembered brightness of the saccade target even though its presence did not reach to awareness.

33.421 PRE-SACCADIC SUPPRESSION IS REDUCED IN THE ANTI-SACCADE TASK

Matthew Smith[1], Chris Scholes[1], Neil W Roach[1]; [1]University of Nottingham

Our ability to detect visual stimuli is suppressed immediately before and during saccadic eye movements, but the mechanisms underlying this change in sensitivity are not fully understood. One proposal is that saccadic suppression is driven by extra-retinal signals generated during motor preparation which act to cancel the perceptual consequences of rapid eye movements. In primates, neuronal activity in superior colliculus prior to a saccade has been identified as a candidate extra-retinal signal. Pre-saccadic activity in superior colliculus build-up activity is stronger before pro-saccades (where the eyes move towards a target) than before anti-saccades (where the eyes are instead diverted away from a target). Here, we asked whether behavioural suppression, measured in humans, also differs between the two saccade types. To test this, we employed a pro/anti-saccade task while measuring participants’ contrast sensitivity for detecting a transient horizontal grating presented in the upper or lower half of the screen. The time course of saccadic suppression was characterized by flashing the grating at three different time points relative to an individual's median saccadic reaction time and exploiting natural variations in saccadic reaction time to sample the entire perisaccadic time window. We found that suppression differed between the saccade types in the pre-saccade window (from 50ms before the eye starts to move up to saccade onset), with suppression beginning significantly later for anti-saccades than for pro-saccades. This effect is not explained by differences in saccade amplitude or peak velocity. Given that our experimental paradigm was designed to minimize passive suppression mechanisms (e.g., blur, masking), our results suggest that active saccadic suppression is altered by behavioural context. Changes in behavioural suppression may be linked to differences in superior colliculus build-up activity for pro- and anti-saccades.

33.422 ADAPTING SACCADES IN OPPOSITE DIRECTIONS REVEALS INTER-HEMISPHERIC INTERFERENCE IN MOTOR LEARNING

Patrik Polgári[1] (patrik.polgari@uni-marburg.de), Alexander C. Schütz[1]; [1]University of Marburg

Position errors after saccades lead to adaptive changes in their programming. Two underlying mechanisms have been described, relating to the direction of error: Backward adaptation i.e., shortening of amplitudes, is stronger and has a quicker dynamic profile than forward adaptation i.e., lengthening of amplitudes. Moreover, saccade adaptation is described as direction-specific, meaning no transfer of adaptation from leftward to rightward saccades. However, classic studies adapted saccades only in one direction and from the same initial starting point. This paradigm limits a potential transfer of error information between adaptation mechanisms acting in opposite directions. In the first experiment, we tested whether adaptation in one saccade direction can interact with adaptation in the opposite direction. In a horizontal random walk paradigm, we simultaneously adapted observers' leftward and rightward saccades in three conditions: (1) both-backwards, (2) both-forwards, and (3) backwards in one direction and forwards in the other ('mixed'). We found an asymmetrical interplay between backward and forward adaptation occurring in opposite directions. While forward adaptation was similar in the 'both-forwards' and the 'mixed' conditions, backward adaptation was stronger in the 'mixed' condition compared to the 'both-backwards' condition. In the second experiment, we sought to discern whether these asymmetrical effects reflect backward adaptation being facilitated by forward adaptation or impaired by backward adaptation in the opposite direction. In three conditions, we combined backward adaptation in one direction with different adaptations in the opposite: (1) forward ('mixed'), (2) backward ('both-backwards'), and (3) no adaptation. We found reduced backward adaptation when coupled with backward adaptation in the opposite direction compared to forward or no adaptation. This replicates findings of the first experiment and suggests that simultaneous backward adaptation in opposite directions poses a particular challenge to the oculomotor system. Generally, saccade adaptation in opposite directions may not be as independent as previously assumed.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project Number 222641018 - SFB/TRR 135 TP B2.

33.423 UNDERSTANDING VISUAL PROCESSING DURING SACCADES USING ZEBRAFISH

Lisa Bauer[1], Joseph C. Donovan[1], Herwig Baier[1]; [1]Max Planck Institute for Biological Intelligence

Saccadic eye movements are fundamental to vertebrate visual perception, yet it is still unclear exactly how neurons handle abrupt shifts in the visual scenery. The zebrafish model, with its experimental tractability, offers an ideal system for investigating the underlying sensorimotor circuits. Zebrafish perform spontaneous and visually induced saccades as early as 4 days post fertilization (dpf). Here, we combined eye tracking and two-photon calcium imaging to investigate the neuronal correlates of saccades in larval zebrafish (6-8 dpf). We focus on the optic tectum (OT), the fish homolog of the mammalian superior colliculus (SC), as well as the largest visual brain area. Using a two-photon microscope custom-modified with a remote focusing path to enable rapid multi-plane imaging, we record single-cell resolution neuronal activity across the OT at 5 volumes per second. Even in the absence of visual stimuli we find neurons in the OT that showed increased activity correlated with spontaneous saccades. To investigate how visual stimuli are integrated, we recorded from the same neurons during various visual stimulus paradigms. Our findings reveal that most spontaneous saccade-correlated neurons in the OT respond similarly regardless of visual environment. Moreover, spontaneous saccade-responding neurons are a subset of visually induced saccade neurons. Notably, while many neurons were active around the time of a saccade, certain neurons' activity peaked before, suggesting a role in anticipatory motor planning. Our results underscore the effectiveness of the larval zebrafish as a model for functional investigation, enabling experimental approaches that are challenging to implement in primate models.

33.424 DYNAMIC READ-OUT OF TRANSSACCADIC ATTENTIONAL REMAPPING FROM EEG ALPHA BAND OSCILLATION

Tzu-Yao Chiu[1] (chiu.315@osu.edu), Zitong Lu[1], Julie D. Golomb[1]; [1]The Ohio State University

Humans experience the world as stable despite saccade-induced displacements in retinal input. To maintain stability across saccades, the focus of spatial attention needs to be remapped, such that attention at the previous retinotopic (eye-centered) location is removed and updated to reflect the spatiotopic (world-centered) location. Previous behavioral studies showed mixed evidence for both pre-saccadic predictive remapping of attention and retinotopic attentional trace lingering after saccade completion. Here, we used multivariate decoding of EEG alpha-band oscillation to investigate the timecourse of attentional remapping in the peri-saccadic time window. Subjects performed a spatial memory task, in which they memorized a cued peripheral location during sustained fixation periods (no-saccade trials) or across a saccade, with eye-tracking and EEG recordings. Fixation location, saccade direction, and memory cue location randomly varied between trials. Subjects participated in two sessions: one in which they reported the memory cue's spatiotopic location and another in which they reported its retinotopic location. We trained classifiers to decode the location of spatial attention based on alpha-band oscillation on no-saccade trials, and tested them time-by-time on saccade-trial data to track dynamics of attentional allocation in the peri-saccadic period. In the spatiotopic memory task, we found reliable decoding of the memory cue's retinotopic location until 400ms after saccade offset, and decoding of the task-relevant spatiotopic location emerging late after saccade completion, consistent with the previously observed retinotopic attentional trace and gradual ramp up of spatiotopic attention. Furthermore, in the retinotopic memory task, we found significant and prolonged decoding of the task-relevant retinotopic location and no spatiotopic decoding, indicating task-dependent remapping of spatiotopic attention. Overall, the current results based on human non-invasive neural recordings provide converging evidence for the gradual remapping of spatial attention after saccades. The current approach further opens ample opportunities to investigate roles of stimulus factors and cognitive control in remapping.

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

33.425 PERCEPTUAL COMPLETION SUPPORTS TRANSSACCADIC OBJECT CORRESPONDENCE

Fabian Parth[1], Alexander C. Schütz[1,2]; [1]Sensorimotor learning unit, University of Marburg, Germany, [2]Center for Mind, Brain and Behavior, Marburg & Gießen, Germany

Saccadic eye movements bring objects from the peripheral visual field to the fovea for further inspection. To combine presaccadic with postsaccadic information, specific features—such as shape, texture, displacement, size, and contrast—are used to establish transsaccadic object correspondence (TOC). However, during saccades, external factors like partial occlusion can alter object information, raising the question of how the visual system copes with such changes and whether perceptual completion supports TOC. We conducted a postsaccadic target selection task to explore the dynamics of TOC and examine whether observers can establish TOC with partially occluded and amodally completed objects. Observers initiated a horizontal saccade towards a presaccadic peripheral object (circle or pacman), which was replaced perisaccadically by two vertically displaced candidates. These candidates were either fully visible and identical to the presaccadic object, partially occluding a triangle, or partially occluded by the triangle. Various combinations of candidates were tested to isolate the effects of shape, occlusion, configuration complexity, and other factors. At the end of each trial, observers had to indicate which postsaccadic candidate matched the presaccadic object. Behavioral responses were almost perfect and observers performed corrective saccades to the corresponding candidates with an accuracy of 80%. This indicates that TOC was possible even with dynamic changes in the available object information. Notably, corrective saccade latencies were longer when targeting candidates overlapping with the triangle, reflecting a cost for processing more complex configurations. Furthermore, latencies were shorter when perceptual completion facilitated the detection of the partially occluded target, compared to configurations where completion hindered selection. These findings suggest that perceptual completion supports TOC by resolving ambiguities in scenarios with partial occlusion. This highlights the visual system’s ability to integrate incomplete information and underscores the role of perceptual completion in maintaining object continuity across saccades.

FP and ACS were supported by Deutsche Forschungsgemeinschaft (DFG), Project number 290878970-GRK 2271/Project 5.

33.426 PERCEPTUAL AND OCULOMOTOR DISSOCIATION OF ASSIMILATION AND CONTRAST EFFECTS OF OPTIC FLOW

Hongyi Guo[1], Alexander Schütz[2], Robert Allison[1]; [1]York University, [2]Marburg University

The motion of targets needs to be taken into account for saccades to land accurately on the target. In an optic flow field, a saccade directed toward a briefly presented stationary target is biased in the direction of radial and rotational optic flow (OF), an effect called assimilation of optic flow (Guo, Schütz & Allison, 2024). We observed that this assimilation disappeared when we tested saccade adaptation to intra-saccadic steps. At least two possible causes were proposed: (1) Saccade adaptation diminished the influence of optic flow, and (2) The visible target after the saccade provided a stationary prior. Two experiments with OF were conducted to test these explanations. Experiment 1: subjects performed saccades to a target that remained visible after the saccade. Experiment 2: subjects performed saccades to a target moving at different speeds left or right, or a stationary target and reported the perceived motion direction of the target. In experiment 1, saccades were smaller while viewing inward OF than no OF; conversely, they were larger with outward OF than no OF. Thus, we replicated the assimilation effect, now when the target remained on screen after the saccade. This means the lack of assimilation during intra-saccadic step adaptation cannot be explained by postsaccadic information about the target. Preliminary results of experiment 2 showed that the perceived motion direction of the stationary target was repelled from the optic flow direction, consistent with the well-known induced motion illusion. Surprisingly, saccade amplitudes were still assimilated in the direction of radial optic flow. Our results show a dissociation in motion perception and eye-movement behaviour, which might indicate separate channels or different reference frames for processing optic flow for perception and behaviour.

This study was supported by the IRTG The Brain in Action program, funded by NSERC and DFG.

33.427 THE DEPENDENCE OF SACCADIC REACTION TIME ON THE PRIOR PRESENCE OF VISUAL TARGETS

Jay A Edelman[1], Amy Sultana[1]; [1]The City College of New York, Dept. of Biology

Saccadic reaction times (SRTs) are normally in the range of 150-200ms when a visual stimulus is presented on an otherwise blank screen. Edelman et al (2023, 2024) have demonstrated that in an antisaccade (AS) task SRTs were ~30 msec longer when stimuli were present at the two possible target locations (10 deg to the left/right of central fixation) for several hundred msec prior to the peripheral visual stimulus triggering the saccade (peripheral imperative stimulus -- IS). These two studies showing such a prior presentation effect (PPE) used participants > 40 y.o. Moreover, the central fixation duration had an exponential distribution with a mean time of 1500 msec and ISs were small in size. We determined whether a PPE was evident 1) for prosaccades (PS) as well as ASs 2) when the fixation duration was short as well as long 3) for ISs of both large and small size and 4) when participants were young. We collected eye position data from eight undergraduate participants using an Eyelink 1000 video eyetracker (SR Research) at 1000 frames/sec. A three-factor ANOVA was run separately for PSs and ASs with factors of 1) prior target presentation (yes/no) 2) fixation duration (long & exponential / short & uniform) and 3) IS size (large – 3 deg square / small – 1 deg square). For PSs, we found very strong main effects of prior target presentation and fixation duration (p < 0.0001) driven by a PPE and shorter SRTs for long fixation duration, as well as a strong main effect of IS size with shorter SRTs for large stimuli (p=0.004). For ASs, we corroborated our earlier work, but, overall, there was a main effect only of fixation duration (p=0.004). Prior stimulus presence may attenuate the impact of suddenly appearing stimuli on prosaccades, and in certain circumstances, antisaccades.

PSC-CUNY grant

33.428 SPATIOTEMPORAL TASK PARAMETERS MODULATE MULTISENSORY RESPONSE ENHANCEMENT IN SACCADIC LATENCY

Jessica Chalissery[1,2] (jesschal@student.ubc.ca), Anthony Herdman[1,2,3,4], Miriam Spering[1,2,3,4,5], Philipp Kreyenmeier[1,2]; [1]University of British Columbia, [2]Graduate Program in Neuroscience, University of British Columbia, [3]School of Audiology, University of British Columbia, [4]Djavad Mowafaghian Center for Brain Health, University of British Columbia, [5]Institute for Computing, Information, and Cognitive Systems, University of British Columbia

Saccadic latencies are typically faster in response to multisensory as compared to unisensory stimuli. However, studies using different saccadic protocols result in varying degrees of multisensory response enhancement. This study investigates which spatial and temporal task parameter configuration yields the greatest multisensory response enhancement for saccadic latency. Human observers (n=9) made saccades to either visual targets (small black dot presented 10 degrees to the left or right from the screen center), auditory targets (burst of white noise played to the left or right ear through headphones), or combined, congruent audiovisual targets. We measured saccadic latency with the Eyelink1000 eye-tracker as indicators of sensory processing speed and to determine the efficacy of multisensory integration across three task configurations: (1) gap & placeholders, (2) no-gap & no-placeholders, and (3) no-gap & placeholders. As expected, a gap reduced saccadic latency to visual targets, whereas placeholders sped up saccades to auditory targets. Across all task configurations, we observed faster latencies to audiovisual targets compared to visual or auditory targets, indicating multisensory response enhancement (all p<0.001). To assess whether this enhancement was explained by statistical facilitation of the redundant audiovisual signal, we compared observers' audiovisual latencies to the upper bound of a race model based on the unisensory target conditions. Task configurations strongly modulated the degree to which audiovisual latencies outperformed the race model ($F(2,16)=6.42$, $p<0.01$) and showed that only the no-gap & placeholder-condition resulted in reliable and strong multisensory integration. Our findings demonstrate that the spatiotemporal parameters of a simple saccade task modulate the magnitude of multisensory enhancement on saccadic latency. These findings provide a reliable foundation to allow exploration of the underlying mechanisms of multisensory perception and orienting behavior.

33.429 SACCADE SYSTEM ACTIVATES PURSUIT TO COVER FOR ITS ERRORS

Krischan Koerfer[1], Markus Lappe[1]; [1]University of Münster

Often the saccade system and the smooth pursuit system have to work in conjunction for optimal gaze behavior. This is especially true for intercepting saccades to a moving target with consecutive pursuit. The extent to which the two systems share a control system and a common neurological representation of motion signals has been debated. Recently, we reported a type of non-rigid motion that cannot be pursued smoothy. Here, we show that the saccade system can briefly activate the pursuit system to compensate for the landing error of the saccade even in absence of any motion signal in the pursuit system. We presented a vortex motion moving horizontally at 5 or 10 degrees per second. Participants were instructed to fixate a red cross and then intercept and pursue the vortex once the red cross vanished. Despite the general inability to pursue the vortex, we measured a substantial pursuit activity within the first 200 milliseconds after the intercepting saccade. After this period, the pursuit gain dropped close to zero across participants. Analysis of the initial pursuit reveals striking trial-by-trial variability and a strong correlation with the saccade landing error. When saccades landed behind the target, pursuit gain was high, whereas when it landed ahead of the target pursuit gain was low or even negative (pursuit direction opposite to target motion). This suggests that the initial pursuit is not driven by the motion signal of the vortex but instead by the expected position error of the saccade. We propose that even in the absence of any motion signal in the pursuit system, the saccade system can activate the pursuit system to compensate for its shortcomings.

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

33.430 SUSTAINED DYNAMICS OF SACCADIC INHIBITION AND ADAPTATION DURING CONTINUOUS EXPLORATION

Antimo Buonocore[1] (antimo.buonocore@unisob.na.it), Celeste Cafaro[1], Alessio Fracasso[2]; [1]Suor Orsola Benincasa University, Italy, [2]University of Glasgow

In natural environments, stimuli often recur over time and space, requiring the visual system to adapt to predictable patterns while maintaining sensitivity to novel information. Within the visual system, neural responses to repeated stimuli at the same location in rapid succession often diminish (neuronal adaptation). However, it is less clear how these adaptation mechanisms persist over longer timescales as behavior evolves, and whether different modalities adapt differently. To address this, we developed a free-viewing paradigm in which participants explored a screen displaying simple shapes, such as squares and circles, for about eight seconds and compared their relative quantities. During visual inspection, a black circle was flashed five times consecutively, either in the fovea or the periphery, four degrees left of the fovea. Each flash was separated by a minimum of 700 ms and a maximum of 1200 ms. Across both retinal locations, saccadic inhibition occurred approximately 100 ms after flash onset. As expected, foveal transients elicited stronger inhibition than peripheral ones. We further examined whether initial transients were more effective at inducing inhibition than subsequent ones. While the magnitude and latency of inhibition remained consistent across flashes, the rebound phase —characterized by the reinstatement of saccades— showed signs of adaptation. Specifically, the proportion of saccades following the first transient decreased for subsequent flashes. This effect was more pronounced for peripheral stimuli, though a similar trend was observed for foveal ones. These findings suggest that during rapid successions of stimuli, the visual system sustains a robust inhibitory response to transients. However, the recovery of oculomotor plans adapts over time, reducing the likelihood of repeated motor execution, particularly for peripheral stimuli. This highlights the dynamic interplay between attentional prioritization and motor planning, where the visual system balances responsiveness to novel information with the efficient suppression of redundant signals in complex environments.

A.B. and C.C. were supported by a grant from the Italian Minister of Research and University (PRIN 2022_PNRR-P2022ST78T)

Eye Movements: Pursuit, learning, vergence

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.431 DOES A SIMILAR MECHANISM CONTROL VERTICAL AND HORIZONTAL BINOCULAR PURSUIT?

Stephen Heinen[1], Devashish Singh[1], Arvind Chandna[1], Scott Watamaniuk[1,2]; [1]Smith-Kettlewell Eye Research Institute, [2]Wright State University

There is evidence that horizontal and vertical smooth pursuit are subserved by different channels. Last year we showed that the eyes were well-aligned during binocular pursuit of a horizontally moving target. However, with one eye covered, the eyes still moved together but the covered eye was horizontally displaced. These results suggest that the eyes are driven by a unitary conjugate controller but have separate independent controllers that use visual feedback to maintain eye alignment. Here we ask if the eyes move equally well together when pursuing vertically-moving targets. Neurotypical observers pursued a small (0.4 deg) spot on a tangent screen that moved vertically with either a predictable sinusoidal profile, or an unpredictable combination of sinusoids (noise condition). 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 eye recording. We found that during monocular pursuit, while the eyes moved together vertically, the covered eye was again horizontally displaced, but to a greater degree than during horizontal pursuit. Furthermore, the horizontal displacement between the eyes was greater in noise condition. The results suggest that the eyes are driven together vertically using a similar conjugate mechanism as is used to drive them horizontally, but that independent control is less able to maintain alignment during vertical pursuit.

Supported by NIH R01EY034626 and Smith-Kettlewell Eye Research Institute

33.432 PLAN IT TO LEARN IT : MOTOR PLANNING DRIVES CONTEXTUAL ADAPTATION IN THE OCULOMOTOR SYSTEM

Maxime Martel[1] (maxime.martel@univ-lille.fr), Daniel M. Wolpert[2,3], Laurent Madelain[1,4]; [1]Université de de Lille, CNRS, CHU Lille, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, Lille, France, [2]Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10027, USA, [3]Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Cambridge, UK, [4]Aix Marseille Université, CNRS, Institut de Neurosciences de la Timone, UMR 7289, Marseille, France

Research on arm movements demonstrates that differences in motor planning - the process of organizing and sequencing movements - play a critical role in contextual motor learning. In contrast, motor execution - the physical performance of movements - does not drive the learning process. This study replicates these findings and extends them to the oculomotor system. Using a double-step paradigm, we implemented a sequence of two saccades to induce contextual adaptation. This was achieved by pairing the direction of the intra-saccadic step occurring during the first saccade with that of the second saccade. We assessed contextual saccadic adaptation across four experimental conditions (n=32): (1) planning and executing the second saccade, (2) viewing the second target cue without planning or executing the second saccade, (3) planning the second saccade without executing it (movement inhibited), and (4) executing the second saccade without prior planning (its direction was cued only after completion of the first saccade). The results revealed significant contextual saccadic adaptation in both the Planning & Execution condition and the Planning Only condition, with minimal to no adaptation observed in the other two conditions. These findings extend previous research by confirming that motor planning is a critical component of efficient contextual motor learning. We provide strong evidence that saccade planning plays a pivotal role in driving adaptation, whereas saccade execution or mere cue perception alone is insufficient.

33.433 HUMANS LEARN TO TRACE MEANINGFUL RELATIONSHIPS WITH THEIR EYES

Ataol Burak Özsu[1], Marcel Linka[1,2], Diana Kollenda[1,2], Benjamin de Haas[1,2]; [1]Experimental Psychology, Justus Liebig University, Giessen, Germany, [2]Center for Mind, Brain and Behavior (CMBB), Marburg and Giessen, Germany

Computational models of gaze in natural scenes evolved to consider ever more intricate features, from local contrast (Itti et al., 1998) to local meaning (Henderson & Hayes, 2017). However, even state-of-the-art models (Kümmerer et al., 2022) are blind to context and interrelations. In contrast, humans can extract global meaning from a single glance (Fei-Fei et al., 2007) and are sensitive to context effects (Murlidaran & Eckstein, 2024). Here, we ask whether human scan paths trace meaningful relationships between elements of a scene. We re-analyzed a large-scale dataset (n > 6,500; Linka et al., 2024) of children and adults free-viewing 40 complex scenes (Linka & de Haas, 2020). For each observer and scene, we simulated scan paths by randomly shuffling the order of empirical fixations, with or without spatial biases (Clarke et al., 2017). This simulation informed us about the frequency of transitions between scene elements to be expected under a (spatially biased) random walk. Adult observers were much more likely than this model to make direct saccades between certain pairs of scene elements. To understand the nature of these saccades, we asked a separate set of participants to describe the same set of scenes (Kollenda et al., 2024). Crucially, direct saccadic transitions were biased to trace relationships picked up in the descriptions. For instance, observers showed a strong tendency to make direct saccades from agents to patients of transitive actions featured in the descriptions. Finally, children's gaze was closer to the random walk model and systematic deviations slowly increased up until early adulthood. We conclude that humans learn to trace meaningful relationships with their gaze. This likely reflects and aids the understanding of global meaning (Murlidaran & Eckstein, 2024) and formation of event models (Loschky et al., 2020, 2024), which are informed by experience-dependent priors.

This research was supported by ERC Starting Grant INDIVISUAL to BdH and an Erasmus+ Traineeship to ABÖ

33.434 SPATIAL FREQUENCY, NOT ORIENTATION, CAN BE ASSOCIATED TO SPEED STATISTICS: EFFECTS ON PURSUIT EYE MOVEMENTS

Cléo Schoeffel[1], Anna Montagnini[1], Guillaume S Masson[1]; [1]Aix-marseille Université - Institut de Neurosciences de la Timone - CNRS - Marseille, France

Smooth pursuit eye movements reflect several biases identified in velocity estimation, such as the aperture problem for direction and the spatial frequency-dependent bias for speed. These biases are best explained by the properties of prior knowledge within the Bayesian inference framework. Some, but not all, previous studies have shown that global motion perception can be adapted through repeated exposure, suggesting that perceptual Priors may be updated based on target motion statistics. We investigated whether overrepresenting visual feature-speed associations could affect pursuit initiation. 11 participants tracked a Gabor (5°) moving horizontally at 10°/s, with varying orientations (0°,+/-45°) or spatial frequencies (0.3, 1cpd). Across five sessions, pre- and post-training baseline blocks (240 trials each) surrounded training blocks where one orientation or spatial frequency was paired with a higher speed (20°/s) and repeated across 720 trials. We compared pre- and post-training pursuit latency and initial eye velocity. Results showed that exposure to motion stimuli with a particular feature/speed association led to changes in pursuit initiation. There was no learning effect on latencies or initial eye velocity for the orientation/speed pairing. However, low spatial frequencies had later pursuit initiation compared to high spatial frequencies. After training, this delay disappeared as (i) latencies for low spatial frequencies were shortened over time and (ii) latency distributions now overlapped between the two spatial frequency conditions. Moreover, after latency realignment of pursuit responses, repeating high spatial frequencies at a higher speed increased initial pursuit eye velocity for high spatial frequency at slow speed. These results suggest that overrepresenting visual features during pursuit could bias pursuit initiation, consistent with an update of the sensory Priors used for 2D global motion estimation. The impact of target spatial frequency, but not its orientation, suggests that biases in speed and direction might be associated with different Priors and/or decoding of global motion.

33.435 DIFFERENT LATENCIES IN THE TWO EYES DURING MIDLINE PURSUIT

Scott Watamaniuk[1,2], Devashish Singh[2], Arvind Chandna[2], Stephen Heinen[2]; [1]Wright State University, [2]The Smith-Kettlewell Eye Research Institute

Chandna et al. (2021) reported that during pursuit of a target moving on the midline, a covered eye moved out of phase with the viewing eye. However, whether the eyes moved ahead or behind the stimulus was ambiguous since the stimulus moved with a continuous sinusoidal pattern. Here we explicitly evaluate the temporal initiation (latency) of midline pursuit in binocular and monocular conditions in discrete trials in which the target moved towards or away from the observer. Observers pursued a physical accommodative target (circular letters array) that moved in depth on a track aligned with the optical midline (peak vel. = 30cm/s). The stimulus started 50 cm from the observer and then moved towards (50-33 cm) or away (50-67 cm) from them at a random time and direction. After a second random fixation period (1-2 sec), the target then moved in the opposite direction to its initial position. Viewing was binocular, or monocular with one eye obscured with an infrared pass filter. An EyeLink 1000+ measured both eyes' movements. During monocular viewing, the occluded eye often had a longer latency than the viewing eye, but sometimes anticipated and moved before it, or in other trials remained stationary. Interestingly, even during binocular viewing pursuit initiation in the two eyes was asynchronous, with latencies often differing by 100 msec or more. The data suggest that a single motor command does not drive the eyes during midline pursuit, rather each eye appears to be controlled independently. These results are consistent with a new hybrid model of binocular control with independent eye commands components.

NIH 1R01EY034626-01

33.436 HOW DOES DEFOCUS LIMIT CONTINUOUS VERGENCE TRACKING OF OBJECTS IN OTHERWISE STATIC NATURAL SCENES?

Fuseina Safianu[1], Eswar Kurni[1], Bhagya L Marella[1], T Rowan Candy[1]; [1]School of Optometry, Indiana University, Bloomington, IN,USA

Objective: Eye alignment is driven by the distribution of retinal disparities across scenes. The ability to maintain alignment depends on scene structure and the observer's ability to encode disparity. This study asked how defocus and peripheral sensitivity combine across the visual field in limiting maintenance of eye alignment. Method: Dichoptic grayscale natural images (76x42deg SYNS database) were presented at a 70cm viewing distance using a PROPixx projector (RB3D mode, VPixx) and computationally-simulated bilateral or unilateral defocus (IRIS software, for participants' pupil size and 0, 1, 2, and 4D). A circular textured disk of 1/f noise with varying radii (1, 2, 4, 8, 12, and 16deg) was presented in the center of the image. The disk followed a 30s random walk in velocity between disparities of ±4deg. Two continuous tracking vergence trials to the moving disk were recorded from 5 adult functionally emmetropic subjects (age: 28 to 56yrs) using an Eyelink-1000 (SR-Research). Results: Peak correlations were computed by cross-correlating velocities of the stimulus and responses. For bilateral defocus, vergence peak correlations increased with disk radius ($p<0.001$). For anisometropic defocus, vergence correlations increased with disk radius and decreased with defocus (p<0.001). To further understand the anisometropia, peak correlations were computed for the defocused and focused eye gaze responses separately. Both the focused and defocused eye peak correlations changed significantly with both radius and defocus (p<0.001), although with a bigger effect of defocus in the defocused eye. Restricting the defocus to the central disk only did not change the qualitative results. Conclusion: The largest reduction in vergence performance occurred with smaller disks and differences in defocus between the eyes. These conditions are of concern for anisometropic children and patients using monovision, whose defocused eye would have both limited sensory input and ocular motor performance.

NEI RO1 EY014460

Eye Movements: Neural mechanisms

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.437 EXPLAINABLE AI-AIDED EXAMINATION OF SACCADE PREPARATION IN HUMAN EEG SIGNALS

Mojan Izadkhah[1] (mojanizadkhah@gmail.com), Cathy S Zhong[1], Jason J S Barton[1], Ipek Oruc[1]; [1]University of British Columbia

We investigated spatiotemporal patterns of neural processing underlying saccadic planning using explainable AI techniques. We recorded scalp EEG using a 64-channel BioSemi setup while participants (N = 20) completed pro-saccades to the left or right, randomly selected on each trial. A 3D adaptation of EEGNet trained with a novel data augmentation technique we developed to address the relatively small number of EEG trials (mean: 436.45 per participant, range: 273-492), successfully predicted saccade direction prior to onset with performance significantly above chance (mean AUC=0.77, p<.001, 95% CI: [0.67, 0.88]; mean accuracy=0.70, p<.001, 95% CI: [0.61, 0.79]) based on the held-out validation data. Model performance correlated strongly with dataset size for both AUC and accuracy (r=0.68, p=0.001), suggesting that sufficiency of training data was the primary limiting factor for model generalization performance. We applied a modified GradCAM algorithm to identify spatial and temporal features informing CNN predictions. Right frontal electrodes (FP2, AF4, F2, FC2, C2, CP2) were critical for predicting left saccades, while left electrodes (F1, FC1, C1, CP1) were important for right saccades. Of note, right frontal electrodes (FP2, AF4) remained critical for both directions, possibly reflecting a right-lateralized 'cognitive motor planning' signal. Temporal analysis of this signal revealed harmonic oscillations, peaking at 30Hz for both left and right saccades, consistent with low gamma band activity with a ~10ms lag for left saccades compared to right. A lateralized 'motor preparation signal' peaked during the final 32-16ms for right saccades at the left frontal electrodes. This motor preparation signal was more broadly distributed temporally for left saccades, appearing 36-24ms and 72-56ms prior to the onset of the saccade at the right frontal electrodes. We conclude that our explainable AI analysis can identify saccadic preparation signals and reveals significant asymmetries in the generation of left versus right saccades that require exploring.

This work was supported by an NSERC Discovery Grant RGPIN-2019-05554 (IO) and an NSERC Accelerator Supplement RGPAS-2019-00026 (IO). JB was supported by Canada Research Chair 950-232752. CZ was supported by the Neurological Foundation of New Zealand Grant 2246 CHF.

33.438 OCULOMOTOR CONTRIBUTION TO SYNCHRONIZATION OF CORTICAL ACTIVITY

Alessandro Benedetto[1,2,3] (alessandro.benedetto@unifi.it), Michele A. Cox[1,2], Jonathan D. Victor[4], Michele Rucci[1,2]; [1]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA, [2]Center for Visual Science, University of Rochester, Rochester, NY, USA, [3]Department of Neurosciences, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy, [4]Feil Family Brain and Mind Research Institute, Weill Cornell Medical College, New York, NY, USA

The human eyes are never at rest. Incessant movement also occurs in the intervals between saccades, when fixational drifts move the stimulus across many receptors on the retina. These drifts continually modulate the visual signals to the retina, and a growing body of evidence shows that the visual system uses spatial information provided by the resulting luminance modulations. In this study, we combined electroencephalography with high-resolution eye-tracking and gaze-contingent display control to directly estimate the influences of ocular drift on cortical responses in human observers. Subjects (N=15) maintained fixation on a high-frequency (10 cycles/deg) full-screen grating (~37x21 deg) modulated at ~7 Hz. The stimulus was presented at maximum contrast and either viewed normally, with the retinal motion caused by fixational drifts, or under retinal stabilization. In this latter condition, the stimulus moved on the display according to the observer’s eye movements, so to remain immobile on the retina. In both conditions, as expected, temporal modulation of the stimulus evoked a steady-state potential at twice the stimulus frequency, which primarily reflects responses from early visual areas. Since the image motion caused by eye movements redistributes the power of the stimulus on the retina, the spectral density distribution of the visual input was narrower under retinal stabilization, with more power concentrated at 7 Hz. Yet the magnitude of the EEG response was reduced by approximately 17% under retinal stabilization compared to natural viewing (t(14)=2.97, p=0.010). The result indicates that the continuous motion of the retinal image during fixation enhances early visual responses, even when the input is already temporally modulated. The dissociation between input power and response magnitude suggests the involvement of mechanisms driven by visuomotor contingencies that modulate early visual responses.

This work was supported by National Institutes of Health grants EY018363, EY07977, P30 EY001319, and University of Florence (Progetti competitivi 2025-2026)

33.439 INTRACRANIAL ELECTROENCEPHALOGRAPHY REVEALS THE NEURAL CORRELATES OF EYE-MOVEMENT GUIDANCE TO OBJECTS

Arish Alreja[1,2], Matthew Boring[1], Michael Ward[3], Mark Richardson[4], Taylor Abel[1], Avniel Ghuman[1]; [1]University of Pittsburgh, [2]Carnegie Mellon University, [3]University of California, Los Angeles, [4]Harvard University

When we move our eyes, it is to look at something: your daughter's face, the person who just walked into the room, a cup of water, your keys, etc. Yet our understanding about the neural guidance of eye movements is shaped by studying where we look in space and little is known about the role objects in our environment play in guiding eye movements. To address this gap, we recorded simultaneous intracranial electroencephalography with eye tracking in individuals with epilepsy during free viewing of scenes and during unscripted natural social interactions in the real world. In both paradigms, the category of object participants looked at next could be predicted based on their brain activity from prior to the onset of saccades to that object - even after controlling for the spatial location of the eye movement. Neural activity that enabled prediction of what a person was going to look at next arose from regions of parietal cortex traditionally associated with eye movement guidance, as well as category selective regions of ventral temporal cortex. The influence of para-foveal responses to the object of the next fixation was examined by comparing saccades traversing greater than 5 degrees of visual angle to smaller saccades, and prediction accuracy did not diminish for the larger saccades. Preliminary results show 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. 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.

National Institutes of Health (R01MH132225, R01MH107797, P50MH109429), the National Science Foundation (1734907), the Richard King Mellon Foundation, and the Beckwith Foundation.

33.440 OBJECT-BASED SACCADIC SIGNALS REVEALED IN SUPERIOR COLLICULUS LOCAL FIELD POTENTIALS

Christopher Conroy[1], Abe Leite[2], Hossein Adeli[3], Gregory J. Zelinsky[2], Robert M. McPeek[1]; [1]SUNY College of Optometry, [2]Stony Brook University, [3]Columbia University

Natural visual scenes are composed of spatially extended visual objects, and objects are often the targets of saccades. However, little is known about how such objects are represented by the saccadic system or selected as the targets of saccades. Recently, we discovered that when a saccade is made to one part of an extended object, superior colliculus (SC) spiking activity is enhanced throughout the object. Here, we examined local field potentials (LFPs) in the SC to better understand the population neuronal activity underlying this object-based saccadic enhancement. Rhesus monkeys fixated a central point and made saccades to peripheral target locations that were either part of an extended object (connected condition) or separated from the object by a small gap (disconnected condition). LFPs were recorded from SC sites with response fields (RFs) that overlapped the extended object. Crucially, however, the RFs of those sites were spatially remote from the saccade target locations. We found that peri-saccadic LFPs differed in the connected and disconnected conditions, despite the fact that the saccades made in the two conditions were identical. Because LFPs, unlike spiking activity, are thought to reflect primarily synaptic activity associated with the input to neurons, this suggests that the object-based saccadic enhancement of SC spiking activity that we previously demonstrated is mediated by upstream neural structures and/or by intra-collicular interactions.

NIH-NEI R01-EY030669

33.441 TRACING PERISACCADIC CHANGES IN LOCATION PERCEPTION IN EXTRASTRIATE AND PREFRONTAL RESPONSES

Geyu Weng[1,2], Behrad Noudoost[2], Neda Nategh[2,3]; [1]Department of Biomedical Engineering, University of Utah, [2]Department of Ophthalmology and Visual Sciences, University of Utah, [3]Department of Electrical and Computer Engineering, University of Utah

The brain generates a perception of visual space through processing and interpreting the location of stimuli projected onto the retina, yet how spatial information is represented and read out in neural responses remains unclear. To understand the brain's spatial code, it is essential to investigate alterations in neural activity that accompany changes in visual perception. In this study, we measure location perception in rhesus macaque monkeys using combined behavioral and physiological experiments, linking perisaccadic modulations in extrastriate responses to their behavioral counterparts. We use array and single electrodes to simultaneously record spiking activity and local field potentials (LFPs) in area V4 and the Frontal Eye Field (FEF) from sites with overlapping receptive fields while the monkey's eye movements are monitored. In each trial, the monkey makes a saccade from a fixation point to a peripheral saccade target. During fixation and saccade execution, a 50-ms visual probe stimulus is presented in one of nine possible locations in a 3×3 grid placed around the V4 neuron's receptive field. When the probe disappears, the monkey makes another saccade to the perceived stimulus location. We measure perisaccadic mislocalization, defined as the shift in the perceived location of a stimulus when presented around the time of the first saccade, compared to when it is presented during fixation. We correlated mislocalization magnitudes with spike rates in populations of FEF and V4 neurons and showed that trials associated with the smaller magnitudes of mislocalization corresponded to higher firing rates in these areas. We then analyzed FEF and V4 signals, recorded simultaneously during the same session, to extract the oscillatory components communicated between the two areas around the time of saccades. We show that behavioral mislocalization can be traced in FEF and V4 signals, demonstrating the possible role of prefrontal and extrastriate areas in location perception.

33.442 NEURAL ENCODING OF FIXATION LOCATION AND ORDER OF GAZE PATHS

Radha Nila Meghanathan[1], Stefan Pollmann[1,2]; [1]University of Magdeburg, [2]Center for Behavioral Brain Sciences, Magdeburg

Gaze paths, like the images that prompt them, are represented in the cortex. Previous research from our lab has shown that gaze paths associated with faces and houses are distinctly represented in the fusiform face area, the parahippocampal place area, the superior parietal lobule and frontal eye fields. Although face and house gaze paths differ spatially, they encode both location and order information. Do neural representations of gazepaths contain both spatial and temporal information? To investigate whether and where location and order information in gaze paths are represented in the cortex, first, we recorded eye movement patterns while participants viewed images of faces and houses. Second, we recorded fMRI data while a different sample of participants followed these gaze paths by tracking dot sequences presented on a gray background in both ordered and shuffled sequence. Using multivariate pattern classification analysis with a linear support vector machine estimator on a brain parcellated according to the Schaefer atlas, we could distinguish activation patterns elicited by face and house-related gaze sequences in ventral occipitotemporal regions, specifically, the fusiform, lingual and occipital gyri, for both ordered and shuffled gaze paths. In contrast, along the borders of the horizontal and descending segments of the left intraparietal sulcus, we were able to classify face from house activation patterns only for ordered gaze paths. Notably, eye movements recorded while participants tracked dot sequences showed better dot tracking for ordered than shuffled sequences. Our findings show that the order of gaze paths is represented in the dorsal stream, while location information is encoded in the ventral stream. Ventral occipitotemporal cortex, which is involved in shape perception, may process predominantly spatial fixation locations but not temporal order. In contrast, dorsal regions, involved in visuomotor control, may use order information to predict upcoming saccade goals.

33.443 UNVEILING THE FUNCTIONAL ARCHITECTURE OF HUMAN EYE FIELDS

Marco Bedini[1] (marco.bedini@univ-amu.fr), Uriel Lascombes[1], Sina M Kling[1], Guillaume S Masson[1], Martin Szinte[1]; [1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Université, Marseille, France

While extensive research has elucidated the neural networks underlying eye movements in non-human primates, the functional architecture of these networks in humans remains relatively unclear. Here, we analyzed high-field resting-state fMRI data using reproducible pipelines in the context of a larger dataset targeting oculomotor functions and visual field mapping in 20 healthy participants. We computed correlations of BOLD time series between all surface vertices belonging to retinotopic regions along the cortical visual hierarchy and macro-anatomical clusters in the parietal and prefrontal cortex localized along the medial, superior, and inferior branches of the precentral sulcus. These clusters were defined by merging regions of the Glasser multimodal parcellation and spanned the supplementary and cingulate eye fields, the frontal eye field, and the premotor eye field, respectively. We refined the selection of our vertices in oculomotor areas in parietal and frontal clusters to those strongly activated during visual, smooth pursuit, and saccadic eye movement tasks. Using a seed-based functional connectivity approach, we contrasted pairwise correlations between all clusters and systematically evaluated their contributions to the overall connectivity patterns. Our analyses revealed consistent functional connectivity patterns linking the oculomotor parietal clusters to the three major prefrontal clusters. Combined, our results provide a comprehensive framework for understanding the functional architecture of the human eye fields.

This research was supported by a postdoctoral fellowship awarded by the FRM to M.B. (SPF202409019664), and an ANR JCJC and a Fyssen Foundation grant to M.S..

33.444 DYSCONNECTIVITY OF SOMATOMOTOR AND VISUAL NETWORKS EMERGES CLEARLY IN THE FUNCTIONAL BUT NOT STRUCTURAL CONNECTOME IN EARLY PSYCHOSIS

Brian Keane[1], Yonatan Abrham[1], Kirsten Peterson[2], Michael Cole[2], Paul Geha[1], Skylar DeWitt[1]; [1]University of Rochester, [2]Rutgers University, Newark

Psychosis patients functionally exhibit thalamo-cortical hyperconnectivity and cortico-cortical hypoconnectivity in somatomotor and secondary visual (visual2) networks. These dysconnectivity patterns have been combined to form a “somato-visual” biomarker that is highly robust, reliable, and generalizable across samples (Keane et al., 2024, Mol. Psych.). Can traces of this biomarker be found in white matter? To address this question, we leveraged diffusion MRI data from the Human Connectome Early Psychosis project, which included 33 healthy controls (22 males; mean age=25) and 86 early-stage psychosis patients (50 males; mean age=24). Diffusion data were preprocessed using QSIPrep. Tractography was performed using DSI Studio. Structural connectomes (361x361) were derived from the thalamus plus 360 cortical parcels. Thalamo-cortical connectivity was computed by averaging streamline counts between the thalamus and visual2/somatomotor networks. Cortico-cortical connectivity was computed by averaging streamline counts between cortical parcels of the same two networks. To compute the structural variant of the biomarker, we normalized and subtracted the two averaged values (thalamo-cortical – cortico-cortical). To provide a comparison, we re-derived the functional variant of the biomarker from the same subjects in the same atlas space using already-published connectomes (ibid). Patients exhibited expectedly higher functional biomarker values (p=5e-07, Hedges’ g=1.0) but non-significantly lower structural biomarker values (p=.10, g=-.3; BF01=14.1, one-tailed). The functional and structural biomarker variants were uncorrelated across subjects (r=-.05, p=.66). The foregoing null results could not be ascribed to noisy data since the connectomes passed numerous quality checks, e.g., showing larger structural/functional correlations within- than between-subjects (p=9e-07) and larger thalamo-cortical structural connectivity with primary versus secondary visual networks (p=6e-06). To conclude, somato-visual functional dysconnectivity in early psychosis may emerge synaptically (e.g. due to pathology to the NMDA receptor) without corresponding large-scale changes in white matter. A next goal is to understand the behavioral consequences of this dysconnectivity (e.g., oculomotor control, visually-guided reaching, perceptual organization).

This work was supported by a Del Monte Institute pilot grant to BPK

33.445 DECODING FOVEAL FEEDBACK OF SACCADE TARGETS FROM EARLY VISUAL CORTEX

Luca Kämmer[1,2,3] (lucakaemmer@gmx.de), Lisa M. Kroell[3], Tomas Knapen[4,5,6], Martin Rolfs[3,7,8,9], Martin N. Hebart[1,2,10]; [1]Vision and Computational Cognition Group, Max Planck Institute of Human Cognitive and Brain Sciences, [2]Department of Medicine, Justus Liebig University Giessen, Giessen, Germany, [3]Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany, [4]Spinoza Center for Neuroimaging, KNAW Netherlands, [5]Computational Cognitive Neuroscience and Neuroimaging, *Netherlands Institute for Neuroscience, [6]Experimental and Applied Psychology, Vrije University Amsterdam, Netherlands, [7]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Germany, [8]Exzellenzcluster Science of Intelligence, Technische Universität Berlin, Germany, [9]Bernstein Center for Computational Neuroscience Berlin, Germany, [10]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt, Germany*

Humans sample their visual environment with frequent, rapid eye movements to make best use of the high resolution of their fovea. Despite the constant shifts in visual input, perception remains remarkably stable. Foveal cortex has been proposed as a key component in maintaining this stability by receiving feedback from cortical areas that encode peripheral information during saccade preparation. Here we aimed to uncover (1) whether foveal feedback can be found in early visual areas, (2) whether the information encoded in this activation is stimulus-content specific, and (3) which neural regions might be involved in driving this feedback. To dissociate neural processes elicited by direct visual input from those related to foveal feedback, we designed a gaze-contingent fMRI study where the saccade target is removed before it reaches the fovea. As targets, we used natural images and independently manipulated object shape and category. Even though targets were never presented in the fovea, the results showed reliable decoding from foveal parts of early visual cortex, suggesting feedback of feature-specific information during saccade preparation. Breaking the analysis down by stimulus type revealed a drop in decoding across category but not across shape, demonstrating that the classifier was sensitive to shape, but not semantic stimulus information. Cross-decoding to a control condition with foveal stimulus presentation was above chance, indicating a similar neural representation of foveal feedback signals to direct stimulus presentation. An eccentricity-dependent analysis revealed a dip in decoding between foveal and peripheral processing, indicating that the results are unlikely to be explained by spillover from peripheral cortex. Exploring the regions involved in modulating foveal feedback, we found a specific involvement of the intraparietal sulcus. By characterizing the neural representations underlying the processing of peripheral information in foveal retinotopic areas, these findings underscore the importance of foveal processing in active visual perception.

This work was supported by a doctoral scholarship of the German Academic Scholarship Foundation ("Studienstiftung"), a Max Planck Research Group grant, the ERC Starting Grant COREDIM and the Hessian Ministry of Higher Education, Science, Research and Art.

Binocular Vision: Rivalry and bistability, stereopsis, models, neural mechanisms

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.446 MECHANISMS UNDERLYING THE CONTRIBUTION OF SPATIAL FREQUENCY OF DISPARITY AND VERGENCE TO DEPTH PERCEPTION IN STEREOPSIS

Joonsik Moon[1] (moon.joo@northeastern.edu), Peter Bex[1]; [1]Northeastern University

The distribution of visual cues across depths and retinal eccentricities changes with gaze direction and distance. However, how the stereoscopic system utilizes these covarying distributions to maintain a stable perception of 3D structure remains underexplored. This study uses classification image analysis with naturalistic images to examine the selectivity of stereoscopic disparity mechanisms across changes in vergence. Three observers distinguished between convex and concave targets over 1,000 to 2,000 trials. Targets were disparity-defined Gaussians (σ = 0.4°) centered ±2° left or right of fixation (compliant within 1°), 100% contrast 1/F noise, on a polarized 3D display (LG 42LM6200, Lmean=75cd/m^2), with gaze compliance enforced with a 2000Hz Eyelink. In alternating blocks of twenty 200msec trials, participants verged at near (40cm) or far (infinity) distances and reported the location of the concave or convex target (at random across subjects). In a training phase, stimulus presentation duration decreased in 5 steps from 10 s to 200 ms. Stereoscopic disparity was controlled by a 3/1 staircase, thresholds from the training phase seeded the test phase. The 2D Fourier transform amplitudes of the interocular difference images were classified based on responses. Classification images contained peaks at high spatial frequencies for vertical and oblique orientations. For near-convergence, oblique orientations were correlated with correct responses, whereas for far-divergence, vertical orientation were correlated with correct responses. This suggests that stereoscopic performance shifts from vertical to oblique content as gaze shifts from far-divergence to near-convergence. These findings provide new insights into the influence of vergence signals in orientation tuning during stereoscopic depth encoding that could be related to ocular torsion during convergence.

This research was supported by NIH grant EY029713.

33.447 TRACKING VISUAL AWARENESS IN CFS

Jan Drewes[1], Mingjie Gao[2], Weina Zhu[2]; [1]Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, China, [2]School of Information Science, Yunnan University, Kunming, China

What happens in our brain when a perceived target enters our visual awareness, and what factors regulate this process? For years, b-CFS has been an effective tool for studying conscious awareness by using high-contrast dynamic masks in one eye to suppress awareness of a static target stimulus with gradually increasing contrast in the other eye for relative long time. Participants' response times serve as a behavioral indicator of conscious awareness. However, the brain dynamics underlying this conscious process remain a “black box.” In this study, we address this gap by combining EEG/SSVEP frequency tagging and sensory entrainment in a b-CFS paradigm to track and modulate the process of conscious perception. Sensory entrainment was achieved by presenting beep sounds at the same frequency as the target or mask stimuli. Through time-frequency analysis, we found that the power and time course of the target and mask frequencies was modulated by cross-modal sensory entrainment. Sensory entrainment enhanced the target frequency energy in the target-entrainment condition and the mask frequency energy in the mask-entrainment condition, relative to the no-entrainment condition. Regarding the time course, the mask frequency initially dominated but decreased over time, while the target frequency energy began to increase after a certain point, with a crossover occurring eventually. Notably, such crossover time occurred earlier in the target-entrainment condition compared to both no-entrainment and mask-entrainment conditions. These findings suggest that visual awareness can be tracked using CFS-Tagging, and its progression can be modulated through frequency energy.

WZ was supported by the National Natural Science Foundation of China (61263042, 61563056)

33.448 **A NEW 'TRACKING' VERSION OF CONTINUOUS** FLASH SUPPRESSION TO QUANTIFY SUPPRESSION STRENGTH: CONSTANT CFS SUPPRESSION FOR ALL IMAGE TYPES & TWO TIMES THE SUPPRESSION STRENGTH OF BINOCULAR RIVALRY

David Alais[1], Jacob Coorey[1], Randolph Blake[2], Lina Ye[1], Will Swann[1], Annie Wang[1], Matthew Davidson[3]; [1]University of Sydney, [2]Vanderbilt University, [3]University of Technology, Sydney

A dynamic stimulus presented to one eye can suppress a static target in the other for long periods (continuous flash suppression: CFS). The suppressed target eventually becomes visible and this duration (bCFS) is often used to index unconscious processing. Controversially, faster breakthroughs are considered evidence of visual processing without awareness while opponents claim breakthrough times simply vary with low-level stimulus properties. bCFS times alone cannot solve this: suppression thresholds are needed as a baseline to compare with breakthrough thresholds. Our new 'CFS tracking' paradigm (tCFS) quickly measures contrast thresholds for breakthrough and suppression so that suppression strength can be calculated for any image. Participants simply track their changing perceptual states as a suppressed image steadily increases in contrast until visibility (i.e., breakthrough) and then decreases until re-suppression is reported, then increases again (and so on, in a continuing cycle). Using tCFS we confirm that: (i) there are some differences in breakthrough thresholds across target types (e.g., grating vs face), as bCFS has shown, but (ii) suppression thresholds show a parallel pattern of differences, thus (iii) suppression strength is the same for all images (~14-15 dB for gratings, noise, objects, food, faces, biological motion). Uniform CFS suppression strength indicates a single mechanism of CFS suppression, likely early in visual cortex where left and right eyes combine, and prior to processing of objects and image identity. Using the tracking paradigm with binocular rivalry reveals half as much suppression: ~7-8 dB.

33.449 FASTER BREAKTHROUGHS OF FEARFUL FACES FROM CONTINUOUS FLASH SUPPRESSION ARE ACCOMPANIED BY FASTER EYE RESPONSES

Junchao Hu[1], Stephanie Badde[2], Petra Vetter[1]; [1]University of Fribourg, Switzerland, [2]Tufts University

In breaking continuous flash suppression (bCFS), a flickering high-contrast mask displayed to one eye suppresses conscious awareness of a visual stimulus shown to the other eye, and one measures the time it takes participants to report the suppressed stimulus becoming visible. Fearful faces have been shown to break through continuous flash suppression faster than neutral faces, which is often interpreted as prioritized processing of fear during unawareness. However, it is unclear whether faster reactions to suppressed fearful faces genuinely reflect prioritized unaware processing, or whether speed differences in reports of breakthrough arise from post-perceptual processes, such as variations in participants' manual response speeds after stimuli reach awareness. Here, we combined bCFS with eye-tracking. By continuously tracking observers' eye movements as they detected suppressed, invisible fearful and neutral faces, we examined the temporal dynamics of eye responses to these faces during suppression and breakthrough. Participants pressed a button to indicate the location of the suppressed face image as soon as they saw parts of it. Additionally, they reported the face's emotional expression and its visibility after they gave the localization response. Behavioral results showed that participants localized initially suppressed fearful faces faster than neutral faces. Eye-tracking revealed that even before reported awareness of faces, the eyes moved earlier towards suppressed fearful faces than neutral ones. When the faces were superimposed on the flashing mask, thus always visible, manual reaction times and oculomotor responses to fully visible fearful and neutral faces were indistinguishable. Our findings suggest that faster breakthrough RTs of fearful faces are driven by faster eye responses to fearful faces during unaware processing. Our approach avoids potential confounds of post-perceptual and decisional factors associated with RT measures in bCFS. We propose that fearful faces' ability to attract the eyes, even in the absence of awareness, may facilitate their perceptual detection.

This work was supported by a PRIMA grant (PR00P1_185918/1) from the Swiss National Science Foundation to Petra Vetter.

33.450 MASK CONTRAST, BUT NOT MASK SIZE OR PARTICIPANT EXPECTATIONS, MODULATES SUPPRESSION DEPTH IN CONTINUOUS FLASH SUPPRESSION.

Jacob Coorey[1], Matthew Davidson[2], David Alais[1]; [1]The University of Sydney, [2]University of Technology Sydney

In continuous flash suppression (CFS), a high-contrast dynamic mask in one eye suppresses a small, static, low-contrast target in the other. While mask size and contrast have been shown to affect suppression duration, their impact on suppression depth—a measure of the strength of suppression—remains unexplored. Additionally, the influence of high-order predictive factors, such as expectations of target location and identity, on suppression depth is unknown. We investigate these questions using a new tracking-CFS (tCFS) method. We measured breakthrough thresholds (a weak target increasing in contrast until visible), suppression thresholds (a strong target decreasing in contrast until suppressed), and suppression depth (the ratio of suppression-to-dominance). In our first study, we manipulated participant expectations by having stimuli unexpectedly change location or identity while suppressed. In our second study, we manipulated the size and contrast of the masking stimulus. Our findings reveal three key points: (i) suppression depth is unaffected by target predictability; (ii) suppression depth does not change with mask size; (iii) suppression depth increases with higher mask contrast. The lack of significant effects from mask size suggests an early interocular suppression process, likely occurring in V1, where interocular conflict is first detected in smaller receptive fields. Similarly, the absence of effects from target expectation aligns with our previous findings of consistent suppression depth irrespective of target features. These results suggest that suppression occurs before global target properties are processed. Finally, the change in suppression depth with increased mask contrast indicates that the ask masking strength increases, so to does the required contrast change from the target image. Future research will explore whether the surround of a small mask impacts suppression depth in CFS, as it does in binocular rivalry, to further examine and distinguish between these phenomena.

33.451 IMAGERY PRIMING OF BINOCULAR RIVALRY IS NOT A RELIABLE METRIC OF INDIVIDUAL DIFFERENCES IN THE SUBJECTIVE VIVIDNESS OF VISUALISATIONS.

Loren N. Bouyer[1], Dietrich S. Schwarzkopf[2], Blake W. Saurels[1], Derek H. Arnold[1]; [1]The University of Queensland, [2]School of Optometry & Vision Science, The University of Auckland, New Zealand

There are marked individual differences in peoples' capacity to construct imagined visual experiences within their minds. This capacity ranges from aphants who report an inability to construct mental images at all to hyper-phantasics who report having imagined experiences that are as realistic as seeing. This intensity of peoples' imagined experiences is most often measured via subjective self-report. However, Chang and Pearson (2018) have suggested that probe detection in a binocular rivalry (BR) protocol can be used as an objective measure. They found that pre-imagining a moving input could prime participants to suppress the non-imagined input. This suppression was objectively measured through a detection task of probes embedded in thesuppressed and non-suppressed inputs. They found that pre-imagining enhanced performance on an objective probe detection task when probes were embedded in imagery-consistent inputs, as opposed to imagery-inconsistent inputs. Our aim is to assess if this type of objective imagery priming can be used to predict the vividness of visualisations between different people. In our study using static BR inputs, we find that imagery primes the detection of probes within imagery-consistent inputs, as opposed to imagery-inconsistent inputs. However, the priming of objective sensitivity to probes does not correlate with the typical ratings people use to describe the vividness of their visualisations (a between participants effect). In contrast, objective priming of sensitivity to probes embedded in BR inputs was greater on trials when participants reported that their pre-imagined experience had been relatively vivid (a within participants effect). Overall, our data suggest that while imagery can prime objective sensitivity to probes during BR, there is currently no strong evidence that this effect can be used as an objective method of predicting the subjective vividness of different people's visualisations.

33.452 FUSIBLE IPSILATERAL SURROUND REDUCES MAGNITUDE OF INTEROCULAR SUPPRESSION WHEN SURROUND CONDITIONS RANDOMLY CHANGE

David Nichols[1], Kiana Watt[1], Margaret McCray[1]; [1]Roanoke College

Interocular suppression can render invisible an image shown to only one eye with the sudden onset or sustained presentation of a stimulus in the other eye. While the effect is strongest for overlapping images, interocular suppression extends to nearby areas. Prior research with distinct adaptation and testing phases found the onset of a contralateral surround resulted in a similar amount of suppression on a nonoverlapping monocular grating regardless of the presence or absence of an ipsilateral surround. With prolonged viewing, fusible surrounds with matching images in the two eyes suppressed less than rivalrous surrounds with nonmatching images, theoretically due to alternations in the strength of the contralateral surround. The current study used continuous perceptual monitoring of a monocular target grating for sixty seconds while the presence or absence of a surround in the ipsilateral and contralateral eye varied independently and randomly every two seconds. Data has been collected from over twenty participants across two experimental setups (mirror stereoscope and shutter glasses), with different presentation parameters (including varying the size of the monocular grating and contrast of the surrounds) to establish the consistency of the observed effects. A monocular target remained visible when there was only an ipsilateral surround, but would occasionally become invisible when there was a contralateral surround present, establishing interocular suppression. Novel findings were that suppression from visibility was substantially more likely when there was only a contralateral surround than when both the contralateral and ipsilateral surrounds were present in a fusible way, even at their time of onset. Furthermore, removing the ipsilateral surround when the contralateral surround was already present resulted in an increased amount of suppression. Therefore, while a contralateral surround is necessary for interocular suppression, its presence alone is not sufficient and its suppressive strength is impacted by the presence of a fusible ipsilateral surround.

33.453 QUANTIFYING LUNING: PERCEPTUAL FADING AT MONO-BINOCULAR BOUNDARIES IN AUGMENTED REALITY

Zhetuo Zhao[1], Gerrit Maus[1]; [1]Magic Leap, Inc.

In near-eye display systems for augmented reality (AR), partial binocular overlap between the left and right fields of view (FOV) can arise due to factors such as the depth of virtual content, users interpupillary distance (IPD), and efforts to expand perceived FOV coverage. This partial overlap divides the FOV into three regions: a central binocular zone and two peripheral monocular zones. At the binocular-monocular (BM) borders, users may perceive a fading of virtual content, commonly known as Luning artifacts. Most prior research has measured these artifacts by the proportion of time users detected fading, which provides limited insight into the severity and spatial profile of the effect. Here, we introduce a luminance comparison paradigm to quantitatively measure perceptual fading as a function of horizontal position relative to the BM borders. Measurements conducted on AR wearables demonstrate that this method effectively quantifies the Luning artifacts and enables objective assessment of potential mitigation strategies, such as vignetting of the FOV near the nasal boundary in each eye. Additionally, our findings reveal that the spatial profile of the perceptual fading differs between the left and right BM borders, which suggests close links to users ocular dominance.

33.454 INCREASES VS. DECREASES: ASYMMETRIC EFFECTS OF CONTRAST CHANGES DURING BINOCULAR RIVALRY MODULATED BY AWARENESS OF PERCEPTUAL SWITCH

Changzhi Huang[1,2] (2023023770@m.scnu.edu.cn), Rong Jiang[1,2], 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, Guangzhou, 510631, China

The human visual system efficiently processes moving or changing stimuli, as these stimuli are known to attract attention and facilitate breaking suppression to reach perceptual awareness. We investigated how contrast changes—increases and decreases in stimulus intensity—affect breaking interocular suppression during binocular rivalry. In Experiment 1a, participants viewed dichoptic sinusoidal gratings of differing orientations. When participants reported full perception of one grating, the contrast of the suppressed grating increased or decreased over 100 ms. Reaction times for the suppressed stimulus to gain perceptual dominance were recorded. For comparison, static baseline conditions with constant contrasts matching the final intensities of the changing contrast conditions were included to isolate effects of intensity changes from constant intensity differences. Results showed that contrast increases led to significantly faster breakthroughs into perceptual dominance compared to decreases. Increases also resulted in faster breakthrough than the constant higher-intensity condition, indicating a specific facilitative effect of increasing intensity. Conversely, decreases delayed breakthrough than the constant lower-intensity condition, demonstrating an asymmetry in subliminal processing favoring increases. To further validate these findings, Experiment 1b involved faster intensity changes over only 10 ms, and Experiment 1c recorded the time for partial breakthroughs of the changing stimulus. Both results were consistent with Experiment 1a. Experiment 2 presented random dots moving in different random directions dichoptically, making interocular conflict imperceptible and non-reportable. Under these conditions, any change in intensity—regardless of increase or decrease—led to easier breakthrough of interocular suppression. Together, our findings reveal an asymmetric effect of contrast changes during binocular rivalry. When observers are aware of rivalry switches, increases in intensity facilitate perceptual breakthrough while decreases hinder it. When switches are not perceived, any contrast change enhances perceptual breakthrough. These results highlight the interplay between subliminal sensory processing of contrast changes and conscious awareness, shedding lights on developing theoretical models of binocular rivalry.

This work is supported by (1) the National Nature Science Foundation of China (Grant No. 31871136), (2) the Science and Technology Ministry of China (Grant No. 2021ZD0204200), and (3) the Sino-German Center for Research Promotion (Grant No. M-0705).

33.455 MODELING THE INFLUENCE OF OSCILLATION FREQUENCY ON PERCEPTUAL STABILITY DURING BINOCULAR RIVALRY

Oakyoon Cha[1] (oakyoon@gmail.com); [1]Sungshin Women's University

Embedding oscillations in the decision criterion of the one-choice drift-diffusion model led to the emergence of oscillatory patterns in the response time distribution (Cha, 2024). Notably, even minimal noise in the oscillation amplitude greatly impacted these patterns, with the influence growing stronger at higher oscillation frequencies. An intuitive prediction from these results is that lower-frequency oscillations could provide better stability when used for sensory gating mechanisms. Consistent with this prediction, a study using binocular rivalry found a negative correlation between the peak alpha (7–13 Hz) frequency and stable percept durations (Katyal et al., 2019). This relationship between oscillation frequency and perceptual stability, although intuitive, has never been predicted by any model for multistable perception. To simulate binocular rivalry durations, the present study utilized the hierarchical Brownian model (HBM; Albert et al., 2017), which models perceptual evidence as continuously drifting between two boundaries representing competing percepts. I made three key modifications to this model: First, I introduced two intermediate boundaries to incorporate mixed perceptual dominance states, allowing the simulation of transitions where both percepts are partially dominant. Second, I embedded oscillations into the perceptual evidence and intermediate boundaries to reflect the rhythmic fluctuations in neural activity influencing perceptual dominance. Third, I reset the phases of these oscillations at the onset of each exclusive perceptual dominance period, mirroring the neural reinitialization that occurs with perceptual switches. This modified model confirmed the intuitive prediction that lower alpha peak frequencies lead to longer stable percept durations. Additionally, the model predicted that oscillations during mixed dominance periods would be stronger than during exclusive dominance periods, consistent with previous findings (Cha & Blake, 2019). These results suggest that lower oscillation frequencies afford advantages in achieving stable perception, enhancing our understanding of the neural mechanisms underlying perceptual stability.

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).

33.456 NEURAL PROCESSING OF BINOCULAR RIVALRY ACROSS THE VISUAL HIERARCHY

Junxian Rao[1] (rao.jun.xian@gmail.com), Chuan Hou[1]; [1]Smith-Kettlewell Eye Research Institute

When incompatible images are presented to each eye at the exact spatial location, perception alternates between them as they compete for dominance, a phenomenon known as binocular rivalry. While previous studies suggest that the neural representations of the rival stimuli compete for visual awareness beyond the primary visual cortex (V1), the precise contributions of different brain regions in the visual hierarchy remain unclear. Here, we investigated the neural mechanisms underlying binocular rivalry by presenting a pair of orthogonal gratings tagged with different temporal frequencies (F1=7.5Hz, F2=6Hz) in each eye to normal vision observers. The participants indicated their perceptual experience (i.e., horizontal or vertical) by pressing one of the two buttons. Neural activity during binocular rivalry and monocular stimuli was recorded using fMRI source-localized 128-channel EEG to provide high spatial definition to the neurophysiological signals. We correlated neural responses from different brain regions with the perceptual reports to examine activity during the dominance and suppression phases. Our findings showed that extrastriate visual areas, including hV4, the middle temporal (MT) area, lateral occipital cortex (LOC), and higher-level cortical regions (e.g., frontal pole) exhibit more robust neural responses during binocular rivalry compared to V1. Notably, the frontal pole showed the highest neural response, indicating that cognitive cortical regions, such as frontal cortices, are involved in rivalry processing. Moreover, compared with the monocular stimuli, the magnitude of neural suppression increased from V1 to extrastriate visual areas (MT and LOC) and the frontal pole. Suppression was again highest in the frontal pole, suggesting that the cognitive cortices might mediate visual awareness during binocular rivalry. Our results support a three-tier model of binocular rivalry processing, with V1 at the first level, extrastriate areas at the second level, and the frontal region at the third level.

This research was supported by NIH grant R01EY035346 to C. H.

33.457 NEURAL RESPONSES TO BINOCULARLY IN-PHASE AND ANTI-PHASE STIMULI

Bruno Richard[1] (bruno.richard@rutgers.edu), Daniel Baker[2]; [1]Rutgers University - Newark, [2]University of York

Binocular vision fuses similar stimuli into a single percept, yet incompatible stimuli result in other experiences such as rivalry, lustre and diplopia. We measured neural responses to binocular stimuli with different phase relationships, intending to understand them using contemporary binocular models. Steady-State Visually Evoked Potentials (SSVEPs) were recorded from 15 observers in response to monocular and binocular stimulation at 3Hz, using either on-off or counterphase flicker. Across the eyes, binocular stimuli could be (i) in spatial and temporal phase, (ii) in temporal phase but spatial antiphase, (iii) in spatial phase but temporal antiphase, or (iv) in spatial and temporal antiphase (for counterphase flicker this is identical to condition(i)). Responses to monocular on-off flicker showed peaks at the fundamental frequency (3Hz) and its harmonics (integer multiples of 3Hz). In contrast, counterphase flicker produced responses only at twice the flicker frequency (6Hz) and its harmonics. Binocular in-phase stimulation resulted in a similar pattern of responses, consistent with 'ocularity invariance' – the observation that binocular and monocular stimuli appear equal at high contrasts. Changing the phase relationship modulated the harmonics pattern in complex ways: in particular, on-off flicker in temporal antiphase reduced the fundamental response, but there was no such effect for counterphase flicker. We modelled the data using a progression of binocular combination algorithms that increased in complexity from a simple linear sum to a two-stage binocular gain control model with parallel monocular and binocular phase-selective channels (the Lustre model; Georgeson et al., 2016). The most complex model (lustre) outperformed all other models in capturing the variance of our SSVEP data, although simpler phase-insensitive models performed similarly well in most experimental conditions. Simpler models struggled to capture the response magnitude to counterphase stimuli. Our findings suggest explaining neural responses to binocular stimuli with different phase relationships requires phase-selective channels.

Binocular Vision: Clinical, perception

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.458 A COMPARISON OF METHODS FOR MEASURING INTEROCULAR DELAYS

Brooke Lim[1], Anna Kosovicheva[1]; [1]University of Toronto Mississauga

Many everyday tasks rely on binocular vision, which is impaired in individuals with amblyopia. Impairments in visual-spatial processing normally characterize amblyopia, but previous work has shown deficits in temporal processing as well, including processing delays in the amblyopic eye. Many techniques have been developed to measure interocular timing delays behaviourally by showing different images to the two eyes and recording participant responses. However, agreement between these measures has not been previously investigated. We compared four different assessment measures in normally-sighted observers: depth-based judgments (using the Pulfrich effect), interocular flicker integration, reaction time to monocular targets, and interocular temporal order judgments. Stimuli were presented using a high-speed projector with passive polarized filters (240 Hz per eye), enabling precise temporal control for dichoptic presentation. We also included a measure of sensory eye dominance to determine how eye dominance is related to each of the timing-based measurements. Pairwise comparisons of temporal delays measured across methods showed that the best-correlated pair of measures was between interocular flicker integration and temporal order judgements (r = 0.50). For each measure, we additionally calculated the average correlation between it and the remaining three measures. The Pulfrich effect was the best-correlated measure for examining timing delays between the eyes (Fisher Z = 0.24). In contrast, the measure that was least correlated with the other three measures was reaction time (Fisher Z = 0.09). Eye dominance was not correlated with the four temporal delay measures (Fisher Z = -0.01). Together, these results suggest that methods that rely on binocular integration are more reliable than monocular measurements. This highlights the importance of selecting appropriate tools for measuring interocular delays, and that suggests that combining specific methods may better characterize temporal delays seen in visual impairments.

This work was supported by an NSERC Discovery Grant to AK, and a UTM Undergraduate Research Grant to BL.

33.459 VEP MEASURES OF MONOCULAR ATTENUATION IN AMBLYOPIA ARE EXTREMELY RELIABLE ACROSS SESSIONS AND PREDICT PERCEPTUAL RESPONSES

Kimberly Meier[1] (kmeier@uh.edu), Mark Pettet[2], Kristina Tarczy-Hornoch[2], Geoffrey M Boynton[2], Ione Fine[2]; [1]University of Houston, [2]University of Washington

Introduction: Binocular integration requires integrating information across the two eyes, while ensuring that the final percept is non-rivalrous. We examined binocular integration across non-rivalrous and rivalrous conditions, across both VEP and psychophysical measurements. Methods: Amblyopia and control observers (n=11 per group) viewed dichoptic gratings (orthogonal vs. same orientations, 2cpd, 7.5 Hz flicker) and reported perceived contrast via joystick while occipital VEP signals were recorded. Grating contrasts modulated independently in each eye (1/8 and 1/6 Hz). VEP and joystick responses were predicted using a model that captured attenuation (k), and a continuum between averaging and winner-take-all behavior (w), by finding the best-fitting values of k and w, using the equation (1-w)[mean(kA,F)] + (w)[max(kA,F)], where F=Fellow and A =Amblyopic eye. Results: Attenuation in the weaker eye (k) was similar across orthogonal and same orientation gratings for both VEP and joystick measurements. Amblyopic individuals had higher levels of attenuation than controls (p=0.030). In amblyopia, estimates of attenuation were highly consistent across the two visits, for both perceptual (r=0.77) and VEP (r=0.92) responses. This consistency across sessions and stimuli suggests that attenuation reflects a fundamental neural property that can be objectively and accurately assessed using VEP; offering a promising pathway for clinical assessment in populations unable to provide reliable behavioral responses. Measures of binocular integration (w) were similar across individuals with amblyopia and controls and did not differ for orthogonal vs. same orientation stimuli. Values of w were close to a mean rule for VEP responses (w=0.06), while perceptual responses were close to the max (w=0.66), suggesting that VEP responses reflect an earlier stage of binocular integration than joystick responses. Results suggest that a winner-take-all mechanism, that does not differ across individuals with amblyopia and controls, results in a moderate level of early attenuation having a disproportional effect on which eye dominates perception.

Knights Templar Eye Foundation, Research to Prevent Blindness

33.460 FOOTSTEPS ILLUSION IN BINOCULAR INTERACTION

Eun Cho[1] (choeun@unist.ac.kr), Chang-Yeong Han[1], Oh-Sang Kwon[1], Hyosun Kim[2]; [1]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, South Korea, [2]R&D center, Samsung Display, South Korea

The footsteps illusion (FI) is a phenomenon where the perceived speed of a moving stimulus changes depending on its background (Anstis, 2001). FI provides a unique tool for studying motion and positional integration in binocular vision because it allows independent control over motion and position signals. Here, we used FI stimuli to investigate how the visual system integrates motion and positional information from binocular inputs, and whether the binocular differences contribute to the depth perception. Three experimental conditions were tested. In FI condition, the physical speed of the moving stimulus was constant while the perceived speed fluctuated due to the FI effect. In constant FI condition, the physical speed followed a square wave patten, while the perceived speed remained nearly constant because of the FI effect. In control condition, the physical speed followed a square wave pattern without FI effect. Motion phase differences between the two eyes ranged from 0° to 360°. At 180° phase, the perceived or physical speed of the stimulus in one eye was maximumly different from that in the other eye. In Experiment 1, participants observed these stimuli in their peripheral vision and reported whether the stimulus appeared to move with constant velocity. Experiment 2 repeated the same conditions, but the stimulus was viewed in the fovea. In Experiment 3, participants observed the same set of stimuli in the fovea and reported whether they perceived motion in depth. The results showed that, although FI stimuli generate strong FI effects in monocular vision and produce significant interocular velocity differences (IOVD), they fail to contribute to depth perception. In contrast, under constant FI conditions, depth was perceived in the fovea despite the absence of detectable IOVD. These findings suggest that differences in position and motion uncertainty between the fovea and periphery play distinct roles in binocular perception.

This work was supported by the Samsung Display Research Center.

33.461 ABERRANT NEURAL INFORMATION ACCUMULATION DURING BINOCULAR RIVALRY IN SCHIZOPHRENIA

Scott Sponheim[1], Joshua Stim[2], Victor Pokorny[3]; [1]Minneapolis VAMC / University of Minnesota, [2]Johns Hopkins University, [3]Northwester University

Introduction: Psychotic disorders, such as schizophrenia, are in part defined by disturbances in perception (e.g., hallucinations). Despite decades of research on visual perception in people with psychotic psychopathology the neural processes that underlie abnormal percept formation are unknown. Methods: To better understand the neural functions that may lead to abnormal perception we had 57 people with psychotic psychopathology (41 schizophrenia, 16 bipolar disorder with psychosis) 25 siblings of these people, and 43 healthy comparison participants complete a binocular rivalry task while magneto-encephalography (MEG) was recorded. Average percept durations during rivalry were computed from the behavioral responses of each participant. We also used SSVEPs derived from frequency-tagged stimuli of the two rival percepts to investigate neural activity related to information accumulation during perception. We submitted the time course of SSVEPs to a frequency decomposition to characterize the dominant frequency of the rise and fall of SSVEP power associated with perceptual switching. Results: Individuals with psychotic psychopathology and their siblings failed to differ from controls in their percept durations based on button presses related to changes in percept. Inspection of the frequency composition of the SSVEP time courses revealed a dominant .5 Hz frequency that appeared reduced in people with psychotic psychopathology compared to their siblings and control participants. Paired comparisons revealed that individuals with schizophrenia had lower power than controls at .5 Hz (t56.3 = -2.13, p =.038). This reduction was absent in individuals with bipolar disorder (t24.5 = -1.57, p =.129). Discussion: Our analysis of neural activity reflective of processing frequency-tagged stimuli during binocular rivalry revealed anomalies in neural information accumulation in people with schizophrenia. Reduced power of the dominant frequency within SSVEP time courses may represent reductions in neural activity that contribute to atypical percept formation in schizophrenia.

This work was funded by VA Merit Grant I01CX000227 and NIMH grant R01MH112583 (SRS as PI)

33.462 HOW DOES TEMPORAL FREQUENCY INFLUENCE BINOCULAR BALANCE IN NORMAL AND AMBLYOPIC OBSERVERS?

Seung Hyun Min[1], Chenyan Zhou[1], Jiawei Zhou[1]; [1]School of Ophthalmology and Optometry and Eye Hospital, and State Key Laboratory of Ophthalmology, Optometry and Vision Science, Wenzhou Medical University, Wenzhou, Zhejiang, China

This study examined how temporal frequency influences binocular balance in normal and amblyopic observers. Twenty-three controls and 13 individuals with amblyopia participated in this study. Binocular balance was measured using an onset binocular rivalry task with sinusoidally flickering gratings at various temporal frequencies and static gratings paired with monocular attentional cues. For the flickering gratings, different combinations of temporal frequencies (2, 4, or 10 Hz in one eye vs. 2, 3, 4, 6, 10, 15, or 20 Hz in the other) were presented. Their effects were then compared, and their relationships were analyzed. The results showed no link between shifts in binocular balance due to temporal frequency and those caused by monocularly directed attention in either group. Intermediate temporal frequencies (6.5 - 9 Hz) in one eye maximized its perceptual dominance, with a larger shift due to temporal frequency in amblyopes than in controls. In normal observers, the balance shifts due to temporal frequency and attentional (active and passive) modulation were similar, whereas amblyopic observers exhibited a greater shift from temporal frequency than from monocularly focused passive attention. The findings indicate that intermediate temporal frequencies in one eye, rather than specific temporal frequency differences between the eyes, maximized perceptual dominance in both groups. Additionally, the influence of temporal frequency on binocular balance was stronger than that of monocularly directed passive attention in amblyopic individuals.

Attention: Neural mechanisms

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.463 CAUSES AND CONSEQUENCES OF LOSING CONSCIOUSNESS: INTRACRANIAL ELECTROENCEPHALOGRAPHY EVIDENCE IN HUMANS

sheng lin[1], qi chen[1]; [1]South China Normal University

In recent years, there has been extensive research on human consciousness in response to stimuli near the perceptual threshold. However, studies focusing on consciousness elicited by suprathreshold audiovisual stimuli remain scarce. Through simple visual and auditory experiments employing intracranial electroencephalography (iEEG), we investigated the mechanisms underlying the loss of consciousness in humans and the regulation of primary cortical areas by higher-order cortical areas following consciousness loss.Our findings reveal that the excitation-to-inhibition (E/I) ratio in primary cortical areas is a direct determinant of consciousness loss. Furthermore, the E/I ratio is influenced by the connectivity between higher-order and primary cortical regions. Following consciousness loss, not only do the temporal dynamics of responses across brain networks change significantly, but higher-order cortical areas also regulate primary cortical areas in preparation for subsequent stimuli.

33.464 EFFECTS OF COVERT ATTENTION ON THE PERCEPTUAL ENCODING OF NATURALISTIC SCENES

John E. Kiat[1] (jekiat@ucdavis.edu), Steven J. Luck[1]; [1]University of California-Davis

Prior research has shown that covert attention enhances the perceptual processing of attended locations and stimuli, boosting behaviorally measured discriminability and neural response amplitudes. However, the time point during perceptual processing at which attended information becomes more discriminable than unattended information is unclear. Magnitude-based neurophysiological measures with high temporal precision (e.g., the N2pc or Alpha Power Suppression) provide useful information about attention but have important limitations. For instance, observing an N2pc in response to a visual search target is certainly evidence for when covert attention becomes focused on the target, but it does not provide direct evidence for when information about the attended object becomes more discriminable than information about unattended objects. To address this issue, we applied multivariate pattern analysis to ERPs to analyze the timecourse of the separability of the neural representations of covertly attended versus unattended naturalistic scene images. Participants (N = 28) were presented with a series of trials, with each trial consisting of two different scene pairs presented simultaneously at fixed locations. Each pair was preceded by a cue indicating which of the two scenes should be attended. We found greater representational similarity between the ERPs and a computational model of early visual processing for the attended scenes beginning at 80 ms poststimulus. At approximately 250 ms, we found greater representational similarity between the ERPs and a computational model of more abstract visual representations as well as greater categorical decoding accuracy for the attended scenes. These findings show that covert attention can boost the neural representation of simple visual features during early visual processing, with more categorical-based separability likely emerging later.

Grant support: This work was made possible by grant R01EY033329to SJL.

33.465 EFFECTS OF INACTIVATING PRIMATE SUPERIOR COLLICULUS ON FRONTAL EYE FIELD NEURONS DURING A SELECTIVE ATTENTION TASK

Leor N Katz[1], Ricahrd J Krauzlis[1]; [1]NIH

The primate superior colliculus (SC) plays a crucial role in visual selective attention, and how it interacts with cortical attention mechanisms remains a controversial area of research. Here we investigated the effects of inactivating primate SC on neurons in the frontal eye fields (FEF), a prefrontal cortical region also important for the control of attention and eye movements. We used multichannel probes to record electrophysiological signals in FEF of two macaque monkeys as they performed a selective attention task, before and during reversible inactivation of SC using muscimol. In the attention task, patches of visual motion were presented on either side of a central fixation point, one of which was positioned within the response fields of the recorded FEF neurons. In “Attend” blocks, monkeys were rewarded for attending to the motion patches and releasing a joystick in response to a change in motion direction. In “Ignore” blocks, reward was delivered for ignoring the change in motion direction and instead, for responding to a change in fixation point luminance. FEF neurons were strongly modulated by the attentional manipulation, exhibiting significantly stronger responses during trials of the Attend blocks versus Ignore (i.e., attention-related modulation). They also exhibited a strong phasic increase in activity in response to the motion stimulus change, and this response was stronger for “hits” versus “misses” (i.e., ‘detect-probability’). Inactivation of SC caused reductions in hit rates during the attention task, as well as reductions in the attention-related modulations of FEF neurons, both during stimulus presentation and during the change in motion direction. The causal effect of SC on FEF neurons implicates the FEF in a network of cortical areas that interact with the SC during attention tasks and suggests that ascending pathways from SC to the prefrontal cortex are important for enabling the FEF contributions to visual selective attention.

33.466 FUNCTIONAL BRAIN NETWORK DYNAMICS CAPTURE CONTEXT- AND MODALITY-GENERAL FLUCTUATIONS IN SUSTAINED ATTENTION

Anna Corriveau[1,2], Jin Ke[1,2,3], Monica D. Rosenberg[1,2,4]; [1]Department of Psychology, The University of Chicago, [2]Institute for Mind and Biology, The University of Chicago, [3]Department of Psychology, Yale University, [4]Neuroscience Institute, The University of Chicago

Sustained attention is supported by distributed brain networks. However, while sustaining focus in the real world requires maintaining multimodal information, studies of the brain networks underlying this ability have almost exclusively examined visual attention. Here, we test whether common neural mechanisms underlie sustained attention across visual and auditory modalities, and across controlled and naturalistic tasks. In a two-session fMRI study, participants performed a continuous performance task in which streams of trial-unique sounds and images were presented simultaneously. They were instructed to attend either images or sounds and press a button when the relevant item belonged to a frequent (90%) but not infrequent (10%) category. To isolate brain networks tracking attention dynamics, we parcellated fMRI data into 400 cortical and 32 subcortical regions and calculated pairwise edge co-fluctuation time series, or the product of z-scored activation time series, between all pairs of brain regions. General linear models were fit to identify pairs of brain regions whose co-fluctuations predicted lapses in sustained attention. Analyses revealed edges whose dynamics were related to visual (5465 positive, 6678 negative edges) and auditory (1849 positive, 2401 negative edges) attention lapses, as well as edges common to both types of errors (~6.4% of selected edges, overlap significant p<.001). We then tested whether fluctuations in the strength of these edges predicted subjective fluctuations in attention to narratives. We correlated edge co-fluctuation strength with continuously-reported narrative engagement during four naturalistic stimuli: two audiovisual movies, one silent movie, and one podcast. While predictions from edges related to visual attention lapses were unreliable, the subset of edges involved in both visual and auditory attention lapses predicted changes in engagement in all narratives (Pearson's rs=.014-.055, all ps<.08). Results suggest that sustained attention relies on modality-general networks which capture attentional fluctuations across contexts.

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

33.467 NONINVASIVE NEUROMODULATION OF HUMAN SUBCORTICAL VISUAL PATHWAYS WITH TRANSCRANIAL ULTRASOUND STIMULATION

Ryan Ash[1], Martin Scott[1], Patricia Limon[1], Morteza Mohammadjavadi[1], Kim Butts Pauly[1], Anthony Norcia[1]; [1]Stanford University

Clinical neuroscience has led to a revolution in our understanding of how different neural circuits contribute to sensory processing, cognition, emotion, and behavior. Unfortunately, the tools to causally investigate these circuits in the human brain are limited by poor focality and depth penetration. Transcranial ultrasound stimulation (TUS) is an emerging tool to achieve noninvasive focal brain-wide neuromodulation with high focality (<1cm) and the ability to achieve high intensities in-depth. This technology is at an early stage of development, and many optimizations are needed. One need is to better understand how different sonication parameters relate to neuromodulatory effects. Key parameters include pulse repetition frequency (PRF), intensity, and duty cycle (DC). Importantly, by varying these parameters it may be possible to have either suppressive/inhibitory effects or facilitatory/excitatory effects on neural activity and synaptic strength. We developed a paradigm to target TUS to the human lateral geniculate nucleus (LGN) as an efficient testbed to evaluate TUS effects. We implemented steady-state visual evoked potential (ssVEP) measures of contrast-response and contrast increment detection psychophysics as neural and behavioral readouts of subcortical visual pathway function. We developed a neuroimaging and simulation pipeline to target LGN, and we are using a neuronavigated depth-steerable 4-element TUS transducer. Our preliminary data suggests that 5 Hz TUS to left LGN reversibly suppresses VEP responses specifically to stimuli in the contralateral hemifield (P=0.001, n=8 participants), while 500 Hz TUS shows a trend toward enhanced responses to contralateral stimuli (P=0.1). We are now measuring the effect of TUS across the full range of visual contrasts. This work provides the foundation for a dissection of the roles of subcortical and deep cortical nuclei in contrast normalization, attention, object recognition, and other visual functions.

National Eye Institute

33.468 DISSOCIATING EFFECTS OF SPATIAL ATTENTION AND MICROSACCADES ON NEURAL ACTIVITY IN EARLY TO MID-LEVEL VISUAL AREAS.

Wojciech Zajkowski[1], Katrina Quinn[2], Lenka Seillier[3], Daniel A. Butts[4], Hendrikje Nienborg[1]; [1]NIH, [2]University of Tubingen, [3]Humboldt Foundation, [4]University of Maryland

Microsaccades are small, involuntary eye movements that occur during visual fixation, and are thought to play a role in maintaining visual perception and preventing sensory adaptation. Microsaccades have long been linked to spatial attention, both as a consequence of attentional states, and as a potential mechanism for enhancing sensory processing. Here, we used statistical modeling to characterize the effects of attention and microsaccades on neural firing rates. We recorded extracellular single and multi-unit activity in early-to-mid visual areas (V1: 88 units, V2:1150, V3: 852, V4: 89) in 2 macaque monkeys, which performed a disparity discrimination task (113 sessions total), in which spatial attention was manipulated by presenting two stimuli simultaneously on both sides of the visual field, but only one of which was task-relevant. We modeled the neural spiking activity of each cell with a GLM, using stimulus and saccade onsets as predictors, and controlling for the effects of neural drift and within-trial adaptation. We show that, unlike prior reports, neural activity is affected by attention independently of microsaccadic modulation, even when considering the effects of microsaccade direction and correlations with attentional state. Across areas, a model where attention modulated stimulus processing had a higher log-likelihood for 98.6% of cells and explained 11% more of the maximal variance captured by the full model, than the model where attentional effects could only be implemented via direction-dependent saccadic modulation. Our results suggest that spatial attention modulates

neuronal responses independently of saccades, consistent with an attention-dependent gain of the stimulus driven response. Moreover, these attentional effects are associated with an additive shift of the microsaccade kernels in early-to-mid visual areas.

This work was funded by the National Eye Institute Intramural grant 1ZIAEY000570-01 (H.N.) and NSF grant NCS-FO 2123568 (D.A.B.)

33.469 VALUE PROCESSING IN THE FRONTAL EYE FIELD OF NON-HUMAN PRIMATES

Xuefei Yu[1], Atul Gopal[1], Okihide Hikosaka[1]; [1]NIH

Value can originate from different sources, such as being associated with specific objects or locations. While extensive research has explored the neural underpinnings of object and location value processing, the precise brain regions, and mechanisms through which these distinct value sources are processed and integrated to motivate actions in non-human primate remain unclear. In this study, we trained two macaque monkeys to perform two value association tasks: an object value task, where objects were associated with high or low rewards, and a location value task, where value was linked to the location where objects were displayed. After training, both monkeys successfully established value associations in both tasks. This was evidenced by faster saccadic eye movements toward high-value object and locations, as well as a preference for high-value conditions in choice behavior. Neuronal population responses in the primate Frontal Eye Field (FEF) showed significant modulation based on value associations. Among the neuronal population, over half encoded a single source of value (either object or location), while a subset of neurons encoded both object and location values. The separation of object value signals emerged approximately 100ms after object onset, whereas location value signals appeared before target onset, indicating stronger anticipatory modulation for location value. These findings suggest that object and location values partially converge in the FEF, supporting a flexible coding scheme that accommodates both integrated and distinct encoding of object and location values.

33.470 REPRESENTATION OF OBJECTS, ATTENTION, AND LOAD IN HUMAN PREFRONTAL CORTEX

Shijie Qu[1] (shijie.qu@yale.edu), Marvin M. Chun[1], Yaoda Xu[1]; [1]Yale University

Visual processing often involves attending to different object features under various task loads. Past studies reported that some object and task features are encoded independently, while others are represented interactively in the human occipitotemporal cortex (OTC) and posterior parietal cortex (PPC). However, the mechanisms underlying their coding in the prefrontal cortex (PFC) are not fully understood. As a hub for cognitive control, it is possible that PFC primarily encodes task-related features, such as attention and load, but not object identity. Alternatively, it may simultaneously encode object identity alongside attention and load. If the latter holds, delineating how these distinct features are coded together in PFC can provide a mechanistic understanding of how it processes diverse information during visual tasks. In this fMRI study, 12 human participants performed an n-back task on colored visual stimuli, with task load (1-back and 2-back) and attention content (color and shape) varied orthogonally. Using multi-voxel pattern analysis, we found that PFC can decode not only task features such as attention and load, but also object identity. Comparison across brain regions showed that load representation was the strongest in PFC while object representation was the strongest in OTC. Attention representation, however, appears to be weaker in PFC than in OTC. Additional analyses revealed that while object representation only interacts with attention in OTC and PPC, it is further modulated by task loads in PFC. Apart from these differences, the three regions converge in that attention representation is modulated by load and object, while load representation remains independent of attention and object. Overall, our results extend prior findings and show a representational gradient across the human brain where object identity information decreases, and load information increases from posterior to anterior regions. We further show that PFC employs distinctive mechanisms to encode objects interactively along with task demands.

This research was supported by NIH grant R01EY030854 to YX.

33.471 TASK-SPECIFIC NEURAL MECHANISMS OF ATTENTIONAL CONTROL: EVIDENCE FROM CROSS-TASK DECODING

Melinda Sabo[1,2] (sabo@ifado.de), Edmund Wascher[2], Manuel Varlet[3], Tijl Grootswagers[3]; [1]Wu Tsai Institute | Yale University, [2]Leibniz Research Centre for Working Environment and Human Factors, [3]MARCS Institute, Western Sydney University

Attentional control (also known as cognitive control) is a fundamental function that regulates information processing to align with goal-directed behavior (von Bastian et al., 2020). Traditionally, its role has been demonstrated through classical tasks like the Stroop, Simon, and Go/No-Go, which have been argued to involve conflict resolution. However, recent behavioral evidence suggests that these tasks do not reflect a single, domain-general construct of attentional/cognitive control but rather task-specific processes. This raises the question: do these tasks rely on a shared neural mechanism orchestrated by a unified attentional/cognitive control network, as traditionally proposed (e.g., Gratton et al., 2017)? To address this, we analyzed data from the ongoing Dortmund Vital study, involving ~500 participants aged 20–70 years. Participants completed tasks traditionally associated with attentional/cognitive control—Simon, Stroop, Go/No-Go, and a perceptual discrimination task—while their brain activity was recorded using electroencephalogram (EEG). A linear classification model was trained on EEG data to distinguish conflict from non-conflict trials within each task. While the classifier successfully learnt this distinction within tasks, cross-task decoding revealed below-chance performance. Specifically, a leave-one-out cross-task decoding procedure (i.e., training on three tasks and testing on the fourth) provided strong evidence against generalization across tasks. Notably, these results remained consistent even when decoding parameters (e.g., input features, different preprocessing steps etc.) were varied. These findings suggest that the neural mechanisms underlying conflict processing are task-specific, likely governed by highly specialized and potentially non-overlapping attentional/cognitive control sub-networks. This aligns with prior behavioral evidence and challenges the traditional view of a shared neural substrate for cognitive control across tasks.

The cross-institute collaboration was supported by the International Visiting Scholar Program initiated by the MARCS Institute (Western Sydney University)

33.472 ATTENTION-RELATED N2PC COMPONENT OF THE VISUAL EVOKED POTENTIALS AS A MARKER OF FINE-GRAIN SHIFTS OF ATTENTION WITHIN THE FOVEOLA

Yue Guzhang[1], Alessandro Benedetto[2], Martina Poletti[1]; [1]University of Rochester, [2]University of Florence

Previous work demonstrated that covert attention can be selectively directed within the 1-deg foveola, enhancing fine spatial vision. The N2pc component, a well-known EEG marker of covert attention, has been extensively studied extrafoveally using large stimuli spanning >1-deg of visual angle. Investigating the N2pc component when attention is selectively shifted within the foveola between stimuli close to the resolution limit is challenging due to difficulties in maintaining precise stimulus lateralization in the presence of constant fixational eye movements. Here, we circumvented these challenges and examined whether the N2pc component is associated with selective attention shifts within the foveola when observers focus on fine spatial detail. Participants (N = 11) were instructed to attend to either a red or green stimulus. The color assignment was counterbalanced across blocks. While maintaining central fixation, a red and a green square (0.12° in size) were briefly presented 0.33° to the left and right of fixation. One square had a small gap positioned either at the top or bottom. Participants reported the gap location as quickly and accurately as possible. In informative blocks, the gap appeared in the cued stimulus, whereas in uninformative blocks, it appeared randomly between the two stimuli. Leveraging high-precision eye-tracking and retinal stabilization, we ensured that the stimuli remained at the same retinal location throughout each trial. Participants discriminated high-acuity stimuli more accurately ($p=0.0467$) and responded faster ($p=0.0019$) in informative than uninformative trials, confirming that selective attention was engaged. Additionally, a clear N2pc component was observed in both types of trials. Yet, its amplitude was larger in informative trials ($-2.05\mu V \pm 1.34\mu V$ vs. $-0.80\mu V \pm 0.87\mu V$, $p = 0.0064$). These findings show that even fine-grain visual attention shifts within the foveola can reliably elicit an N2pc, with amplitude varying based on the attentional focus.

NIH R01 EY029788-01 and Sloan Fellowship

33.473 ONGOING ALPHA OSCILLATIONS BIAS DECISION-MAKING IN WILLED ATTENTION

Changhao Xiong[1], Sungkean Kim[2], Qiang Yang[1], Srinivasan Meyyappan[3], George Mangun[34], Mingzhou Ding[1]; [1]J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, [2]Department of Human-Computer Interaction, Hanyang University, [3]Center for Mind and Brain, University of California, Davis, [4]Departments of Psychology and Neurology, University of California, Davis

It is well established that ongoing alpha oscillations (8 to 12 Hz) influence visual stimulus processing. In this study, we investigated whether and how ongoing alpha oscillations biased decision making in willed attention. Electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) data were recorded from participants performing a cued visual spatial attention task. Each trial started with one of three cues. Two of them instructed the participant to covertly attend a spatial location in the left or the right visual field (instructed attention). The third cue prompted the participant to spontaneously decide which spatial location to covertly attend (willed attention). Following a variable cue-target interval, the participant discriminated the spatial frequency of a grating appearing at the attended location and ignored the grating appearing at the unattended location. Subjecting the distributed pattern of EEG alpha-band power in the pre-cue time period (-500 to 0 ms) to multivariate pattern analysis (MVPA), we found that (1) for instructed attention, the accuracy of decoding the direction of attention was at chance level, as expected and (2) for willed attention, however, the accuracy of decoding the direction of attention was significantly above chance level and significantly correlated with the cue-evoked fMRI activities in the salience and the central executive networks. These results revealed that (1) ongoing alpha oscillations biased decision making in willed attention and (2) it was doing so by modulating the activity in the frontoparietal decision-making networks.

33.474 RECONSTRUCTING MOVING OBJECT LOCATIONS IN AN IMMERSIVE 3D VIRTUAL ENVIRONMENT FROM EEG OSCILLATORY ACTIVITY

Tom Bullock[1] (tombullock@ucsb.edu), Emily Machniak[1], Neil Dundon[1], Justin Kasowski[1], You-Jin Kim[1], Radha Kumaran[1], Julia Ram[1], Melissa Hernandez[1], Stina Johansson[1], Tobias Höllerer[1], Barry Giesbrecht[1]; [1]University of California, Santa Barbara

Many everyday activities require interaction with moving objects, including catching a ball or navigating a busy street. To create a cohesive mental model of the world, we select goal-relevant objects and maintain stable representations of their locations as they move through time and space. Location-selective representations of static objects are supported by EEG oscillations in the alpha and theta frequency bands, but how these representations are constructed and maintained for dynamic objects is not known. To address this issue, we recorded EEG at the scalp while participants (n=34) engaged in an immersive virtual reality (VR) task where colored spheres appeared at a distant location (30m in the VR environment) and moved towards the participant. Participants were provided a pair of virtual lightsabers and required to use one to strike a color-defined target sphere. Participants completed trial blocks where targets were either visible throughout their entire trajectory (control) or briefly disappeared mid-trajectory (500 ms) before reappearing in either a predictable or unpredictable location (1200 ms), just before the lightsaber strike (~1600 ms). Accuracy was high but dropped in the non-predictive condition (mean±SEM: .87±.01) relative to predictive (.95±.01) and control (.96±.01) conditions ($p<.05$). We applied inverted-encoding modeling to alpha-band activity and successfully reconstructed locations of targets throughout their trajectory starting at ~250 ms and even when stimuli disappeared mid-trajectory. Reconstructions were diminished mid-trajectory in non-predictive relative to predictive conditions. This might be interpreted as attention becoming more diffuse when it is likely that the target will appear at another location. We also reconstructed target locations from theta activity immediately before the lightsaber strike in predictive and non-predictive conditions. Together, these results provide insight into how the brain represents predictable and unpredictable goal-relevant moving objects, and validate a new framework for studying dynamic attention in immersive VR.

Funding statement: Research was sponsored by the U.S. Army Research Office and accomplished under contract W911NF-09-D-0001 for the Institute for Collaborative Biotechnologies.

Attention: Spatial

Sunday, May 18, 8:30 am – 12:30 pm, Pavilion

33.475 DISSOCIATION BETWEEN ATTENTIONAL AND OCULOMOTOR HABITS

Emma Holtz[1] (holtz257@umn.edu), Chen Chen[2], Vanessa Lee[3]; [1]University of Minnesota

Attention and eye movements often align in visual tasks, but they can also dissociate, as when people shift attention without moving their eyes. Most studies have examined these systems over short timescales, capturing momentary attention or eye movements. Here, we explored their interaction over a longer timescale using a location probability learning paradigm. In Experiment 1, participants searched for a target that frequently appeared in one quadrant, developing both an oculomotor habit (initial saccades toward the high-probability quadrant) and an attentional habit (faster search when the target appeared in the high-probability region). Both habits emerged simultaneously and persisted in a neutral testing phase with random target locations. This coupling broke down in Experiments 2 and 3, where participants were cued to saccade toward specific quadrants that aligned or misaligned with the high-probability target quadrant. In a spatially-unbiased testing phase without the cue, the oculomotor habit persisted toward the previously cued quadrant, while search speed was fastest in the high-probability area, unaffected by prior cuing. Thus, while oculomotor and attentional habits are often coupled, they arise from distinct mechanisms: oculomotor habits are driven by eye movement history, and attentional habits by search success.

33.476 ENDOGENOUS ATTENTION ENHANCES CONTRAST SENSITIVITY SIMILARLY AROUND CARDINAL MERIDIANS DESPITE DIFFERENTIAL ADAPTATION EFFECTS

Hsing-Hao Lee[1] (hsinghaolee@nyu.edu), Marisa Carrasco[1]; [1]New York University

[Background] Visual adaptation and covert spatial attention help manage bioenergetic resources. Adaptation decreases contrast sensitivity toward repeated features; covert attention increases it at the attended location but decreases it at the unattended location. Visual performance is better along the horizontal than vertical meridian (horizontal-vertical anisotropy, HVA), and better along the lower than upper vertical meridian (vertical meridian asymmetry, VMA). Adaptation reduces contrast sensitivity more at the horizontal than vertical meridian; covert endogenous attention enhances contrast sensitivity similarly around polar angle. Here we investigated whether endogenous (voluntary) attention enhances contrast sensitivity after adaptation and how this relation varies across meridians. [Methods] To equate discriminability across the four cardinal locations, we titrated contrast threshold while observers performed an orientation discrimination task (±2.5° off the vertical axis, 5 cpd) in the neutral condition, with and without adaptation to 100%-contrast vertical Gabor adaptors. We then tested the effect of endogenous attention along the horizontal and vertical meridians (blocked), using a pre-cue indicating the target location (valid cue), a distractor location (invalid cue), or all locations (neutral cue), with and without adaptation. [Results] (1) Titration yielded lower thresholds at the horizontal than vertical meridian (HVA) and at the lower than upper vertical meridian (VMA) in non-adapted trials; (2) With equated discriminability, adaptation effects were stronger at the horizontal than vertical meridian; (3) Attention improved performance (d') at validly-cued locations and reduced it at invalidly-cued locations in both adapted and non-adapted conditions, and the extent of the attentional effect was similar across meridians. [Conclusion] This study confirmed visual field asymmetries and a stronger adaptation effect at the horizontal than vertical meridian (Lee & Carrasco, VSS 2024). Critically, it revealed that, despite this differential adaptation effect, endogenous attention enhances contrast sensitivity, with comparable efficacy across the visual field.

This study was supported by the Ministry of Education in Taiwan to HHL and NIH NEI R01-EY027401 to MC

33.477 HIDDEN IN PLAIN SIGHT: PUPIL CHANGE AND PUPIL ASYMMETRY REFLECT INFANT COVERT ATTENTION DEVELOPMENT

Victoria Jones[1] (vjones27@vols.utk.edu), McKenzie Sheets[3], Micalee Segers[2], Shannon Ross-Sheehy[4]; [1]University of Tennessee

When visually attending to a stimulus, viewers typically direct their eyes towards the item of their attention (overt attention). However, covert attention shifts even before the eyes move, likely facilitating accurate eye movements to the attended location (Posner, 1989). Spatial cueing tasks use a small spatial precue to elicit covert attention shifts, facilitating faster reaction times when targets appear in the cued location. Although adults inhibit eye movements during spatial cueing tasks, infant tasks necessarily rely on eye movements to assess spatial attention (Ross-Sheehy et al, 2015), making it difficult to assess covert attention. Fortunately, recent adult work used pupillometry to assess covert attention prior to a saccade (Salvaggio et al, 2022). Additionally, individual differences of pupil asymmetry may be linked to spatial attention control (Meyyapan et al, 2023). If pupil change (PC) and pupil asymmetry (PA) reveal covert attention, then they may be useful for assessing infant attentional development. To explore this, infants were tested in a cued-attention task and eye-tracking was used to monitor pupil prior to eye movements. If PC reflects covert attention, then we expect strong attention-related pupil changes, particularly for the eye contralateral to the cue and/or target. Infants were tested in a spatial cueing task (N=234 5, 8, and 11-mo-olds). Gaze and pupil were tracked binocularly (300Hz). LME models were used to assess PC and PA. Preliminary model results reveal several significant effects, including significant pre-movement PC ($\beta=0.0004$, $p<0.001$) and PA, with ipsilateral targets producing the greatest dilation ($\beta=0.023$, $p<0.001$). Moreover, PA was greatest for valid and baseline cue conditions ($\beta=0.026$, $p<0.001$), suggesting PA may be a useful assessment both of attention orienting and cue effects. Poster analyses will fully explore pupil change and pupil asymmetry as a function of age and cue condition.

33.478 LARGE FLANKER EFFECT DURING A FOCAL CUE MAY BE THE RESULT OF ATTENTIONAL RESETTING

Joe Opdenaker[1], Miranda Scolari[1]; [1]Texas Tech University

The size and extent of a cued area can induce a focal or diffused spotlight of attention. Theories of visual attention such as the zoom lens and gradient models explain how attention allocation can be guided by pre-stimulus cues and redistributed to accurately process stimuli. We conducted a flanker task (n = 38) where a pre-cue encompassed the upcoming target location (focal cue) or one that encompassed both the target and distractor locations (diffuse cue) to investigate the effects of pre-cue size on performance. As expected, a significant flanker effect in accuracy emerged in both focal and diffuse cue conditions. Unexpectedly, the flanker effect in the focal cue was significantly larger than the diffuse cue, suggesting a breakdown in attentional focus when competing distractors were presented. This pattern of results may have occurred due to a redistribution of attention following the sudden onset of near (<1 degree), salient distractor stimuli. Competing distractors may have broadened the attentional spotlight or shifted the attentional gradient outward to include distractors. In contrast, the diffuse cue preemptively distributed attention more evenly across the target and distractor location, reducing the likelihood of an attentional redistribution following array onset, resulting in a smaller flanker effect. These findings suggest that attentional resetting may be driving changes in resource allocation shaped by the extent of pre-stimulus cue that do not initially include near distractors. Two follow up studies are underway to test this hypothesis.

33.479 MAPPING ATTENTION ACROSS THE VISUAL FIELD WHEN TRACKING MOTION AROUND THE HEAD

Nina M. Hanning[1,2] (hanning.nina@gmail.com), Martin Rolfs[1,3]; [1]Humboldt-Universität zu Berlin, Berlin, [2]New York University, New York, [3]Bernstein Center for Computational Neuroscience, Berlin

How does attending to a location outside our field of view influence what we see within it? To explore this question, participants were seated within a 360° panoramic setup (diameter = 3.14m), consisting of a cylindrical screen and six projectors evenly spaced at 60° intervals. To map visual attention in this scenario, we combined a dot tracking task with an orientation discrimination protocol. A dynamic full-field 1/f noise was displayed across observers' entire (horizontal) visual field, which we determined for each participant. Participants mentally tracked a dot that circled around them horizontally at a constant speed of 45°/s, while keeping their gaze fixed ahead. Fixation was monitored continuously using a head-mounted eye-tracking system. The tracked dot disappeared, accompanied by an auditory cue, either within view, or behind them. Simultaneously, we presented ±45deg orientation-filtered 1/f noise as a local discrimination signal. After a brief masking period, participants then localized the dot's last location and indicated their orientation judgement (cw/ccw, 2AFC). We analyzed their discrimination performance relative to the estimated dot offset as a measure of visual attention. Localization accuracy deteriorated gradually as soon as the dot moved out of sight, showing a pronounced underestimation of distance travelled. Meanwhile, mentally tracking the dot behind the head did not harm visual sensitivity ahead: Participants' performance was high and largely unaffected by the tracking task. Surprisingly, sensitivity showed a spatially specific improvement at the location opposite to the estimated dot position. These results suggest a systematic mapping of locations behind the head on the frontal visual field. We speculate that this mapping may reflect a preparatory mechanism, where attention prioritizes locations ahead in anticipation of movement forward, away from a potential threat.

This research was supported by a Marie Skłodowska-Curie individual fellowship by the European Commission ([898520 – PremotorPerception] to NMH) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme ([865715 – VIS-A-VIS] to MR).

33.480 EXPLORING THE FLIP TILT ILLUSION IN CENTRAL VISION BY IMPAIRING THE TOP-DOWN FEEDBACK VIA BACKWARD MASKING

Li Zhaoping[1] (li.zhaoping@tuebingen.mpg.de); [1]University of Tuebingen, Max Planck Institute for Biological Cybernetics

Consider a gray field comprising pairs of vertically aligned dots; in each pair, one dot is white the other black. When viewed in a peripheral visual field, these pairs appear horizontally aligned. This misperception of an orientation of such dot pairs by 90 degrees is called the flip tilt illusion (Zhaoping 2020). It arises because, according to the central-peripheral dichotomy (CPD) theory (Zhaoping 2019), top-down feedback from higher to lower visual cortical areas is too weak or absent in the periphery to veto confounded feedforward signals from the primary visual cortex (V1). The white and black dots in each pair activate, respectively, on and off subfields of V1 neural receptive fields. However, the sub-fields' orientations, and the preferred orientations, of the most activated neurons are orthogonal to the dot alignment. Hence, V1 reports the flip tilt to higher visual areas, assuming that an information bottleneck admits only information via these misleading V1 signals to feedforward to downstream brain areas along the visual pathway for perceptual outcome. Top-down feedback vetoes such misleading reports by querying for additional information from V1 (e.g., through activities of other V1 neurons), but only in the central visual field according to the CPD theory, making the illusion invisible in central vision. This work investigates whether this illusion can appear in the central visual field when the top-down feedback is compromised by backward masking. This study is ongoing and has so far preliminary findings, a fuller account will be reported at the annual conference.

partly supported by the Max Planck Society and the University of Tübingen

33.481 MEMORY-GUIDED SPATIOTEMPORAL ATTENTION

Keara M. Ginell[1] (ginell.1@osu.edu), David E. Osher[1]; [1]The Ohio State University

Attention serves as a fundamental mechanism for selecting the most valuable sensory information. Attentional control encompasses three modes: goal-driven (top-down), stimulus-driven (bottom-up), and experience-driven (memory-guided). Memory-guided attention reflects how prior experiences shape attentional biases; individuals can attend to specific locations based on learned patterns rather than immediate physical salience or explicit goals. Previous research has used search tasks with repeating experimental parameters (e.g. contextual or probabilistic cueing) to probe memory-guided attention towards static cues. However, these approaches do not investigate attention operating dynamically over time. We developed a novel modification of the serial reaction time task to investigate memory-guided attention in space and time. In each trial, a gabor stimulus—tilted 2° left/right of vertical—appeared briefly in one of four screen locations. Participants reported the gabor tilt (left/right). The experiment alternated between random blocks, where targets appeared in random locations, and repeating blocks, where targets reliably appeared in a sequence of locations across trials. Participants were not informed of the repeating sequence beforehand. We found that participants used statistical learning to guide attention to specific locations at specific times, performing faster and more accurately during repeating blocks than random blocks. This was evident in group-level analyses and at the individual level. Altering the repeating sequence resulted in slower responses and more errors, indicating perseveration. In a post-task, participants reported the next stimulus location after viewing two sequential stimuli from the repeating sequence. We observed significant correlations between this post-task performance and main-task effect sizes; participants whose reaction times and accuracy benefitted most from memory-guided spatiotemporal attention performed best on the post-task. This work provides evidence for memory-guided spatiotemporal attention. Future experiments will use this paradigm to investigate neural underpinnings of memory-guided attention in space and time, and learned spatiotemporal distractor suppression.

33.482 RELIANCE ON ANCHOR OBJECTS AS SPATIAL CUES INCREASES UNDER LOW VISIBILITY

Makayla Souza-Wiggins[1] (mhsouza@ucdavis.edu), Joy J. Geng[1]; [1]University of California, Davis

Much of what we know about visual search in naturalistic scenes comes from images taken under high-visibility conditions. However, real-world searches often occur when visibility is low—such as at night, in rain, or in dimly lit rooms. We hypothesized that in these situations, the ability to use target information to guide visual search is limited, forcing observers to rely more on “anchor” objects (i.e., large objects that predict the locations of smaller related objects). If true, targets should be difficult to find when visibility is low and they appear in unexpected locations relative to their anchors. In Study 1 (N=161), we validated anchor-target spatial predictions by asking participants to indicate where target objects (e.g., a dish sponge) belong within scene images created in Unity. Their responses defined spatially congruent locations (e.g., a sponge on a sink) and incongruent ones (e.g., a sponge on a stove). In Study 2 (N=38), targets appeared in four conditions: congruent or incongruent spatial locations within high- or low-visibility scenes. Visibility was controlled by adjusting the lighting within the Unity models. Each trial concluded with a spatial memory probe, where participants clicked on a blank screen, indicating the target's location. As predicted, target search times were slower in low-visibility scenes, with the slowest times occurring when targets also appeared in incongruent locations. The effect of spatial congruency on the memory probe was also more pronounced in low-visibility scenes. These findings suggest that when low visibility hinders target detection, we rely more on anchor objects as proxies to guide attention. This dependence produces a larger spatial congruency effect during search and memory recall, as expectations based on prior knowledge compensate for perceptual limitations. These results highlight how attentional strategies adapt to visibility constraints, advancing our understanding of how visual search occurs in real-world settings.

NEI T32 Vision Training Grant

33.483 SERIAL PROCESSING OF STIMULUS IDENTITY AND ATTENTION SHIFTING STATISTICAL LEARNING

Anthony Sali[1] (saliaw@wfu.edu), Emily Oor[1]; [1]Wake Forest University

Statistical learning—the mechanism by which the visual system tracks likelihoods in the world around us—allows individuals to flexibly regulate attentional control settings, such as spatial attention shifting readiness (e.g., Sali et al., 2015), across dynamic environments. However, learned adjustments in shifting readiness do not occur in isolation and our understanding of the interaction of different statistical learning processes remains limited. Across two experiments, we investigated the processing architecture responsible for shift readiness predictions (e.g., receiving a cue to shift attention when expecting to hold) and cue stimulus identity predictions (e.g., receiving shift cue B when shift cue A is more frequent). Participants monitored one of two alphanumeric streams for an embedded cue that signaled them to hold attention at the current location or to shift attention to the opposite location and made manual responses to target digits. Experiment 1 employed four cues (two shift and two hold) such that participants could receive a high probability or low probability cue stimulus regardless of the outcome (shift versus hold) and current shift likelihood. There were substantial trial-by-trial priming effects such that exact cue stimulus repetitions were associated with shorter response times (RTs) than the other trial types. When excluding exact repetition trials, we observed additive RT costs for shift readiness and stimulus identity prediction errors. In Experiment 2, we replicated and extended these findings with twice as many shift and hold cues, allowing us to prevent any consecutive exact cue repetitions. A Bayesian analysis revealed that the likelihood of an additive relationship between the cost associated with shift readiness and stimulus identity prediction errors was approximately six times as likely as the alternative that an interaction existed. These results suggest that visual statistical learning processes governing stimulus expectations and shift readiness are distinct and constrained by a shared processing bottleneck.

33.484 SOUNDS LIKE VISION: MAPPING THE SPATIAL AND TEMPORAL CHARACTERISTICS OF CROSSMODAL CUEING

Jamal Williams[1] (jrwilliams@ucsd.edu), Yong Hoon Chung[2], Jonathan Keefe[1], Viola Stoermer[2]; [1]University of California San Diego, [2]Dartmouth College

Salient events in our environment rapidly capture attention, enabling reliable processing of relevant information. However, evidence is mixed regarding whether this exogenous attention operates similarly when cue and target information originate from distinct sensory modalities; with some work suggesting that endogenous (volitional) components of attention may be necessary when sounds cue attention towards locations in space (e.g., Mondor & Amirault, 1998; but see

Mcdonald & Hillyard, 2000). In this work we explore how external cues affect visual processing within (visual) and between (auditory) modalities. In Experiment 1, participants determined the orientation of a rapidly presented, masked Gabor following exogenous cues at one of two peripheral locations from either visual (white circle) or auditory (pink noise burst from externally mounted speakers) modalities. Even though these cues were spatially non-predictive of the target location, for validly cued locations, both visual and auditory cues improved sensitivity (d') and response times (visual: both $p < .001$; auditory: $p = .01$ & $p < .001$). In Experiment 2, we increased the possible target locations to investigate how these effects might be modulated by greater spatial distances and inherently lower cue validity (from 50 to 25%)—conditions that should reduce any strategic deployment of endogenous attention. We replicated the cuing effects, showing only modest, and similar,reductions across experiments for visual (Cohen's dz = 1.51 vs 1.24, Exp 1 vs 2) and auditory (dz = 0.94 vs 0.99) cues. In Experiment 3, we track the time course of these cuing effects and find unique characteristics for both auditory and visual cues. Taken together, these results suggest that exogenous spatial attention operates through similar mechanisms regardless of sensory modality and supporting models of exogenous attention that incorporate direct crossmodal links rather than requiring voluntary attentional control.

33.485 STATISTICAL LEARNING ADAPTIVELY ADJUSTS THE VISUOSPATIAL ATTENTIONAL FOCUS UNDER HIGHLY UNCERTAIN SEARCH CONTEXTS

Andrea Massironi[1], Giulia Spinelli[1], Carlotta Lega[2], Luca Ronconi[3], Emanuela Bricolo[1]; [1]Department of Psychology, University of Milano-Bicocca, Milan, Italy, [2]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy, [3]Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy

Within a highly entropic visual world, experience-dependent mechanisms like statistical learning allow to take advantage of environmental regularities, functionally driving attentional guidance and optimizing visual search. The way attention unravels into space has been linked to a "Mexican-Hat" profile, described as an inhibitory ring encircling the attentional focus to minimize the interference of distractors. While experience-dependent mechanisms like reward have been shown to interact with the inhibitory ring, the role of statistical learning is still underexplored. The current study addresses this gap through a psychophysical task wherein we mapped the attentional profile of 26 participants asking them to report the gap orientation of a "C" letter displayed either as a salient target (Baseline Condition) or as a non-salient probe, with the salient target - now a distractor - located at progressively increasing distances from it. In both conditions, we applied statistical learning, increasing the likelihood of the salient target being in the spatial positions immediately adjacent to the probe, namely at the distance engendering the surround inhibition. Notably, such manipulation highlighted the highly uncertain nature of the salient target, who served with the same likelihood as a real target (Baseline Condition) or as an interfering distractor (Probe Condition). We found that statistical learning cancelled the inhibitory ring, reshaping the "Mexican-Hat" profile into a non-linear gradient, bringing a performance gain in the Probe Condition and maintaining it at ceiling in the Baseline one. Crucially, our findings do not fit with either the expected target vs. distractor behavioural pattern predicted by statistical learning literature. Instead, they fit with a zooming out of the attentional focus, which adaptively reshaped the priority map to resolve uncertainty and optimize search performance across conditions. Overall, our findings bring the intriguing possibility that statistical learning effects are contingent upon the uncertainty inherent to the search context.

33.486 STATISTICAL LEARNING TO INHIBIT VISUAL DISTRACTORS: ROLE OF AWARENESS IN SETTING THE FRAME OF REFERENCE OF LEARNING

Litian Chen[1] (l.chen2@vu.nl), Freek van Ede[2], Heleen A. Slagter[3]; [1]Vrije Universiteit Amsterdam, [2]Institute Brain and Behaviour Amsterdam

Recent research shows that humans can learn to ignore a salient distractor when it more often occurs at a particular location while typically remaining unaware of this regularity. We previously found that such implicit distractor-location learning is viewer dependent, indicating that one learns how to avoid distraction from an egocentric perspective, rather than learns where distractors are more likely in the world. The current study examined whether awareness of the statistical regularity (implicit vs. explicit) affects the spatial reference frame for distractor location learning (egocentric vs. world-centered). In four experiments (n=96), participants performed an additional-singleton visual search task, displayed on a table top, while walking around the table. Critically, participants were either explicitly told (aware) of the high-probability distractor location or not (unaware) and the likely distractor location was either fixed from their perspective (egocentric) or fixed in the world (allocentric). Initial analyses replicated our past work showing that in the unaware group, learning where the distractor location was more likely exclusively occurred from a viewer-dependent perspective. They furthermore showed that awareness of the regularity enhanced the ability to ignore the distractor in a world-centered frame. These findings suggest that unconscious statistical learning is naturally confined to egocentric frames and that awareness may be a prerequisite for learning to inhibit distractor regularities in world-centered frames.

33.487 TARGET DISCRIMINATION AT DIFFERENT VIEWING DISTANCES: THE ROLE OF EXPECTANCY FOR TARGET CONFIGURATIONS

Noah Britt[1] (brittn@mcmaster.ca), Hong-jin Sun[1]; [1]McMaster University

The distribution of attention across depth has recently attracted much research. Research from detection or localization tasks has suggested that attention is distributed more strongly toward near space than far space in the 3D environment. Such 'Near Advantage' has been demonstrated through faster reaction times in localizing near than far targets. However, when the task requires a discrimination response, the results for the depth effect were mixed in the literature. The current experiments sought to examine one potential moderating factor for the effect of depth: expectancy for the target configuration over trials. Recent research has led us to believe that whether participants can predict the task-prioritized target feature on a trial-to-trial basis may result in attention being allocated differently across depth. To investigate this question, using a simulated 3D environment, we implemented an orientation discrimination task where target stimuli were presented pseudorandomly at either the near or far depth plane and either left or right hemifield. In Experiment 1, the magnitude of the orientation differences were randomly selected from three values on every trial. In Experiment 2, the same three magnitudes of orientation differences were implemented between participants, with only one magnitude for a given participant; thus, participants could anticipate the magnitude of orientation difference in the upcoming trials. The results showed that when participants were unable to predict the configuration of the upcoming target stimuli, there was a null effect of depth (Experiment 1). However, when participants could form an expectation pertaining to the upcoming orientation differences, a far advantage was revealed (Experiment 2). These findings reveal that target expectancy could differentially impact attention distribution in a 3D discrimination task. The findings in this study could provide insights into learning in attentional allocation and the possible involvement of dorsal/ventral visual pathways in processing stimuli across 3D space.

NSERC, Canada Foundation for Innovation

33.488 HOW DO RELATIVE FEATURES GUIDE ATTENTION IN VISUAL SEARCH?

Stefanie Becker[1], Zachary Hamblin-Frohman[2], Koralalage Don Raveen Amarasekera[1]; [1]The University of Queensland, Australia, [2]The University of Toronto, Canada

Research has shown that attention is often biased to the relative feature of the target in visual search (e.g., redder / larger / darker) rather than its absolute feature values, in line with a Relational Account of Attention (Becker, 2010). However, it is currently unknown how tuning to relative features is achieved. If we know the feature value of the target (e.g., orange), can the visual system rapidly assess the dominant feature in the visual scene (e.g., red or yellow) and compute the relative feature of the target prior to selection (e.g., yellower or redder)? Or do we need on-task experience to learn how the target differs from the context? Another important question is how search progresses through the search items when there are multiple distractors. When an orange target is redder than the majority of other items, do we first select the reddest item, then the next-reddest item and so forth, until we find the target? The present study tested these questions in a 36-item search display with multiple differently coloured distractors and variable target and non-target colours. The first fixations on a trial showed that these displays still reliably evoked relational search, even when observers had no knowledge of the context or relative feature of the target. This indicates that information about the relative target feature can be rapidly extracted and guide attention prior to the first eye movement. Moreover, the first five fixations within a trial revealed that we tend to select the most extreme items first (e.g., red), followed by the next-extreme (e.g., red-orange), etc., until the target is found. This shows that attention is first guided by relative features and only hones in on the exact target colour after multiple fixations on relatively more extreme distractors.

Australian Research Council

SUNDAY AFTERNOON POSTERS IN BANYAN BREEZEWAY

Visual Search: Models, strategy, sequential effects, context

Sunday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

36.301 CONTEXTUAL CUEING IN COMPLEX STIMULI: ATTENTIONAL GUIDANCE AND RESPONSE FACILITATION BENEFITS

David A. Tomshe[1], Melissa R. Beck[1]; [1]Louisiana State University

Previous investigations of contextual cueing using complex stimuli revealed limited performance benefits despite strong explicit memory for repeated displays. In the current study, we increased the number of repetitions and modified the task from a simple identification task to a more complex orientation judgment, aligning our approach more closely with traditional contextual cueing paradigms. Participants completed a visual search task across 24 blocks, with half of the displays repeated and half non-repeated in each block. Results revealed a robust contextual cueing effect, with significantly faster response times for repeated displays compared to non-repeated displays. This effect emerged early and persisted throughout the experiment. Nonlinear model analyses of eye movements uncovered two mechanisms contributing to this performance benefit: (1) attentional guidance, where repeated contexts facilitated more efficient deployment of attention towards the target location, evidenced by shorter time to the first fixations on the target area, and (2) response facilitation, characterized by reduced time between the first fixation on the target and the subsequent response in repeated displays. These findings suggest that the nature of the visual search task can play a role in facilitating or impeding the contextual cueing effect. Importantly, the current study demonstrates that complex stimuli can produce robust contextual cueing effects. Future research should leverage eye-tracking methodologies to further elucidate the underlying mechanisms of search efficiency in repeated complex stimuli, focusing on the temporal dynamics of attentional deployment and refinement of search strategies over time.

36.302 IMPROVED TARGET DETECTION (FEWER LBFTS ERRORS) IN THE REPEATED DISPLAYS CONTRIBUTES TO CONTEXTUAL CUEING

Jeunghwan Choi[1] (abcdef0518@naver.com), Jeremy Wolfe[2,3], Ava Mitra[2], Nathan Trinkl[2], Sang Chul Chong[1,4]; [1]Graduate Program in Cognitive Science, Yonsei University, [2]Brigham and Women's Hospital, [3]Harvard Medical School, [4]Department of Psychology, Yonsei University

In visual search, people typically find targets more quickly in repeated displays than in novel ones, an effect known as contextual cueing (Chun & Jiang, 1998). For example, you may find your phone more quickly in your own room than in a stranger's room even if you don't explicitly recall the phone's location. Two mechanisms have been proposed to account for this effect: attentional guidance and response facilitation. Attentional guidance posits that repeated displays facilitate the search, guiding observers to the target so they encounter fewer distractors. Response facilitation holds that the search is unchanged but the speed of target identification is increased. Here, we add a third factor. The accuracy of target identification is improved by repetition. People sometimes fail to detect the target although it is clearly in view (known as a Look But Fail To See (LBFTS) error, Wolfe et al., 2022). Observers can also make LBFTS fixations where they fixate near the target, fail to identify it, and only find it on a later fixation. In the present study, we investigated whether participants make fewer LBFTS fixations and errors in repeated displays than novel ones. We conducted a contextual cueing experiment while recording participants' eye movements. Participants were asked to find a T shaped target among L shaped distractors. Half of the displays were repeatedly presented throughout the entire experiment. We found faster reaction times (RTs) and lower LBFTS fixation and error rates in the repeated condition than in the novel one. Moreover, we obtained the same result in a previous contextual cueing study (Choi & Chong, 2020) when we reanalyzed those data. These results suggest that people can identify the target more accurately in the repeated displays and this more accurate target detection contributes to faster RTs in the contextual cueing paradigm.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(RS-2022-NR070542) and NSF 2146617 & NIH-NEI EY017001.

36.303 CHARACTERIZING THE INDEPENDENT AND JOINT IMPACTS OF PREVIOUS CATEGORY EVIDENCE AND VISUAL AND SEMANTIC SIMILARITY ON VISUAL ATTENTION

Justin Grady[1] (justgrad@gwu.edu), Emma Siritzky[1], Dwight Kravitz[1,2], Stephen Mitroff[1]; [1]The George Washington University, [2]US National Science Foundation (SBE/BCS)

Research studies commonly average over trials and/or participants to assess general cognitive effects. In visual search, for example, comparing average response times from multiple individuals across two trial types (e.g., low vs. high set size) provides key insights; however, it can also leave meaningful variability untapped and unexplored. An in-depth understanding of visual attention and other related cognitive processes can be gained by assessing factors that may produce trial history effects. Critically, such efforts leverage the fact that performance can be impacted by participants' experience in previous trials. Previous studies have found that repeating stimulus features can lead to increased accuracy and quicker response times (e.g., Gaspelin et al., 2019), and recent work from our group (Kramer et al., 2022) demonstrated that performance benefits were sensitive to both the amount and the relative proportion of evidence for one stimulus condition over another in previous trials. The current project expanded these efforts by assessing the impact of three different types of evidence accumulation on subsequent performance in a simple visual attention task. In addition to the Binomial Z value used in the previous study (Kramer et al., 2022), the current work also assessed the simultaneous impact on performance of prior evidence for repeated category-level features as well as the impact of prior exposure to visually and semantically related evidence. A big data set (from an object-sorting task within the Airport Scanner mobile game) was used to provide sufficient power to explore the nature of these factors together. Main effects and interactions for accuracy and response times were found for evidence of category membership, visual similarity, and semantic similarity. Taken together, this work highlights how various factors, even within the context of trial history in a simple task, can interact to meaningfully impact subsequent visual attention.

Army Research Lab Cooperative Agreement #W911NF-24-2-0188

36.304 A SALIENT, EXPECTED TARGET IN AN UNEXPECTED SETTING CAN PRODUCE INATTENTIONAL BLINDNESS

Daniel Ernst[1], 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 visual search experiments, search trials are often blocked by difficulty. This enables observers (Os) to adjust their strategies to an expected level of difficulty. However, in natural search settings, observers may not be able to anticipate how difficult the next search will be; even when the search target remains constant. For instance, search for a cancer on one CT may be easier or harder than previous searches for that same cancer. Search strategies geared towards harder search may be suboptimal when the target is now more salient and this might lead to an increased number of miss errors. To test this, 25 Os searched for a circle among Landolt Cs (gap size of 0.09 degrees) in an initial block of 32 difficult search trials. Here, Os produced 37% miss errors. On trial 33, Os still searched for the same circle, but the gap size of the Landolt Cs was now 0.45 degrees. Normally, this new task would have been easy. However, when surprised by this easy search, Os failed to take advantage of the target's higher salience and still produced 36% misses. For subsequent presentations of these easy search displays after the surprise trial, miss error rates dropped to 5%. The 36% of Os who responded "target-absent", even though a salient target was present, can be said to have experienced a form of inattentional blindness (IB). Notice that, in contrast to traditional IB experiments, the missed IB stimulus here was no gorilla. It was the absolutely task-relevant, unaltered, and expected target of the search. The same target had already been searched for 32 times. Nevertheless, on the first trial when the distractors were unexpectedly changed to make the task easier, observers failed to adapt. We will discuss the relationship of this form of IB to more traditional versions.

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

36.305 DO INFREQUENT TASK-IRRELEVANT VISUAL CUES CAUSE DISTRACTION?

Nadja Jankovic[1], Rachel Yapp[1], Aaron Richardson[1], Vincent Di Lollo[1], Thomas M. Spalek[1]; [1]Simon Fraser University

In a visual search task that involves the detection of an oddball amongst distractors (e.g., one red diamond among three green diamonds), responses (RTs) are faster when the search display is preceded by a brief, task-irrelevant stimulus (e.g., white screen flash) on half the trials ("alerting" effect). In a parallel procedure in audition, one of two tones precedes the search display on every trial. Longer RTs are observed if the preceding tone is infrequent (20%) rather than frequent (80%). This is known as the "deviance distraction" effect. We tested the idea that the violation of an expectation determines whether alerting or distraction effects are obtained. In a first experiment, the task-irrelevant stimulus was a brief white flash on only 20% of the trials. On the remaining 80% the screen remained mid-grey until the search display. Two options are considered. If the white flash violates the expectation established by the 80% grey screens, a deviance distraction effect (i.e., longer RTs) should occur. In contrast, if expectation does not play a role, an alerting effect (i.e., shorter RTs) should occur. We found no RT differences supporting the hypothesis that distraction and alerting effects canceled each other. In a second experiment the invariant grey screen was replaced by a brief black flash on 80% of the trials. From Experiment 1 we expected alerting and distraction effects to cancel each other on white-flash trials but alerting to occur on black-flash trials. No RT differences occurred between white- and black-flash trials, however both RTs were faster than those found in Experiment 1, suggesting alerting effects – but no deviance distraction effects – in both conditions in Experiment 2. This pattern of results prompts a systematic investigation of how probability of preceding task-irrelevant stimuli affects performance on the search task.

Natural Sciences and Engineering Research Council of Canada

36.306 EXPLICIT STRATEGY INSTRUCTIONS BOOST VISUAL SEARCH OPTIMALITY, BUT THE BENEFITS ARE SHORT-LIVED

Mackenzie J. Siesel[1], Tianyu Zhang[1], Yin-ting Lin[1], Andrew B. Leber[1]; [1]The Ohio State University

People choose among multiple strategies to tackle everyday visual search challenges. For example, when looking for their shiny red car in the parking lot, the most efficient strategy is to focus on its conspicuous color. However, people can also choose suboptimal strategies, like serially searching every car. Recent work has begun to reveal such suboptimal strategy usage, which may lead one to question: can strategies be improved? We have shown that explicitly informing participants of the optimal strategy boosts optimality (Zhang et al., 2024). But, how durable is the improvement? We compared optimality rates between one group of participants that was explicitly informed of the optimal strategy vs. a group that was not (N=60 per group). We found a significant decline in optimality for the instructed group over time, while there was a numerical but not significant increase over time for the non-instructed group. We then compared the linear slopes of the change in optimality over time between the two groups, revealing a significant difference. This confirmed distinct patterns of change in search strategies over time. Moreover, using a sliding window analysis, we found little evidence for rapid boosts in optimality attributable to sudden insight of the optimal strategy (c.f., see Lin & Leber, in press). Overall, explicit strategy instruction does enhance visual search performance, but this improvement appears to be short-lived. We speculate that the drive to avoid cognitive effort associated with optimal performance may eventually win out, irrespective of optimal strategy awareness. The short-lived nature of the instruction effect may explain strong test-retest reliability in optimality (Irons & Leber, 2018), as visual search strategy optimization may be largely governed by stable traits. In summary, these results shed light on how explicit knowledge and trait variables interact to drive visual search strategy use.

This work is supported by NSF BCS-2021038 to ABL.

36.307 NEGATIVE ATTENTIONAL TEMPLATES DEPEND ON TOP-DOWN CONTROL, NOT SELECTION HISTORY

Haley Bennett[1], Nancy Carlisle[1]; [1]Lehigh University

Prior research on negative templates shows participants can use cues about upcoming distractors feature to ignore those distractors. This is demonstrated by faster reaction times when participants are provided an informative distractor cue versus an uninformative cue, leading to a negative cue benefit. However, it is unclear whether these benefits are from active top down control versus learning through experience during the task because all negative cue studies have explicitly instructed participants to ignore the negative cue. In this experiment, all participants performed the exact same task. For the Explicit Ignore group, we told them to use the negative cues to actively ignore, which should rely on top-down control. In the Memory-Focus group we told them to maintain the cue information for a working memory test, so any negative cue benefits would be driven by selection history. All participants received negative or neutral cues in two separate blocks, allowing us to compute negative cue benefits. In addition, we manipulated task difficulty by changing the similarity between targets and distractors. Prior research has shown a negative cue benefit during hard tasks, but not easy tasks. If negative template benefits depend on top-down control, we would expect to see benefits in the hard condition only for participants who were explicitly instructed to use the cue to ignore. However, if negative template benefits are based on selection history, we would expect to see benefits of the negative cue in the hard condition for both explicit instruction and memory instructions. Our results showed negative cue benefits for reaction time and accuracy in the hard condition only for participants who were explicitly instructed to use the cue to ignore. These results support the top-down control account of negative templates.

36.308 PATCHES HALF-EMPTY: HOW TO FORAGE WHEN SOME PATCHES CONTAIN ONLY DISTRACTORS

Injae Hong[1,2] (ihong1@bwh.harvard.edu), Jeremy Wolfe[1,2]; [1]Brigham & Women's Hospital, [2]Harvard Medical School

In visual foraging, decisions on when to leave a patch have been studied using the Marginal Value Theorem (MVT, Charnov, 1976), which suggests leaving a patch when its instantaneous rate of return drops below the environment's average rate. While lab studies often assume rewards in every patch, real-world scenarios include target-absent patches, such as apple trees without ripe apples but with unripe apples only. This study investigated how the presence of empty patches influenced patch-leaving decisions in target-present patches. The task was to collect as many good targets as possible in 10 minutes, moving from one visual patch to another at will. Travel time between patches was either 3 or 6 seconds, during which no targets could have been collected. Three conditions were tested: 'Always-Present' condition - good targets were present in every patch; 'Half-Present' condition - good targets were present in half the patches; remaining patches contained only bad targets; 'Extended-travel' condition - good targets were present in every patch, but the travel time was increased to match the average time lost in the Half-Present condition. The result shows that observers did not reject bad patches immediately and picked some bad targets, presumably to confirm that the patch was empty. Every pick in an empty patch, by definition, had a rate of return below the overall rate. Ideally, observers should have moved on from those patches before picking anything. In apparent compensation, observers tended to leave good patches earlier, while the rate of return was still high. This "underharvesting" seems to violate the MVT prediction but the overall result was broadly in line with MVT. The result seems to reflect complex, (probably implicit) reasoning about balancing behavior in bad and good patches in the Half-Present situation.

NSF-2146617

36.309 TO CHOOSE OR NOT TO CHOOSE: VOLUNTARY TASK SWITCHING WITHOUT COST IN VISUAL SEARCH

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

In visual search tasks in the lab, participants are typically required to perform blocks of the same type of search tasks repeatedly (e.g., find a T among Ls). In real-world searches, however, you look for your keys, then your jacket, then the doorknob, and so on. You rarely search for your keys 100 times in a row. Real-world searches can offer a degree of choice that is not typically a feature of laboratory tasks (i.e., what do I want to look for next?). For instance, given a worklist of cases, should a radiologist be allowed to determine the order in which they are read? The current study aimed to investigate whether search performance, especially reaction time (RT) and miss rate, would be affected by manipulations of trial ordering and participants' choice. Fifty observers completed 100 trials of each of four search tasks: T among Ls, a shape search for bumpy targets among smoother distractors, a colorXcolor conjunction search, and a search for any animal among other objects. Each participant was randomly assigned to one of five conditions: four fixed blocks of 100 trials, blocks whose order could be chosen, a random mixture of all four tasks, free choice of which task came next, and a 'yoked' condition where trials were presented in the order that someone else had chosen. Interestingly, when given the choice, participants rarely switched between tasks, choosing to run blocks of the same trials. Choice made no significant difference to the RTs across conditions, though, unsurprisingly, tasks differed in difficulty. Similarly, there were only very minor differences in errors between choice conditions. While task-switch imposes a cost in other situations, it does not appear to interfere with the performance of these visual search tasks.

NEI EY017001

36.310 TWO SEPARATE CATEGORY LEARNING SYSTEMS GOVERN CATEGORICAL SEARCH

Joseph Schmidt[1], Steven Ford[1], Ashley Ercolino[1], Mark Neider[1], Corey Bohil[2]; [1]University of Central Florida, [2]Lawrence Technological University

Over a decade of research demonstrates that object categories can guide visual search. However, a prominent category learning theory suggests that two systems govern categorization (e.g., Ashby et al., 1998, 2002). The explicit learning system relies on working-memory (WM) and easily verbalizable rules (Rule-Based; RB), whereas the implicit learning system relies on associative learning and difficult to verbalize rules (Information-Integration; II). We tested these systems' contributions to categorical search by training participants to categorize sinusoidal gratings best characterized by RB or II learning, followed by categorical search which included a concurrent spatial working memory task on half the trials (i.e., search between memory and probe arrays). We found search for II categories relative to RB categories resulted in stronger search guidance as measured by the proportion of trials in which the target was the first object fixated, replicating our prior work (Bohil, et al., 2023). However, guidance was largely unaffected by the addition of the WM task. Taken together with the recent success of computational models of search which utilize inverse reinforcement learning (e.g., Yang et al., 2020; Zelinsky et al., 2021), this suggests search guidance may be governed by implicit systems. Importantly, object categorization times (dwell times on fixated targets and distractors) demonstrated clear effects of separate learning systems. RB classification slowed considerably when paired with the WM task but II classification was largely unaffected by the WM task. Collectively, this suggests that visual search guidance may be governed by implicit systems, whereas, object categorization times show clear impacts of the separate learning systems, with explicit RB categories being strongly affected by concurrent WM demands, whereas implicit II categories remain largely unaffected by a concurrent WM task.

This research was supported by the National Eye Institute of the National Institutes of Health under Award Number R15EY029511.

36.311 WHAT MAKES A SEARCH EFFORTFUL? EVIDENCE FROM A SEARCH DISCOUNTING PARADIGM

Tianyu Zhang[1], Andrew B. Leber[1]; [1]The Ohio State University

In daily life, people often engage in visual search, such as finding a car in a parking lot or a friend in a crowded mall. While search is characterized by low-demand, or even automatic, processes and may not be considered as effortful as working memory (Anderson, 2018), recent studies have shown that search is more effortful than previously thought, with individuals even willing to exert physical effort to reduce search demand (Anderson & Lee, 2023). However, it remains unclear how specific components of search contribute to perceived effort. For example, both set size and target-nontarget similarity can affect performance (Duncan & Humphreys, 1989). But do these factors affect the subjective experience of effort similarly? We investigate these questions using a modified effort discounting paradigm. During the selection phase, participants are offered two T among L search tasks that differ in target-nontarget similarity, and they choose freely which task to perform. Across multiple selections, the set size of the selected tasks were adaptively adjusted to reach a choice indifference point between the two. We quantified participants' preference between easy and difficult conditions by the ultimate difference in set size. A small difference means a preference for searching fewer items, while a large difference means a preference for searching less difficult displays (i.e., lower target-nontarget similarity). Results showed participants reached an indifference point when the easy set size exceeded the hard set size by approximately 20 items, despite performance remaining significantly faster in the easy condition. Thus, the choice reflects a relative discounting of search difficulty and preference for smaller set sizes. These findings show that not all components of visual search are equivalent in their subjective demands, and people's approach to search is not solely to minimize time spent on tasks.

36.312 PARALLEL ACCUMULATORS IN LIP DRIVE BEHAVIORAL CHOICES IN VISUAL SEARCH: EVIDENCE FROM SINGLE-NEURON RECORDINGS

Aviad Ozana[1] (ozanaaviad1@ucla.edu), Yelda Alkan[1], James W Bisley[1,2]; [1]Department of Neurobiology, David Geffen School of Medicine at UCLA, Los Angeles, California, USA, [2]Department of Psychology,UCLA, Los Angeles, California, USA

The lateral intraparietal area (LIP) is thought to be a critical component of the neural network involved in spatial attention and eye movements, acting as a priority map. At the neural level, visual search could be implemented in the LIP priority map as parallel neural accumulators that indicate the location of a target when one of the accumulators reaches a threshold. To test this hypothesis, we adopted a covert target present/absent search paradigm using random dot patterns as stimuli (Alkan, Mirpour & Bisley, 2022). This allowed us to look at accumulation patterns of activity during search. We recorded the activity of single neurons in LIP while an animal performed the task. The animal was required to make a saccade towards a target and maintain fixation on target-absent trials. We varied the set size, coherence level, and had 100% valid attentional cues in some trials. Consistent with the parallel accumulator hypothesis, neural activity was correlated with the animal's behavioral responses and task performance. Specifically, neural activity prior to both false alarm (target absent) and hit (target present) trials—where a saccade was eventually made to a stimulus—ramped up to similar peak response – a common threshold. Conversely, prior to miss and correct rejection trials, responses were significantly lower, presumably failing to reach the threshold. The level of activity preceding motion onset decreased with increased set size and was higher when attention had been cued. Lastly, we observed a different pattern of neural activity between correct and error responses. For example, compared to hits, false alarms were either preceded by elevated responses or had transient bursts and fluctuations before reaching threshold. Our findings are consistent with the hypothesis that covert visual search tasks use a parallel and a unified neural process aimed at target detection, in line with the signal detection model.

36.313 PARALLEL ACCUMULATORS IN LIP DRIVE REACTION TIMES IN VISUAL SEARCH: EVIDENCE FROM SINGLE-NEURON RECORDINGS

Yelda Alkan[1], Aviad Ozana[1], James Bisley[1,2,3]; [1]UCLA

In two-alternative forced choice decision making, activity in the lateral intraparietal area (LIP) correlates with the accumulation of evidence to a threshold for a decision that will result in a saccade to a neuron's response field. It has been proposed that this process could explain behavior in visual search by having an accumulator at the location of each stimulus in the array. We tested this hypothesis by having animals perform a visual search task with random dot patterns as stimuli (Alkan, Mirpour & Bisley, 2022). The animal was trained to make a saccade to a target and withhold a saccade if no target was present. We used set-sizes of 1, 2 and 4, and had conditions with 100% valid attentional cues. Stationary dots were presented before motion onset to let the animals know the set-size and attentional condition. Recording from single LIP neurons, we asked whether set-size or attention affected the starting position of the accumulator, the threshold (end point of the accumulator) and the rate of accumulation. We predicted that set-size and attention would not affect the threshold, but that they would affect the starting point of the accumulator and the rate of accumulation. As expected based on previous studies, the starting point of the accumulator varied as a function of set-size (lower for higher set-sizes) and was higher in the attention conditions, and as we predicted, saccades to targets in all conditions were made when activity reached a common threshold. However, we found no evidence that accumulation rates or target detection time differed across any of the conditions. These data are consistent with the hypothesis that visual search may be driven by parallel accumulators and suggest that the differences in reaction time and behavior in all conditions was due only to the different response levels when search began.

Decision Making: Actions

SUNDAY, MAY 18, 2:45 – 6:45 PM, BANYAN BREEZEWAY

36.314 ANALYZING THE SOURCES OF ERROR IN VISUAL SEARCH OF WHOLE SLIDE IMAGES IN PATHOLOGY

Veronica Thai[1], Meng Ling[1], Jeremy Wolfe[2], Zaibo Li[3], Jian Chen[1]; [1]The Ohio State University, [2]Brigham and Women's Hospital, [3]The Ohio State University Wexner Medical Center

Errors are a problem in pathology: false negatives lead to disease not being treated and false positives lead to unnecessary treatment and resulting risk to the patient. We used eye tracking to investigate pathologists' search patterns and behaviors with the goal of understanding the nature of errors in search for cancer in whole slide images (WSIs) of lymph nodes. Ten pathologists of varying experience levels diagnosed and annotated a set of 60 lymph node WSIs; 45 with metastases and 15 benign, while we recorded their gaze and mouse behaviors. Our pathologists had 100% accuracy on the benign slides. There were no false positives in this data set. The false negative error rate ranged from 17.8% to 73.3% (46.1% avg). Based on eye movement scanpaths, we categorized these errors into "search", "recognition", and "decision" errors based on a taxonomy introduced by Kundel et al (1978). The majority (67.5%) of false negatives can be labeled search errors. Search errors are defined as cases where pathologists never fixated on a tumor region. 11.7% were recognition errors where the eyes landed briefly on or near the malignancy without being noted by the pathologist. Decision errors, where a pathologist scrutinized the malignancy but decided it was benign accounted for 20.4%. For all experience groups except residents, longer viewing time was associated with higher accuracy. Curiously, for residents, the reverse was true: longer viewing time led to lower accuracy. Additionally, residents make more use of zooming in comparison to more experienced groups (avg 34.9 vs. 17.2 zooms). They also tended to view at higher magnification (avg 22.64x vs 15.42x). Compared to more experienced pathologists, residents spent more of their viewing time zooming rather than panning (Residents: 22.0% zoom, 10.0% pan; non-residents: 15.2% zoom, 13.4% pan).

36.315 DEEP REINFORCEMENT LEARNING'S STRUGGLE WITH VISUOSPATIAL REASONING: INSIGHTS FROM THE SAME-DIFFERENT TASK

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

Deep learning has dramatically changed the landscape of computational visual systems. One such prominent example is deep reinforcement learning, which is a type of machine learning solution that has seen numerous applications for a wide range of problems spanning game playing to finance, health care, natural language processing and embodied agents. We are interested in embodied agents that are free to visually examine their 3D environment, i.e., are active observers. We will show that deep reinforcement learning struggles to learn the fundamental visuospatial capability that is effortless for humans and birds, rodents and even insects. In order to collect data for our argument, we created a 3D physical version of the classic Same-Different task: Are two stimuli the same? The task was found to be easily solvable by human subjects with high accuracy from the first trial. Using human performance as the baseline, we sought to determine whether reinforcement learning could also solve the task. We have explored several reinforcement learning frameworks, including SAC, PPO, Imitation Learning and Curriculum Learning. Curriculum learning emerged as the only viable approach, but only when the task is simplified significantly to the point that it has only distant relevance to the original human task. Even with curriculum learning, the learned strategies differed significantly from human behaviour. Models exhibited a strong preference for a very limited set of viewpoints, often fixating on the same location repeatedly, lacking the flexibility and efficiency of human visuospatial problem-solving. Conversely, the outcomes of the human experiment were instrumental in developing a curriculum lesson plan that improved learning performance. Our human subjects seemed to develop correct strategies from the first trial and then, over additional trials, became more efficient, not more accurate. Reinforcement learning methods do not seem to have the foundation to match such human abilities.

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.

36.316 SPONTANEOUS SOCIAL COORDINATION THROUGH VISUALLY GROUNDED THEORY-OF-MIND IN **CHILDREN'S AND ADULTS' DECISION**-MAKING

Shaozhe Cheng[1,2] (chengshaozhe@ucla.edu), Kaelin Main[2], Tamar Kushnir[2], Tao Gao[1]; [1]UCLA, [2]Duke University

Humans can perceive rich mental information, such as intentions and goals, from the mere visual movement of simple shapes. While much research has focused on third-party inferences of mental states, little is known about how people use these inferences to make first-person, agentive decisions—especially in social interaction contexts involving conflicting goals. Here, we explore how spontaneous intention inference influences one's own decision-making processes. Participants played a computer game where they freely chose between two equally desirable goals while another agent, controlled by a reinforcement learning (RL) model optimizing its own expected utilities, played the same game on the same visual display. Participants were explicitly informed that they could pursue any goal, regardless of the other agent's actions. We aimed to test whether the mere presence of another agent would prompt participants to infer its intentions and adjust their decisions accordingly. The results revealed that: (a) Both adults and children (ages 5–7) spontaneously used theory-of-mind to infer the intentions of the other agent, as evidenced by their tendency to choose goals opposite to those of the other player, demonstrating "conflict avoidance." (b) This tendency was not universal but depended on the temporal dynamics of intention presentation. Participants avoided conflicting goals when the other agent revealed its intention first or simultaneously but did not yield when they revealed their own intention first. This conditional social avoidance, based on active intention inference, suggests a sense of intention-based ownership. (c) Humans consistently took longer to make decisions compared to the RL agent, highlighting the deliberative nature of human decision-making in social contexts. Together, these findings suggest that visually grounded theory-of-mind significantly impacts individual decision-making in social contexts. Humans spontaneously infer others' intentions, deliberate more, and tend to avoid conflicting with others' goals, revealing the dynamic interplay of vision, intention inference, and social coordination.

36.317 SCENE VARIABILITY SHAPES THE LINK BETWEEN METACOGNITIVE ESTIMATES AND ACTION DECISIONS

Cristina de la Malla[1], Joan López-Moliner[1], David Aguilar-Lleyda[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

Sensorimotor tasks often require us to rely on visual information to evaluate situational risk while considering one's own perceptuo-motor abilities to decide whether and how to act. To investigate the interplay between these evaluations and actions in complex environments, we designed a virtual reality task where participants (N=30) stood in front of a crosswalk. On each trial, two groups of three cars each approached the crosswalk from opposite sides. All cars reached the crosswalk simultaneously, with times-to-contact (TTC) of 3.35, 4.09, 5.00, 6.11 or 7.46 seconds in different trials. Scene variability was manipulated by varying the uniformity of car speeds within each trial: all cars moved at the same, slightly different or clearly different speeds (no, mid and high variability, respectively). In different conditions, participants either (1) assessed perceived risk, (2) evaluated their confidence in crossing successfully and (3) actually crossed the street if they so decided. Both TTC and speed variability influenced responses across conditions: longer TTC and lower speed variability led to higher confidence, lower perceived risk, and more frequent crossings. Across-participant correlations revealed a strong inverse relationship between confidence and perceived risk. Higher confidence and lower risk were also related with more crosses, but only when speed variability was high. This was produced by higher speed variability decreasing confidence and increasing perceived risk more than it reduced crosses, effectively aligning metacognition and action decisions. Conversely, TTC did not produce any coherent modulation of the relationship between perceived risk / confidence and crosses. Overall, our findings suggest that scene variability influences differently metacognitive estimates and action decisions.

This work was supported by grants PID2023-150883NB-I00 and CNS2022-135808 to CM and by grant PID2023-150081NB-I00 to

JLM, funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

36.318 SEQUENTIAL EFFECTS OF UNCONSCIOUS PRIMING IN A 2AFC CATEGORIZATION TASK OF WORDS

Uri Korisky[1], Khen Heller[2], David Whitney[1], Liad Mudrik[2]; [1]UC Berkeley, [2]Tel Aviv University

Many sequential effects in perception and decision making are thought to require that the stimuli are consciously perceived. We present initial findings suggesting that processing the relations between a subliminal word (prime) and a supraliminal word (target) carries onto subsequent trials, affecting the processing of future prime-target pairs. In a reaching task, participants were asked to reach out to the screen and touch the category representing if the target word was artificial or natural. Awareness of the congruent/incongruent subliminal prime was measured at each trial, using both objective and subjective measures. Analysis of reaching trajectory showed that even in trials in which participants were not aware of the prime, reaching movements deviated towards the primed category. A sequential analysis revealed that this effect was larger when prime-target relations changed from the previous trial (e.g., a congruent trial preceded by an incongruent one or vice versa). If true, this suggests that unconscious processing allows integration over longer time windows than commonly assumed, extending across trials and over several seconds.

Decision Making: Models

Sunday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

36.319 CHARACTERIZING THE INTERNAL REPRESENTATION OF CONTRAST THROUGH MAGNITUDE ESTIMATION.

Cristina Rodríguez-Arribas[1], Joan López-Moliner[1], Daniel Linares[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 can be modeled through an internal representation that includes a transducer function mapping physical intensity into perceived intensity and noise. However, characterizing this internal representation remains elusive. Discrimination experiments aim to infer this representation by measuring discrimination sensitivity across a range of intensities. However, since sensitivity reflects the slope of the transducer divided by noise, any internal representation where this ratio matches the observed sensitivity pattern is consistent with the data. In contrast, magnitude estimation experiments could, in principle, resolve this ambiguity. The mean of intensity judgments can represent the transducer, while the standard deviation can quantify the noise. This approach to identify the internal representation, however, remains largely unexplored, as magnitude estimation studies have focused on recovering the transducer, neglecting the role of noise. In our study, eleven participants performed a magnitude estimation experiment for contrast perception at low contrast levels. We fitted the responses of each participant to different models of the hypothetical internal representation, combining various functional forms for the transducer and the noise. We found that the best model for all participants included a sigmoidal transducer and a sigmoidal noise. Assuming that magnitude estimation can capture the internal representation, sensitivity, derived as the slope of transducer divided by noise, should align with sensitivity directly obtained from discrimination experiments. To test this, we measured sensitivity in the same participants using a discrimination task and compared it to sensitivity calculated from magnitude estimation. We found that both methods estimated similar sensitivities both consistent with the pedestal effect, a well-known non-linearity observed in discrimination experiments, but never revealed previously with magnitude estimation. Critically, we found that the location of the pedestal effect estimated from both methods correlated across participants. These findings highlight the potential of magnitude estimation to characterize internal representations.

Supported by grants PID2020-119843RB-I00 and PID2023-151752NB-I00 to DL, PID2023-150081NB-I00 to JLM and CNS2022-135808 to CM funded by MICIU/AEI/10.13039/501100011033 and the EU NextGenerationEU/PRTR. CRA supported by grant PRE2021-099277 funded by MICIU/AEI/10.13039/501100011033 and by the FSE+

36.320 DO TIME-DEPENDENT DECISION BOUNDARIES EXIST? EVIDENCE FROM EMPIRICAL DATA FROM RANDOM-DOT KINEMATOGRAMS (RDK) AND A DRIFT-DIFFUSION MODEL

Dietmar Heinke[1], Casimir Ludwig[2], Jordan Deakin[1,3]; [1]University of Birmingham, UK, [2]University of Bristol, Uk, [3]Universität Hamburg

A dominant model of perceptual decision making assumes observers accumulate noisy evidence to a decision boundary. Theoretical work has shown that under some circumstances, the optimal strategy is to change the decision boundary dynamically while accumulating evidence. The majority of studies suggests that lowering the boundary (collapsing boundaries) over time would be optimal especially when evidence is difficult to obtain, or a decision deadline is involved. However, Malhotra et al. (2018) demonstrated that a certain mix of easy and difficult (but doable) decisions can actually lead to an increase of the decision boundary. This prediction was based on the following rationale: At the start of a trial, the decision boundary might be low and if the decision difficulty is low, it will be crossed rapidly. If the decision boundary is not crossed early, this itself is evidence that decision difficulty is high, and a higher decision boundary should be adopted. However, to our knowledge there is no evidence for such increasing decision boundaries. Here we present evidence from a decision task with RDK-stimuli of four noise levels (10%, 40%, 70%, 80%) and participants had to identify the direction of the coherent motion (left or right). We used Bayesian model fitting of a drift-diffusion model to identify whether and how the decision boundary varies with noise level. When the noise levels were blocked, decision boundary was larger for higher noise levels which might simply reflect "standard" adaptation of constant decision boundaries to decision difficulty. When the difficulties are randomly intermixed, such strategic adaptation is impossible. Nevertheless, we again found that the decision boundaries increased with the noise level. Hence observers may adjust their decision boundaries dynamically within a decision epoch, confirming

Malhotra's et al. prediction. Future studies need to test these findings further and fit formal models with raising decision boundaries.

The first author was supported by a grant from UK-ESRC ES/T002409/1.

36.321 GRAPHICAL PERCEPTION: ALIGNMENT OF VISION-LANGUAGE MODELS TO HUMAN PERFORMANCE

Jenna Kang[1] (jennakang@nyu.edu), Grace Guo[2], Raj Sanjay Shah[3], Hanspeter Pfister[2], Sashank Varma[3]; [1]New York University, [2]Harvard University, [3]Georgia Institute of Technology

Vision-Language Models (VLMs) show promise in chart comprehension tasks that integrate visual and textual information. However, their alignment with human cognitive behaviors in graphical perception is not fully understood. This study evaluates the performance of a VLM, GPT-4o-mini, on seven graphical perception tasks from seminal behavioral studies (Heer & Bostock, 2010; Cleveland & McGill, 1984). The tasks enable assessment of the model's ability to extract and compare numerical values embedded in visualizations across variations in stimulus design, prompt structure, and task difficulty. Results from 315 visualization stimuli reveal that the VLM achieves human-like accuracy in some conditions. The strongest alignment with human judgments ($\rho = 0.90$) was for default stimuli with default stimuli and color cues or explanation-augmented prompts; here, the model matched the task difficulty profiles of humans. The model's sensitivity to visual design elements, such as segment contiguity and color usage, negatively impacted accuracy, even when the underlying numerical information remains unchanged, decreasing model accuracy by up to 41% compared to the default stimuli. The model exhibited improved accuracy when segments were visually distinct, and when prompts included explicit references to color to provide explanations. These findings offer insights into the alignment of VLMs with human graphical perception and suggest the potential of VLMs for the design and evaluation of data visualizations. The observed decline in accuracy under specific stimulus conditions, such as contiguous versus separated segments, raises potential theories for ongoing behavioral studies. These studies aim to explore whether humans exhibit similar biases and assess the feasibility of using Vision-Language Models (VLMs) to simulate human visual processing of graphical data. This research paves the way for further work integrating AI into the design and evaluation of data visualizations, showing the promise of work at the intersection between the vision science and data visualization communities.

36.322 HUMANS VS. LARGE LANGUAGE MODELS: TOWARDS ADDRESSING VISUAL QUALIA

Ziqian Cui[1,2], Weisa Wu[1,2], Shuai Chang[3], Junting Hu[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, China, [3]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, China, [4]St. Paul's School, 325 Pleasant St., Concord, NH, US.

Recent developments of large language models (LLMs) provide a unique opportunity to investigate how subjective visual experience (qualia) may function in human cognition, such as addressing the "Mary's room" thought experiment. The current study examines how humans, influenced by emotional and social experiences, and LLMs, relying on statistical associations, evaluate imagined faces based on semantic facial feature descriptions. 2304 face descriptions were generated by combining ten facial features (gender, brow ridge, eyes, cheekbones, nose, nasal bridge, mouth, chin, skin, and the shape of the eye and mouth corners). Human participants (N=25) were asked to imagine faces based on descriptions and rate them on a 1–9 scale across dimensions, including maturity, emotional valence, gender, physique, trustworthiness, attractiveness, extraversion, leadership, dominance, flexibility, familiarity, and memorability. We also used API calls to five LLMs—Claude 3.5, GPT-4, GPT-3.5, Kimi, and ERNIE—to perform the same task with identical instructions. Representational similarity analysis (RSA) indicated a significant correlation between human ratings and the evaluations provided by the LLMs, with Claude 3.5 showing the highest similarity and GPT-3.5 the lowest. Uniform manifold approximation and projection (UMAP) dimension reduction demonstrated that both humans and LLMs distinguished faces with high and low scores across each dimension. However, significant differences were found in the distances between faces rated as high and low by humans and LLMs. A mixed-effects model examined the impact of facial features on dimensions, revealing that participants focused on eye and mouth corner shape, skin condition, and eye size, while LLMs considered a broader set of features. The results indicate that humans prioritize facial cues associated with emotional expression and social judgments, reflecting cognitive biases shaped by experiences, whereas LLMs, rely on statistical associations, resulting in a broader reliance on general facial features without the emotional and social stability inherent in human judgments.

This work is supported by (1) the National Nature Science Foundation of China (Grant No. 31871136), (2) the Science and Technology Ministry of China (Grant No. 2021ZD0204200), and (3) the Sino-German Center for Research Promotion (Grant No. M-0705).

36.323 LARGE BEHAVIORAL DIFFERENCES AMONG DEEP NEURAL NETWORKS THAT SHARE THE SAME ARCHITECTURE AND TRAINING

Herrick Fung[1] (herrickfung@gatech.edu), N Apurva Ratan Murty[1], Dobromir Rahnev[1]; [1]School of Psychology, Georgia Institute of Technology

Humans vary greatly in many aspects of behavior. Deep neural network models are often seen as promising models of human perceptual decisions but are typically assumed to lack such individual differences. Recent work has revealed substantial differences in the representational geometries among model instances differing only in their random weight initializations (Mehrer et al., 2020). However, whether these representational differences translate into individual differences in behavior remains unknown. Here we examined three datasets of human perceptual decisions that involve tasks of varying complexity: 16-choice object categorization (N=36), 8-choice digit recognition (N=60), and 2-choice Gabor orientation discrimination (N=20). For each task, we trained deep neural networks of several different architectures. For each architecture, we randomly initialized

many network instances that matched the number of human subjects and trained them in identical fashions to reach similar overall accuracy on the validation dataset. We then tested the performance of each model instance on new stimuli on which the model was never trained. We found substantial individual differences in accuracy, confidence, response bias, and reaction time across model instances. Similar to humans, some instances performed reliably better or had consistent response biases compared to other instances. To quantify these individual differences, we computed pairwise image-by-image correlations among model instances across all behavioral metrics. We found that the average pairwise correlations between model instances of different behavioral metrics ranged from r = 0.203 to r = 0.836, demonstrating large individual differences in the pattern of image-by-image responses. Our findings establish the existence of robust individual behavioral differences among different model instances, highlighting the importance of analyzing multiple network instances to draw reliable inferences about network behavior. These results also open a promising avenue for using individual differences in neural networks as a framework to model human individual differences.

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

Spatial Vision: Natural image statistics, texture

Sunday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

36.324 DESIGN AND CALIBRATION OF A SPECTRO-TEMPORAL LIGHT LOGGER

Zachary J Kelly[1], Vincent Lau[1], Geoffrey K Aguirre[1]; [1]University of Pennsylvania

Measurements of visual experience have largely focused upon spatial visual properties (e.g., orientation, spatial frequency) and have made use of static images and cinematic videos. To support studies of individual differences in visual experience, we created a personalized, wearable, all-day, high-speed light-logger. Our goal is to measure the wide-field, low spatial frequency properties of spectral and temporal visual experience across the full range of photopic vision, and to characterize how these natural statistics interact with eye and head movements. The logger is mounted on inexpensive eye glasses (Pair Eyewear) with custom refractive lenses for each participant; 3D-printed spectacle arms incorporate sensor housings. The World camera (IMX219 chip with an M12 lens) provides 120° FOV, RGB video at 200 FPS (rolling shutter; 8 bit, 30x40 Bayer down-sampled image). The Eye camera (PupilLabs Cam2) provides 120 FPS IR video (global shutter, 400x400). Photopic spectral irradiance is derived from 10 narrow wavelength channels at 1Hz (AS7341 chip); mesopic illuminance is measured for a single wide band (TSL2591 chip). 6-DOF head movement is recorded at 10Hz (LSM6DSV16X chip). A Hall effect sensor detects the presence of magnetic clip-on sunglasses. HDMI and USB cables carry data to a 3D-printed (5.4”x3.4”x2.6”) recording pack worn on a cross-body strap, containing a Raspberry Pi 5, M.2 SSD, and rechargeable battery. We performed sensor calibration against a commercial spectrophotometer (PR670) within a light-integrating sphere receiving output from an 8-channel digital spectral synthesizer (plus varying neutral density filters). Using a custom gain control algorithm, the World camera provides nearly flat temporal sensitivity between 0.25 and 50 Hz across a 4 log illuminance range. Validated spectral sensitivity channels are linear across 5 log illuminance. Eye and World recordings are synchronized to within 10 msecs. We are currently refining the device for subject comfort prior to field trials.

R01EY036255

36.325 LINEARIZING SCREEN GAMMA FOR PRECISE PSYCHOPHYSICAL ONLINE STUDIES IN LESS THAN 5 MINUTES

Francis Gingras[1,2] (francis.gingras16@gmail.com), Alex Cousineau[1], Daniel Fiset[1], Caroline Blais[1], Frédéric Gosselin[3]; [1]Université du Québec en Outaouais, [2]Université du Québec à Montréal, [3]Université de Montréal

Online data collection has multiple advantages, including access to larger, more diverse samples as well as fast data collection. However, a particular challenge for vision science studies is that they require visual stimuli to be standardized. We propose measuring perceived brightness at different luminance values to estimate screen Gamma, thus creating a luminance lookup table so researchers can correct their stimuli when running online experiments. Participants adjust the luminance of a uniform square to match the perceived brightness of flanker stimuli, lines alternating between two luminance levels on every pixel across stimulus width. We validated this task, based on an adaptation of the code included in the PsyCalibrator package (Lin et al., 2023), in an online sample of 19 participants recruited through Prolific and tested using VPixx Pack & Go (VPixx Technologies, 2021). Participants completed five 1-parameter Gamma curve measurements (nPoints= 1, 3, 7, 15 and 31 equally spaced luminance levels between 0 and 1, each measured once), twice to assess test/retest reliability. As 0 and 1 necessarily correspond to minimal and maximal luminance respectively, the estimated Gamma curve is fit using nPoints + 2 data points. The best speed/accuracy tradeoff is found using 31 luminance measurements, allowing for a precise estimate in around 3 minutes . Comparing luminance lookup tables, extrapolated to 256 luminance levels, across measurements reveals the absolute value error between measurements is on average .008 (sd = .012). Corrections applied on the 31 measured luminance levels fit a linear model with an average R2 of .999. We find this short and simple task to be a very robust measure of screen Gamma for online participants. Using it will allow vision scientists to account for varying gray level rendering of participant display configurations in their online studies, increasing their control on the presented visual stimuli.

This work was funded by a Canada Research Chair in Cognitive and Social Vision (CRC-2023-00019) as well as a grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-06201), both held by Caroline Blais.

36.326 FROM SIMPLE EDGES TO CONTOURS IN NATURAL SCENES: AUGMENTING THE CONTOUR IMAGE DATABASE

Lynn Schmittwilken[1], Anna, L. Haverkamp[1], Marianne Maertens[1]; [1]Technische Universität Berlin

Edges are 2d image features that are indicative of object boundaries in the natural world. To understand how humans extract edges from the visual input, many psychophysical studies characterize edge sensitivity in well-controlled tasks and with parametrized stimuli. While this approach allows to test specific hypotheses about the underlying mechanisms, it is not clear how well these insights translate to contour perception in natural scenes. Here we test how contour perception in natural images relates to edge sensitivity when both types of stimuli are perturbed with different types of noise. We compare edge sensitivity data from a 2-AFC task with simple stimuli, and line drawings from a contour tracing task in images. In both cases, we perturb the stimuli with 2d noise to probe the putative underlying spatial frequency (SF) selective mechanisms. We used three broadband noises (white, pink, brown) and three narrowband noises (center SFs: 0.5, 3, 9 cpd). In the 2-AFC task, observers indicated the location of a Cornsweet edge via a button press. We used three Cornsweet edges with different peak SFs (0.5, 3, 9 cpd). In the contour tracing task, subjects drew all visible contours in a variety of natural scenes from the Contour Image Database (Grigorescu et al., 2003) using a drawing tablet. In both cases, we measured psychometric functions varying edge contrast. In the drawing task, we quantified performance as the amount of agreement between the contour traces in the presence and absence of noise. Finally, we compared the corresponding noise masking effects in both tasks. Our results provide an example that early visual processes can be investigated in more behaviorally-relevant and meaningful tasks. Furthermore, our data serves as a useful augmentation of the Contour Image Database to study the mechanisms underlying human edge sensitivity via computational modeling.

36.327 MODELING THE STATISTICAL PROPERTIES OF NATURAL IMAGES AND MEDICAL RADIOGRAPHS

Vicky Cai[1], Jonathan Victor[2]; [1]Stuyvesant High School, New York, USA, [2]Weill Cornell Medical College, New York, USA

Typical of sensory systems, many aspects of human vision are tuned to the characteristics of its natural inputs. These characteristics include global statistics (power spectrum) and local statistics (multipoint correlations). While the basis for the global statistics of natural scenes is largely understood, the origin of the distinctive pattern of local image statistics is not. Medical radiographs constitute a useful contrast: they have a steeper spectral slope (-2.6 to -3, vs. -2 for natural images), a different pattern of local statistics, and analysis of radiographs has important consequences. Moreover, natural images and radiographs are formed by different physical processes: occlusion is typical in natural images; transparency is typical in radiographs. To address the origin of the statistical properties of these image classes, we built a series of generative models. Starting with the “dead leaves model” for natural images (Ruderman 1997), we replaced occlusion by transparency, generalized it to 3D, and considered several object size distributions. We then calculated global and local image statistics for each model across a range of scales. As expected, (Metheany et al., 2008), the shift from occlusion to transparency accounted for spectral slopes. For local image statistics, we considered the means of multipoint correlations (overall characteristics of each image class) and their standard deviations (characteristics that distinguish images within a class). All models accounted for the prominence of pairwise statistics, compared to three-point and four-point statistics, seen in natural and medical images. Occlusion vs. transparency and size distribution had a large effect on the mean values of multipoint correlations. They had a smaller effect on their standard deviations and did not account for the differences between three-point and four-point statistics seen in natural images and radiographs. Overall, simple generative models can explain some, but not all, of the statistical characteristics of natural images and radiographs.

NIH EY07977, Fred Plum Fellowship in Systems Neurology and Neuroscience

36.328 TASK-DEPENDENT MODULATION OF SIMILARITY JUDGMENTS

Jonathan Victor[1], Mary Conte[1]; [1]Weill Cornell Medical College, New York, USA

Representations of visual information can often be described by perceptual spaces: constructs in which stimuli or features correspond to points, and their pairwise distances correspond to perceived dissimilarity. This information can be used in many ways – for example, discrimination, similarity, or grouping -- and correspondingly, a perceptual space’s geometry can change in a task-dependent fashion. The domain of visual textures is a model for probing these transformations. Using synthetic black-and-white textures (Victor & Conte 2015) and the paradigm of Waraich & Victor (2022) we recently showed that the transformation from the representational space for threshold judgments to suprathreshold similarity was characterized by selective stretching along some axes and curvature-producing nonlinear distortions (VSS 2024). Here, we examined a specific judgment, brightness comparison. On each of 1000 trials, subjects (N=5) ranked eight texture patches in order of similarity to a central reference texture. These judgments were used to infer a perceptual space via multidimensional scaling. The texture stimuli varied along many dimensions, including mean luminance and low- and high-order spatial correlations. In one experiment, subjects could judge similarity using any criterion. No further instructions were given. In a second experiment using the same stimuli, subjects were asked to judge similarity solely on the basis of the brightness of the central reference. Two novice subjects, unfamiliar with the texture stimuli and paradigm, carried out the brightness task first. Similar results were obtained from both the practiced and novice subjects. We expected that brightness judgments would yield a one-dimensional representational space, corresponding to luminance. Instead, the representational space remained high-dimensional. Compared to the representational space for suprathreshold similarity, the axis corresponding to luminance was elongated; other axes were selectively compressed; and curvature was less prominent. Thus, the perceptual space of similarity judgments can be selectively expanded or contracted by task.

NIH EY07977, Fred Plum Fellowship in Systems Neurology and Neuroscience

36.329 DETECTION OF ARTIFACTS IN CLEAN AND CORRUPTED VIDEO PAIRS IS INFLUENCED BY ARTIFACT TYPE AND PRESENTATION MODALITY

Niall L. Williams[1] (niallw@umd.edu), Anatolii Evdokimov[2], Budmonde Duinkharjav[1], Anjul Patney[3], Qi Sun[1], Jae-Hyun Jung[3], Ruth Rosenholtz[3]; [1]New York University, [2]University of Richmond, [3]Nvidia

Modern computer-generated videos display a variety of artifacts. While image-computable metrics exist to quantify the visibility of artifacts in images and videos, designers often rely in part on human observers to find artifacts and assess video quality. Furthermore, human labeling of artifacts is often an essential component of building image and video quality metrics. Yet, relatively little research has studied the impact of different video comparison interfaces on an observer's strategies and ability to detect different artifact types. Different presentation modalities may require higher memory load or may make differences more visible, e.g. presenting videos side-by-side forces the viewer to saccade between matching regions of the video to do comparisons, showing one video at a time and allowing the user to toggle between them emphasizes small differences, and a split-screen view of both videos with a movable seam between the two videos affords precise inspection of specific regions of the video. Here we study how artifact search performance and behavior changes as a function of the video playback interface and the types of artifacts. Five participants identified and labeled the locations of artifacts (ghosting, compression, or added noise) in pairs of videos, with and without artifacts, using one of three different interfaces: side-by-side simultaneous viewing, temporal toggling between videos, or split-screen simultaneous viewing with a movable sliding seam. Results showed that observers correctly located ~25% of all corrupted pixels regardless of the viewing interface, but artifact type had a significant effect on detection rate: ghosting artifacts were harder to detect than compression and noise. The side-by-side viewing condition caused viewers to scan a larger percentage of the display while the split-screen condition produced more mouse movements (although cursor position distribution remained consistent across viewing conditions). Finally, task completion time was not significantly different across presentation modalities, except for one participant.

36.330 UNCERTAINTY IN VISUAL DETECTION

Zahra Hussain[1] (zahra.hussain@plymouth.ac.uk), Chinaecherem Nwigwe, Patrick Bennett; [1]University of Plymouth, [2]University of Plymouth, [3]McMaster University

We compared the effects of signal uncertainty in a detection task for two signals: simple sinusoidal gratings, and two-dimensional noise textures. This comparison was prompted by our observation that textures elicit a potentially different strategy for detection than gratings. Observers performed a two-interval forced-choice task in which signal type (gratings vs. textures) and signal uncertainty (one vs. five signals) varied across blocks. The gratings were 1 cycle/degree sinusoids at five orientations (0, 72, 144, 216, 288 deg), and the textures were five band-limited noise patterns (1-2 cycles/degree). Signals were embedded in Gaussian noise that was either fixed to a single sample for both intervals and all trials in a block (fixed noise; five observers), or varied across intervals and trials (variable noise; three observers). In the no-uncertainty condition, a single orientation or texture was used throughout a block of trials. In the uncertainty condition, the signal was sampled randomly from one of five variations on each trial. Observers performed twelve blocks per session (2 signals x 2 uncertainty conditions x 3 repetitions per condition; 50 trials per block), and contrast thresholds were measured in each block using QUEST+. For gratings, contrast thresholds were higher under uncertainty, and this effect was larger in variable than fixed noise. Interestingly, no uncertainty effect was was found for textures in either noise condition. This result supports the idea that signal uncertainty has distinct effects on the detection of spatially complex and spatially predictable patterns. We currently are investigating how these uncertainty effects vary across a larger group of naive and experienced observers.

Face and Body Perception: Body

Sunday, May 18, 2:45 – 6:45 pm, Banyan Breezeway

36.331 AN ASIAN-BASED DATABASE OF EMOTIONAL BODY MOTION INDUCED BY DIVERSE DAILY SITUATIONS

Chiahuei Tseng[1], Miao Cheng, Ken Fujiwara, Yoshifumi Kitamura; [1]Research Institute of Electrical Communication, Tohoku University, Japan, [2]Interdisciplinary ICT Research Center for Cyber and Real Spaces, Tohoku University, Japan, [3]Department of Psychology, National Chung Cheng University, Taiwan

Emotion understanding in body movements is crucial for non-verbal communication, yet our progress is far behind our knowledge of facial expressions. One reason is the lack of a robust database that is representative enough to cover the complexity of emotional expression in the real world. Most current databases for bodily emotion expressions are culturally skewed to Western countries, and many are simple and repetitive actions (e.g. walking, knocking) focusing on parts of the body only. We expanded the current repertoire of human bodily emotions by recruiting Asian professional performers varied in personification information (e.g. gender, age, performing experience) to wear whole-body suits with 57 retro-reflective markers attached to major joints and body segments and express 7 basic emotions (joy, sadness, anger, fear, contempt, disgust, and surprise) and 5 social emotions (gratitude, jealousy, pride, shame, guilt) with whole-body movements in a motion capture lab. For each emotion, actors performed three self-created scenarios that cover a broad range of real-life events to elicit the target emotion within 2-5 seconds in 3 different intensities (low, medium, and high). No actor saw any others' performances, ensuring that each actor's performance remained uninfluenced by others, thereby allowing for a wide range of interpretations and patterns of movements. After all the motion capture sessions were completed for each performer, we interviewed them to elaborate on their performing scenarios and acting strategies. We evaluated this database with human evaluation and found (1) that the emotion discrimination accuracy was comparable to other Western databases containing standardized performance scenarios; (2) own-race advantage in emotional recognition accuracy between Asian and non-Asian participants; (3) accurate situational information enhances body emotional understanding. The results suggest that a database using a novel emotional induction approach based on personalized scenarios will contribute to a more comprehensive understanding of emotional expression across diverse contexts.

36.332 TWO PREVIOUSLY UNDESCRIBED CORTICAL REGIONS SELECTIVE FOR BODIES IN THE VISUAL PERIPHERY

Yu Zhao[1] (ruose1314@gmail.com), Matthew W. Shinkle[1], Mark D. Lescroart[1]; [1]University of Nevada, Reno

Recent neuroimaging work has used sensitive scans and analyses to describe several small, previously overlooked regions that respond to scenes, faces, and words. This suggests that there may be elements of the representation of these or other categories that known regions do not capture. Inspired by this work, we investigate whether there may be additional body-selective regions by analyzing a large-scale 7T fMRI dataset, the Natural Scenes Dataset. We first analyzed the high-field category localizer data in NSD. Contrasts of bodies and limbs versus objects revealed small and as-yet undescribed patches of selectivity in consistent anatomical locations across subjects: one located posterior to the parahippocampal place area (PPA) and another superior to the retrosplenial complex (RSC) along the parieto-occipital sulcus. We provisionally term these areas the Ventromedial Body Area (VMBA) and the Medial Body Area (MBA), respectively. Next, to validate selectivity and characterize tuning in these regions in naturalistic contexts, we analyzed responses to the NSD main experiment. We developed a voxelwise encoding model with features capturing body locations and counts in each stimulus image. This model accurately predicted responses in both VMBA and MBA and in established body-selective areas EBA and FBA. Variance partitioning against models of low-level visual features (Gabor filters) and scene category labels showed that body-related features explained unique variance in VMBA and MBA responses. Finally, analyses of encoding model weights and highly stimulating NSD images suggested that, while EBA and FBA are selective for centrally-presented bodies, VMBA and MBA are selective for multiple bodies in the peripheral visual field. Together, our results reveal two novel cortical regions which play a complementary role to established body-selective areas. These new regions may provide a basis for understanding bodies or actions in their spatial contexts and/or for directing attention to bodies in the visual periphery.

Supported by NSF CAREER # 2340895 to M.D.L.

36.333 ARE THEY ONLY ACTING? PERFORMING AND OBSERVING PANTOMIMED ACTIONS

Sholei Croom[1], Chaz Firestone[1]; [1]Johns Hopkins University

Ordinary observers can infer the goals of others' actions, such as where someone is walking or which object they are reaching for. However, actions are more than their goals; underlying visually-guided behavior are complex dynamics between an agent's body and the environment. What do observers know about such dynamics and the behaviors that emerge from them? Here, we explore this question through "pantomimed actions", in which people perform actions with imaginary objects. Pantomimed actions differ kinematically from genuine object-directed actions because the absence of visual information disrupts the typical perception-action feedback cycle. If observers can distinguish real actions from pantomimed actions, this would reveal finer-grained intuitions about the dynamics underlying visually-guided action. We created a set of videos in which agents performed object-directed actions involving no physical contact with the objects (e.g., stepping over a box, ducking under an overhang, or weaving between poles). In half of the videos, actors interacted with real boxes, real overhangs, etc.; in the other half, they were instructed to move *as if* interacting with these objects (in both cases, a censor box occluded where the object was or would have been). Then, independent subjects watched these videos and had to determine which was which; which videos showed real actions and which showed pantomimes? Collapsing across all actions, observers discriminated real actions from pantomimes at rates above chance. However, certain actions were more easily discriminable than others; for some actions, actors successfully 'fooled' observers into thinking an object was present. Our work supports two conclusions: (1) Observers are sensitive to kinematic differences distinguishing genuine visually-guided actions from their pantomimed counterparts (revealing surprisingly fine-grained intuitions about visuomotor processing); (2) The ability to "fake" actions may be more robust than previously suggested (e.g., findings that actors cannot make a light box seem heavy).

NSF BCS #2021053

36.334 INTERPERSONAL DISTANCE MODULATES THE DETECTION OF SOCIAL GROUPS

Luowei Yan[1] (luowei.yan@mail.mcgill.ca), Clara Colombatto[2], Jelena Ristic[1]; [1]McGill University, [2]University of Waterloo

Human life is inherently social, with people frequently assembling in groups of two or three individuals. Accordingly, human perception is attuned to detect social groups; for example, observers are faster at detecting individuals facing towards versus away from each other, an effect known as the 'facing advantage'. However, group configuration is not defined solely by the orientation of its members, but also by the distances between them. The study of proxemics identifies four interpersonal distance zones — intimate, personal, social, and public — with only the intimate and personal distances supporting close social interactions. Here, in a preregistered experiment, we investigated the role of interpersonal distance in the perception of social groups. Participants viewed displays containing four dyads or triads and searched for facing groups (among non-facing distractors) or non-facing groups (among facing distractors). Group members were positioned at either personal or public distances. Participants were faster to find dyads than triads, and groups with members positioned at closer interpersonal distances compared to farther ones (personal vs. public). Most importantly, while overall facing groups were found faster than non-facing groups, this facing advantage also varied with interpersonal distance such that for groups of two the facing advantage was more pronounced at the personal relative to public distance, while for groups of three the facing advantage did not vary across distances. This suggests that distance between group members may play a greater role in the perception of smaller dyads, with distances conducive to close social interactions enhancing perceptual grouping. Social perception thus appears to be attuned not only to the overall structure of groups, but also to their intrinsic interactive properties, highlighting a fundamental capacity of the human perceptual system to incorporate diverse environmental cues into the perception of complex social structures.

This project was supported by funding from the Natural Sciences and Engineering Council of Canada (NSERC), Social Sciences and

Humanities Council of Canada (SSHRC), G. W Stairs, and William Dawson Funds.

36.335 DEVELOPMENT AND VALIDATION OF A NEW DYNAMIC BODY EXPRESSION DATABASE

Lilly Kelemen[1] (lilly.kelemen@nih.gov), J. Brendan Ritchie[1], Chris Baker[1], Shruti Japee[1]; [1]National Institute of Mental Health

Facial expressions are a crucial, nonverbal means of social communication. While much research has focused on the role of facial expressions in social interactions, the role of body expressions has been less studied. Further, studies examining the role of body cues often rely on static images of actors posing various body expressions with their faces digitally obscured or cropped. To address this gap, we developed the Facial and Body Movement Database (FaBMoD), which comprises 352 videos of nine (5 male) racially diverse professional actors making a variety of body movements. The actors were instructed to portray six basic emotional movements and four neutral movements. To understand the relative role of face and body expressions in social communication, we manipulated the visibility of facial information in two ways. First, each movement was performed in two orientations—once facing the camera and once facing away from the camera such that only a limited portion of the face was visible. Second, each movement was performed with and without masking using a green screen head covering. To validate the stimulus set, 10 participants completed a computer task where they were shown each video and asked to identify the expression portrayed and rate each video on a scale of 1-9 for intensity, genuineness, and valence. Preliminary results indicate good overall correspondence for most videos between the intended expression and the expression perceived by participants. This was especially the case for videos depicting anger and happiness, while videos intended to depict surprise were sometimes perceived as fear. Neutral movements were classified as neutral in almost every case. This new database of body movements, along with the validation data, will be used to understand the role of body expressions in social communication. We also intend to share this database with other researchers world-wide.

36.336 BEYOND THE MU RHYTHM: SENSORIMOTOR REPRESENTATION OF EXPRESSIVE MOVEMENT BY ONESELF VS. OTHERS

Alison M. Harris[1], Chandlyr M. Denaro[1], Eric J. Moody[2], Catherine L. Reed[1]; [1]Claremont McKenna College, [2]University of Wyoming

The human ability to rapidly infer others' mental states from their body movements has been hypothesized to depend on motoric representations, whereby observed actions are internally simulated within our own sensorimotor circuits. Supporting this idea, studies using electroencephalography (EEG) have reported reductions in the mu (8-14 Hz) and beta (15-20 Hz) rhythms over sensorimotor cortex both when executing one's own movements and when observing the actions of others. Yet, this idea suggests that actions closer to one's own motor repertoire should be easier to simulate, particularly for more idiosyncratic expressive movements that convey emotional information. To test this prediction, we measured behavioral performance and brain activity using high-density 128-channel EEG in participants (N =56) while they viewed point-light display (PLD) videos of their own movements vs. those of unfamiliar controls. Every participant first completed a motion capture session modeling three examples each of expressive movements (happy, angry, sad) and non-expressive neutral movements (marching, sidestep, trunk twist). Individual PLDs were then created using 3D animation software, removing any visual form cues that could be used to distinguish identity (e.g., height, body shape). In a subsequent testing session, participants were able to behaviorally discriminate PLDs of their own vs. others' movements, but only for expressive movements (d' >1). However, EEG analysis found no significant differences in the central mu rhythm for observation of expressive movements as a function of self vs. other. Instead, significant reductions were seen at frontocentral sensors in the beta band (15-17 Hz). These results support the idea that observers are differentially sensitive to their own motor repertoires, even without other identifying visual information. However, this self-other discrimination appears to occur beyond previously-reported central sites and alpha-band frequencies, suggesting the involvement of broader networks in the analysis of motor representations during action observation.

This research was supported by NSF BCS #1923178

36.337 DETECT AI-GENERATED FACES IN A GLANCE? EFFECT OF STIMULUS DURATIONS ON DEEPFAKE DETECTION AND EYE MOVEMENTS

Zihao ZHAO[1] (zihao005@e.ntu.edu.sg), Mingzhi LYU[2], Adams Wai Kin KONG[2], Hong XU[1]; [1]School of Social Sciences, Nanyang Technological University, [2]College of Computing and Data Science, Nanyang Technological University

Artificial intelligence (AI) advancements have led to hyper-realistic facial images. Recent studies found that humans could not tell deepfakes from real faces. However, whether increasing viewing duration can improve deepfake detection is unclear. This study investigated the effects of viewing durations and deepfake algorithms on detection accuracy and eye movement patterns. Thirty-eight participants viewed 112 facial images (6 types: Stable Diffusion (SD 1.5 & SD XL 1.0) generated, Style Generative Adversarial Network (StyleGAN) generated, face-swapped images, and face images from Deepfake Detection Challenge (DFDC), Karolinska Directed Emotional Faces (KDEF), and Tsinghua Facial Expression datasets), judging each as fake or real. Images were presented at 6 viewing durations (16.67, 33.33, 50, 100, 500, 1000 ms) across 6 blocks in random order, with eye movements recorded by EyeLink 1000 Plus. One participant's data was excluded from analysis for extremely low accuracy. Generalized Linear Mixed Models showed that image type and viewing duration significantly influenced detection accuracy (p's < 0.001 for main and interaction effects). SD-generated and real images were easily detected (accuracy > 83%), even at 16.67 ms, while StyleGAN and face-swapped images were difficult to detect (accuracy ~7% and 20%, respectively). For face-swapped images, accuracy improved with longer durations, suggesting extended exposure aids in identifying subtle artifacts. StyleGAN images did not improve with longer viewing times; accuracy even decreased. Eye movement analysis showed that real and face-swapped images from DFDC elicited fewer fixations and saccades with viewing durations of 500 and 1000 ms, whereas SD-generated images elicited more saccades and higher saccade amplitudes. Area of Interest analysis revealed predominant fixation on the nose region across all image types. Our results found that deepfake algorithms and viewing duration significantly affect deepfake detection and eye movement patterns. It sheds light on our understanding of face perception mechanisms in the age of AI.

36.338 LEVERAGING SUBJECTIVE RATINGS FOR DEEPFAKE FACE CLASSIFICATION USING MACHINE LEARNING

Lillie C. del Real[1] (lillie.delreal@ucf.edu), Jessica N. Goetz[1], Mark B. Neider[1]; [1]University of Central Florida

Several studies have demonstrated that human deepfake classification accuracy is often near chance levels and synthetic faces are rated as more trustworthy than real faces (e.g., Nightingale & Farid, 2022). These findings may suggest the potential for subjective judgments like trustworthiness and related attributes, like attractiveness and symmetry, to influence classification accuracy of real and synthetic faces (Bascandziev & Harris, 2014; Perrett et al., 1999). In the current study (N = 418) participants rated (e.g., trustworthiness, attractiveness, symmetry, facial expression, racial ambiguity, threat, competency, masculinity/femininity, likeableness, and unusualness) and classified 160 high-quality face images (80 real; 80 synthetic generated using NVIDIA's StyleGAN2 (Karras et al., 2020)) as real or synthetic to examine whether a participant's perceptual experience of a face contributes to how they classify that face. Overall classification accuracy was slightly above chance performance (~52%) ($p < .001$) with synthetic faces rated more trustworthy, attractive, and symmetrical than real faces (all ps < .001). Two binary logistic regressions of the subjective ratings as predictors of classification accuracy of synthetic faces and real faces revealed aggregate ratings of attractiveness, symmetry, unusualness, and likeableness as positive predictors of correct classification of synthetic faces and negative predictors of correct classification of real faces (all ps ≤ .05). Additionally, to explore more complex patterns in the data, we trained a support vector machine (SVM) using the image ratings as features. The SVM was able to correctly classify a face as real or synthetic with 65% accuracy (relying heavily on attractiveness, unusualness, and symmetry ratings). Combined, our findings suggest that while humans struggle to classify synthetic faces, their percepts of those faces may hold some predictive power with respect to accurate classification and potential for intervention.

SUNDAY AFTERNOON POSTERS IN PAVILION

Visual Memory: Capacity and encoding of working memory

Sunday, May 18, 2:45 – 6:45 pm, Pavilion

36.401 WORKING MEMORY CAPACITY LIMITS ARE MORE IMPORTANT THAN FEATURE SIMILARITY WHEN REMEMBERING DUAL INFORMATION STREAMS

Chenye Bao[1] (cbvm5@umsystem.edu), Kyle Hardman[1], Jake Sauer[1], Nelson Cowan[1]; [1]University of Missouri

Modern life often requires managing dual information streams, such as driving while following GPS directions. This study examines why dual sources are challenging by investigating visual working memory's capacity and precision limits. Forty-eight adults (mean age = 20.21 years, SD = 4.88) performed recall tasks involving one or two sequentially presented arrays of colors and line orientations. We manipulated inter-array similarity to explore interference effects. Participants reported stimuli features using response rings, enabling assessments of both capacity and precision. Results show consistently poorer performance when recalling two arrays compared to one, regardless of their similarity. Our psychological process modeling revealed that this difficulty arises from reduced item availability in visual working memory rather than precision degradation. Additionally, the absence of significant inter-array similarity effects challenges interference-based theories, suggesting that capacity limitations rather than similarity drive working memory constraints. These findings contribute to understanding visual working memory by highlighting capacity as a fundamental limitation, overriding similarity effects. Future research could investigate how factors like temporal organization or item distinctiveness influence interference and capacity constraints.

The work was funded by NICHD Grant R01-HD021338. We thank Bret Glass for assistance. The program, results, and analyses are available online at https://osf.io/hu83r/?view_only=11c45592c6804e70aa332e635b2dae81

36.402 THE COST-BASED BALANCING OF SAMPLING VERSUS REMEMBERING IN A NATURALISTIC TASK

Candice Koolhaas[1] (candice.koolhaas001@umb.edu), Jade Zack[1], Zsuzsa Kaldy[1], Erik Blaser[1]; [1]UMass Boston

Naturalistic tasks, like a successful grocery run, allow individuals to modulate their use of external resources (a shopping list) versus internal resources (memory); a sampling-remembering trade-off (Ballard et al., 1995; Van der Stigchel, 2020; Liang et al., 2025). In our view, we exploit this trade-off to maintain a preferred balance: as the cost to use internal resources would increase (a longer to-be-remembered shopping list), we compensate by sampling more; as the cost to use external resources would increase (the list is in your partner's pocket), we compensate by remembering more. In our tablet-based 'Shopping Game', participants (N=27) chose items in a store from a 10-item shopping list. The list and the store were not visible simultaneously. In the low-cost baseline condition, participants could toggle freely between them. In the sampling-delay condition, sampling cost was increased by introducing an annoying 5s lag before list exposure. In the maintenance-delay condition, sampling cost was similarly increased, but now the 5s delay occurred before store exposure, meaning, critically, that remembering cost was also increased (since list items had to be maintained in working memory during the lag); i.e., opposing effects. We measured participants' study time of the list and memory usage (number of correct selections, per store visit). As expected, there was a shift toward remembering - greater study time and memory usage - in both delay conditions. Our results also reflected the opposing effects present in the maintenance-delay condition, with study time and memory usage midway between the baseline and sampling-delay conditions (a post-hoc test for linear trend (baseline, maintenance-delay, sampling-delay) was significant for both measures, F(1, 50)=19.8, p<.001 and F(1, 50)=4.95, p=.031, respectively). An ongoing analysis of concurrent task-evoked pupillometry (Tobii Nano) will be used to corroborate this cost-based rebalancing of sampling versus remembering.

The project was supported by NIH R15HD115244 and NDSEG 12084CNBS grants

36.403 TRANSFORMATIONS IN VISUAL WORKING MEMORY IMPOSE FIDELITY COSTS

Eva Lout[1] (evalout@utexas.edu), Jarrod Lewis-Peacock[1]; [1]University of Texas at Austin

In our dynamic world, working memory (WM) plays a critical role in storing, updating, and using goal-relevant information to enable adaptive behavior. Previous research on WM updating has demonstrated that updating slows significantly when a transformation of the item is required before substitution (e.g., “add 4” to the “25” stored in WM vs. “replace 25 with 29”). Transformation processes require the original information to be retained throughout the updating process. Critically, recent EEG findings suggest that this original information may linger in WM even after the transformation is completed. Retaining the original copy may help stabilize the newly transformed representation, but it can also tax the limited capacity of WM storage. Performing multiple transformations may impose an accumulative burden, ultimately degrading the fidelity of the representation. This study aimed to test this hypothesis by varying the number of mental rotations participants performed on the visual orientation of a remembered teardrop shape (zero, one, three, or five rotations of either 30°, 60°, or 90° in either direction). We conducted two experiments with the same task but different memory reports: In Experiment 1, participants reported the updated teardrop shape’s precise orientation, while in Experiment 2, they made a same/different judgment between the remembered orientation and a probe orientation. Our preliminary results revealed that as the number of transformations increased, both memory precision and recognition accuracy progressively declined towards chance levels, while reaction times remained relatively stable. These results suggest that the amount of uncertainty in the reported memory representation increases with each successive transformation. Updating the contents of WM through transformations of existing representations appears to be a lossy process that reduces memory fidelity. Fidelity costs may arise from interference caused by the continued storage of outdated representations or from imprecise transformation processes that degrade sharp representations into blurred ones.

36.404 REWARD SHAPES RESOURCE ALLOCATION IN WORKING MEMORY

Qingqing Yang[1], Hsin-Hung Li[1]; [1]The Ohio State University

Efficient coding theories posit that neural resources are optimized according to environmental statistics, successfully explaining stimulus-dependent variability in perception and working memory (WM). Recent studies further propose that rewards also influence stimulus encoding (Schaffner et al., 2023). However, past experiments provided conflicting evidence on whether trial-wise rewards modulate WM resources (Cho et al., 2022; Van Berg et al., 2023). Here, we test the hypothesis that the learnt association between stimulus features and rewards impact the allocation of WM resources. Across two sessions on separate days, participants learned two reward contexts through a 4AFC decision task (100 trials per day). In each trial, four oriented gratings (sampled from naturalistic distributions with more cardinal orientations) were presented on different quadrants and participants chose one of them to receive immediate reward. In the baseline context session, rewards were uniform across orientations, whereas in the reward-manipulated context session, rewards increased linearly as the orientations became more diagonal. After finishing the decision task, participants performed a delayed-estimation task (300 trials per day). In each trial, participants viewed a central grating (uniformly sampled orientations; set size=1; 4 dva in width, 500 ms duration) and reproduced its orientation after a 1s or 5s delay. The rewards in each trial were based on the memory error, and scaled by the orientation-reward mapping learned during the decision task. Across sessions, behavioral estimations showed higher variability at more diagonal orientations (45°, 135°), consistent with the oblique effect. Critically, we found an interaction between orientations and reward contexts. Orientations with higher diagonality were remembered with greater precision—lower variability and smaller error—when associated with higher reward compared to the baseline context. This effect is stronger with a longer (5s) delay. Our results indicate that participants actively adjusted the allocation of WM resources during maintenance based on the expected utility of features.

36.405 PERCEPTUAL AND CONCEPTUAL CONTRIBUTIONS OF THE REAL-WORLD OBJECT BENEFIT IN VISUAL WORKING MEMORY: IS LOOKING LIKE AN OBJECT GOOD ENOUGH TO ENHANCE MEMORY?

Yong Hoon Chung[1] (yong.hoon.chung.gr@dartmouth.edu), Sam (Seho) Jung[1], Viola Störmer[1]; [1]Dartmouth College

Recent studies demonstrate that visual working memory capacity is greater for real-world objects compared to simple features like colors and scrambled objects. This led to the proposal that conceptual meaning plays a critical role in structuring visual working memory (Chung, Brady, & Störmer, 2024). However, one challenge in comparing memory performance across stimulus sets is that they vary not only in conceptual meaning but also in perceptual similarity. Thus, some of the working memory benefits for real-world objects may arise from these perceptual differences – for example whether the visual system interprets inputs as objects or not. Here, we provide a strong test of this by using novel objects generated by generative adversarial networks designed to resemble real objects (Cooper et al., 2023), and compare memory performance across novel and familiar real-world objects. Across experiments, participants remembered sets of four objects drawn from one of four stimulus types: familiar objects, novel objects, scrambled familiar objects or scrambled novel objects. After a short delay, they completed a two-alternative-forced-choice task, selecting between a target and a foil object. Results revealed enhanced working memory performance only for familiar objects, with no differences among the other conditions. Importantly, convolutional neural networks analyses confirmed comparable perceptual similarities between familiar and novel objects relative to scrambled stimuli. Thus, although novel objects closely resembled familiar objects, they did not enhance memory performance. Further correlation analyses revealed that subjective familiarity ratings are correlated with memory performance for familiar objects, while low-level features like colorfulness are correlated with memory performance for novel objects, suggesting that visual memory relies on different aspects to best remember each stimulus type. Overall, these results demonstrate that “object-ness” alone is insufficient to enhance visual working memory. Instead, familiarity and conceptual knowledge are critical in improving working memory performance.

36.406 RETRIEVAL FROM LONG-TERM MEMORY DOES NOT BYPASS WORKING MEMORY

Michael Mugno[1], Emma Sutcliffe[1], Timothy Vickery[1]; [1]University of Delaware

Information retrieved from long-term memory (LTM) enters working memory (WM), and the amount of information that can be retrieved is constrained to the limits of WM (Fukuda & Woodman, 2017). Liu, Li, Theeuwes, and Wang (2022) presented evidence that when WM is full, retrieved information bypasses WM. However, it is possible that retrieved LTM items can still enter WM when it is already holding many items online, albeit at the cost of the fidelities of some or all items. We investigated this by introducing continuous reporting to their paradigm. If retrieved items bypass WM, then there should be no impairment of item fidelity for either the WM or retrieved items. If retrieved items enter WM, then item fidelities should suffer compared to when retrieval is not cued. We found that when WM is full (4 items), the fidelity of LTM items suffered compared to when WM was not full (2 items). Also, the contents of WM suffered when retrieval from LTM was cued compared to baseline, under both 2-item and 4-item WM load. An additional experiment showed evidence that these effects were not due to the task design, where LTM retrieval was nested within the WM task. However, the order in which LTM or WM items were cued did affect performance – the items that were cued first and tested last suffered from reduced precision. We conclude that LTM retrieval does not bypass WM, as it robustly interferes with WM representations (and vice-versa).

Supported by NIH NIGMS grant 5P30GM145765-03

36.407 HOW DOES WORKING MEMORY STABILIZE PERCEPTUAL BIASES?

Kuo-Wei Chen[1] (kchen172@asu.edu), Scott Janetsky[1], Mckay Turner-Colling[1], Keimya Pennington[1], Gi-Yeul Bae[1]; [1]Arizona State University

While perception and working memory (WM) are fundamental functions for visually guided behaviors, perception often exhibits systematic biases, and these biases are exaggerated in WM. For example, the cardinal bias in orientation perception increases when the report is made after a delay. This increase in bias in WM suggests that the visual system recruits correction mechanisms to mitigate neural noise accumulated over time by drifting the representation toward more stable states (i.e., attractors). However, WM involves both encoding and maintenance processes, and how these two processes coordinate to mitigate neural noise is not well understood. To address this question, we investigated how cardinal bias varies over time by measuring the bias during perception (i.e., when the stimulus is in view), WM encoding (i.e., a 0-second delay between the stimulus and response), and variable WM delays (2, 4, and 6 seconds) in orientation estimation tasks. We found that the magnitude of cardinal bias was significantly greater in the longest delay condition (6 seconds) compared to the perception condition, consistent with findings from previous studies. Critically, however, the majority of the increase in bias magnitude (69% of the total increase) occurred between the perception and WM encoding conditions, rather than gradually increasing over time. In follow-up experiments, we replicated this effect with another perceptual bias (i.e., the visual tilt-illusion), demonstrating the generalizability of this effect. We confirmed that this effect persisted even when different delay conditions were randomly mixed within a block of trials, ruling out the possibility that it was driven by task-dependent strategies. Together, our results provide converging evidence that perceptual biases are stabilized in WM in two discrete steps: the WM encoding process rapidly pushes the representation toward an attractor with greater magnitude, followed by a slower, smaller drift toward the attractor during WM maintenance.

36.408 DIFFERENCES IN INTER-ITEM SALIENCE AT ENCODING PRODUCE IMBALANCED ATTENTIONAL AND REPRESENTATIONAL VISUAL WORKING MEMORY BIASES

Ryan Williams[1,2] (ryanwilliams@cmu.edu), Susanne Ferber[1,2]; [1]Carnegie Mellon University, [2]University of Toronto

We examined whether the competitiveness of a given visual working memory representation, with regard to memory-driven capture (Study 1) and inter-item memory distortions (Study 2), is modulated by salience disparities present at encoding. In both studies, individuals encoded the colors of two slanted rectangles embedded within a broader cluster of vertically oriented rectangles. For the task-relevant stimuli, salience was operationalized according to the rotation of each item, with one (high-salience) item oriented to a 45-degree angle, which popped-out from the background, and one (low-salience) item oriented to a 12-degree angle, which was less differentiated from the background items. In Study 1, individuals had to either report the color of a probed item or perform a visual search task (unpredictably). To measure memory-driven capture, a task-irrelevant color-singleton was always present in the visual search displays, which could match the color of the high-salience item, the color of the low-salience item, or neither of the two memorized colors. Here, we found memory-driven capture to be limited to the high-salience item. To examine the role of encoding salience on inter-item memory distortions, in Study 2, the colors of the two memorized items were held to a constant difference of 45-degrees in a circular hue space. Additionally, probed items were tested using 2AFC judgments in which a visually similar lure was positioned either towards or away from the unprobed item in hue space. Through these judgements, we found representations of the low-salience items to be distorted by high-salience items (i.e., attractive bias), while low-salience items had no effect on representations of high-salience items. Overall, we demonstrate that, when salience-based differences are present at encoding, the most salient item holds a competitive advantage, in that, it more strongly influences which visual inputs are prioritized in the environment and skews the representations of less salient items.

36.409 SPATIAL LAYOUT INFLUENCES THE DYNAMICS OF VISUAL WORKING MEMORY FOR BOTH CATEGORICAL AND CONTINUOUS STIMULI

Shaoying Wang[1], Srimant Tripathy[2], Haluk Öğmen[1]; [1]University of Denver, [2]University of Bradford

Working memory (WM) has limited capacity, making efficient updating crucial for everyday tasks. Prior studies using categorical stimuli (consonants) showed that updating WM involves an active removal process that is dependent on the spatial layout of stimuli (Ecker et al., 2014; Wang et al., 2024). We investigated whether these findings generalize beyond categorical stimuli that are easy to rehearse and maintain in WM. In Experiment 1, we used an articulatory suppression (verbally repeating "the") approach to minimize the effect of rehearsal. Participants viewed three random consonants in black-outlined separate boxes. Before new consonants appeared, some box outlines turned red (removal cue) to indicate which consonants participants (N=13) needed to update in WM. The interval between the onset of the removal-cue and the appearance of new letters varied from 50 to 1500 ms. Participants pressed a key when the removal and update were complete and their Reaction Times (RTs) were recorded. In Experiment 2, instead of categorical stimuli, we used a continuous orientation-judgment task. Procedures were similar to Experiment 1, but instead of consonants, we used randomly oriented Gabor-patches as stimuli and did not include articulatory suppression. Participants viewed these patches and reported their orientations via a method of adjustment. Results were similar between the two experiments: accuracies were high; RTs decreased with increasing removal-cue durations; the presence of a spatial gap between items to be updated increased RTs; under similar gap conditions, updating the leftmost item took more time compared to the rightmost item. For continuous stimuli, global updates required more time than partial updates (contrary to previous findings for categorical stimuli), likely because continuous stimuli require shorter removal time but longer encoding time. The results largely replicate those found using categorical-stimuli and suggest that similar spatio-temporal properties of WM removal- and update-processes apply to both categorical and continuous stimuli.

36.410 IS THERE A BUFFER DEDICATING TO BOUND REPRESENTATION IN WORKING MEMORY?

Xinyu Zhang[1] (12239027@zju.edu.cn), Jinglan Wu[1], Tengfei Wang[1], Mowei Shen[1], Zaifeng Gao[1]; [1]Zhejiang University

The binding problem is a central issue in cognitive science, addressing how integrated information is processed and stored in mind to support cognition and behavior. Despite its significance, whether a dedicated storage buffer exists for bound representation in working memory (WM) remains an open question. This study aims to address this question through latent variable modelling (Study 1) and psychophysical experiments (Study 2). In study 1, we adopted 13 tasks to tap the WM storage of visual objects, verbal stimuli, and bound representations, the central executive, and the storage of bound representations in long-term memory. A total of 184 participants completed these tasks. Confirmatory factor analysis revealed no evidence for an independent buffer dedicated to bound representations storage in WM. The storage of bound representations in WM closely linked to the buffer that holds their constituent elements. Meanwhile, a clear distinction emerged between WM binding and long-term memory binding. Study 2 extended the findings of Study 1 by conducting three psychophysical experiments. We employed a dual-task paradigm in which participants were required to simultaneously memorize two types of stimuli: single features and bound representations. If these two types of representations share the same buffer in WM, increasing the load on one would reduce memory performance for the other (i.e., dual competition effect). We observed dual competition between orientation and letter-location binding (Experiment 1) and between orientation and color-letter binding (Experiment 2). Meanwhile, no competition was observed between location and color-letter binding (Experiment 3), suggesting that location may have a unique role in WM. These findings align with the prediction of Study 1 and provide complementary evidence. Taken together, these studies suggest that there is no dedicated buffer for bound representations in WM, offering new insights into the structure of current WM models.

36.411 USING AI-GENERATED REAL-WORLD OBJECTS TO UNCOVER THE STRUCTURE OF VISUAL MEMORY

Jiachen Zou[1], Chen Wei[1,3], Quanying Liu[1], Maria Robinson[2]; [1]Southern University of Science and Technology, [2]University of Warwick, [3]University of Birmingham

Studying people's memory for real-world concepts remains a key challenge in vision science. Here, we introduce a novel AI-driven generative model capable of dynamically creating visual stimuli with new concepts, enabling the study of human perception and memory with previously unattainable experimental control (Wei, et al., 2024). In two visual working memory experiments we generate "morph wheels" of real-world objects (e.g., animal and plant morphs), with smooth interpolations between different object instances. We demonstrate that the similarity structure predicted by the model can be used to make parameter-free predictions of people's memory errors across different morph types (Schurgin et al., 2020). Specifically, distributions of memory errors from one morph wheel were predicted by the model's similarity structure from a different morph wheel (Brady & Robinson, 2023). Furthermore, the model can be used to generate novel exemplars and states of real-world objects with text prompts. Exemplars capture variability within a category (e.g., different kinds of chairs), while states capture dynamic transformations of an object (e.g., the same chair rotated in different ways). Critically, people's memory confusions replicated well-documented differences in memory for exemplar versus state stimuli. Finally, the model can be used to generate exemplar and state wheels with a controlled similarity structure that allows parameter free predictions of memory errors from exemplars to states. These findings highlight the versatility of the AI model, and we discuss how it provides a powerful experimental tool and theoretical framework for probing the structure of human cognition across visual domains.

Plasticity and Learning: Adaptation

SUNDAY, MAY 18, 2:45 – 6:45 PM, PAVILION

36.412 DIFFERENTIAL CHANGES IN VISUOCORTICAL ENGAGEMENT DURING AVERSIVE VERSUS NON-AVERSIVE ASSOCIATIVE LEARNING TASKS

Sarah Gardy[1] (sgardy@ufl.edu), Laura Ahumada[1], Faith Gilbert[1], Hannah Engle[1], Christian Panitz[2], Andreas Keil[1]; [1]University of Florida, [2]University of Bremen

Aversive conditioning experiments repeatedly pair a neutral stimulus (CS+) with a stimulus (US, e.g., loud noise) that reliably elicits an unconditioned response. Other stimuli are not paired with an aversive stimulus. Eventually, the CS+ becomes predictive of the US, which is accompanied by visuocortical facilitation and heightened sensory gain when viewing the CS+. However, it is unclear whether this visuocortical response is specific to defensive/aversive association formation or reflects associative memory formation in general. To examine the possible impact of non-aversive associative learning on visual processing, a soft tone (65 dB) was paired with one of two possible high-contrast black-and-white circular gratings. Young adult (n=22) participants viewed four Gabor gratings with variations in orientation. Either a 15° or 75° grating was paired with the tone (counterbalanced across participants), while 35° or 55° orientations were never paired (generalization stimuli, GSs). The gratings were presented for 3000 ms and the soft tone was played only during the last 1000 ms of the CS+, co-terminating with the grating. Analyses were conducted on the first 2000ms of grating presentation (i.e., timespan before tone). A Morlet wavelet analysis was used to quantify activity in the alpha-band. Spectral estimates of the ssVEP amplitude were obtained. Results indicate that alpha power increases, rather than decreases, when expecting the tone (CS+). The ssVEP showed a small effect with the non-predictive cues prompting higher amplitude than the conditioned stimulus—the opposite of what has been found during aversive conditioning. Together these findings suggest that the visuo-cortical changes observed during aversive conditioning are specific to situations in which the US elicits a defensive response. Neutral US by contrast prompt responses similar to those observed in working memory, imagery, and anticipation tasks.

36.413 IMPLICIT LEARNING OF SOCIAL INFORMATION IN CONTEXTUAL CUEING

Lijeong Hong[1] (lijeong_hong@yonsei.ac.kr), Min-Shik Kim[1]; [1]Yonsei University

People implicitly learn target-distractor co-occurrences, a process that facilitates visual search performance, known as contextual cueing (Chun & Jiang, 1998, 1999). While perceptual schemas—like expecting to find a pen on a desk—are acquired through visual experience, real-world tasks often involve social contexts, such as recognizing familiar individuals in crowded environments. This study investigates whether social information about interpersonal associations can be implicitly learned as a contextual cue and used to guide attention during visual search tasks. Notably, the contextual cue in this study was independent of the spatial configuration of all stimuli, including the target's location. Participants were asked to find a female face (target) among ten male faces (distractors). In the consistent mapping condition, specific pairings between target and distractor faces were maintained across all blocks, while in the variable mapping condition, target-distractor pairings were randomized. For example, in the consistent condition, a target face (e.g., Kim Kardashian) was always paired with specific distractor faces (e.g., the other Kardashian family members), whereas no such pairing existed in the variable condition. Importantly, spatial configurations of all stimuli were randomized in both conditions. Results showed that participants implicitly learned the associations between target and distractor identities, encoding interpersonal contextual information and using it to guide attention during the visual search. These results indicate that social information about interpersonal associations can be implicitly learned and served as contextual cues, enabling individuals to anticipate the target's identity and enhancing visual search performance. This study suggests that contextual cueing research could extend to other dimensions of social information, such as emotional expressions or group dynamics.

36.414 INVARIANCE OF VISUAL STATISTICAL LEARNING

József Fiser[1], Zoltán Rácskay; [1]Central European University

A staple feature of object representation and perception is invariance: stable representation of new objects is learned from perpetually varying visual inputs and objects are identified as same despite of large differences in the appearance of their local features. Yet, visual statistical learning (VSL) paradigms investigating the formation of such object representations typically use static images of standardized visual patterns as local features. In three experiments using a Virtual Reality environment, we explored how human VSL is affected when the stable visual structures to be learned appear in a more natural dynamic manner by where the classical spatial VSL patterns appeared in a plane that perpetually changed its 3D position, rolled, pitched and yawed in a random fashion. First, we replicated the classical static results showing that adults have a significant familiarity preference (59%) for shape-pair structures that were used to generate composite scenes in the preceding task-free exposure phase of the experiment. Next, we rerun the experiment with moderate dynamic movement of the plane during the exposure phase and found that preference became indistinguishable from chance performance indicating the disappearance of implicit learning under these conditions. In the third experiment, we increased further the magnitude of the dynamic movement of the plane holding the composite patterns. Surprisingly, instead of staying in the ”no-learning” regime, participants developed a significant familiarity preference for the alternative lure shape pairs during the test trials (40%). This preference was significantly different not only from chance performance but also from the preference in the moderate condition. These results suggest that instead of being simply eliminated, the effects of dynamic variations in the input during exposure are associated with the features and can take part in active interference during the familiarity judgement.

This work has been supported by Grant ANR-FWF I 6793-B.

36.415 REPRESENTATIONAL DRIFT AND REPETITION SUPPRESSION: IDENTICAL OR DISTINCT PHENOMENA?

Zvi N. Roth[1], Elisha P. Merriam[2]; [1]Gonda Brain Research Center, Bar-Ilan University, [2]Laboratory of Brain and Cognition, National Institute of Mental Health

Several recent studies have reported that stimulus representations are not temporally stable, but rather exhibit cumulative, directional changes over hours, days, and months, a phenomenon termed 'representational drift'. Many of these studies used a stimulation protocol in which stimuli were repeated multiple times within each recording session. Repetition suppression is a well-established phenomenon whereby repeated presentations of a stimulus evoke decreasing neural responses. Here, we ask whether repetition suppression affects measures of representational drift, or alternatively, whether the two are entirely distinct phenomena. Representational drift has recently been demonstrated in human V1 using the Natural Scenes Dataset (NSD), which consists of fMRI responses to a large number of unique images repeated up to 3 times over the course of a year. Computer simulations revealed that repetition suppression could, in principle, produce changes similar to those we have attributed to representational drift. We next tested whether repetition suppression does indeed account for representational drift. First, we reanalyzed the NSD, limiting ourselves to first-repetition trials only, while subsampling each session to equalize the number of trials across sessions. While this subsampling procedure drastically reduced the amount of data used in the analysis, we nonetheless found robust evidence for representational drift. Second, we analyzed the THINGS dataset, which contains a large set of images that were not repeated (we didn't analyze the subset of 100 stimuli that were repeated). Remarkably, two out of three subjects exhibited significant drift, even though the THINGS dataset is considerably smaller than NSD. These findings indicate that although repetition suppression poses a major potential confound in representational drift studies, the two phenomena are distinct. Our results suggest that neural representations are subject to multiple time-dependent changes at different time scales. The role of each of these processes in cognition, perception, and learning remains an important and unresolved question.

36.416 THE EFFECT OF CONTOUR ERASURE ON ARTIFICIAL SCOTOMA FILLING-IN

Yih-Shiuan Lin[1] (yihshiuan.lin@gmail.com), Serena Castellotti[2,3], Maria Del Viva[3], Chien-Chung Chen[4,5], Mark Greenlee[1]; [1]Institute of Psychology, University of Regensburg, [2]Department of Translational Research on New Technologies in Medicine and Surgery, University of Pisa, [3]Department of Neurosciences, Psychology, Drug Research and Child Health (NEUROFARBA), University of Florence, [4]Department of Psychology, National Taiwan University, [5]Neurobiology and Cognitive Science Center, National Taiwan University

Perceptual filling-in describes how the visual system compensates for missing information by interpolating from surrounding features. Two examples are artificial scotoma (AS) filling-in, perceived within blank regions embedded in dynamic noise after prolonged fixation, and contour erasure (CE), where brief adaptation to a flickering outline renders an object instantaneously invisible. Our study examines the interaction between AS and CE filling-in effects to reveal the underlying mechanisms. AS was induced using a high-contrast (-1dB) dynamic noise background (10 Hz) with two 7°-eccentric blank disks (3° diameter) in the left and right upper quadrants. With steady fixation, AS filling-in could be induced within these blank regions. CE was induced by counterphase flickering high-contrast (-1dB) contours (6 Hz) that outlined these blank regions. After 8 seconds of adaptation to mean luminance, CE, AS, or both, two stimuli (random noise disks, duration 100ms) were shown at the filling-in locations: participants had to discriminate the target, superimposed randomly on either side of the pedestal locations. Pedestal contrasts ranged from -∞ to -10 dB. Target thresholds showed a dipper shape, decreasing at low pedestal contrasts and increasing at high contrasts. Both AS and CE adaptation increased target thresholds when presented alone. However, superimposing CE on AS resulted in minimal threshold increase, and participants reported less AS filling-in, when CE was present. When CE preceded AS adaptation by 2 seconds, the percentage of subjective filling-in reports increased (from 58% to 73%), and the time to perceived filling-in was reduced (from 6.09s to 5.11s). These findings suggest that CE can facilitate AS filling-in when presented sequentially, but not simultaneously. Our results can be explained by a divisive inhibition model in which an inhibitory sensitivity parameter and a normalizing constant capture the adaptation effects. Our findings highlight the interplay between these two filling-in effects.

This study is supported by the Deutsche Forschungsgemeinschaft (DFG, GR 988/27-1) and the National Science and Technology Council (NSTC, 110-2923-H-002-004-MY3).

Development: Neural

Sunday, May 18, 2:45 – 6:45 pm, Pavilion

36.417 TRAINING WITH DEGRADED DATA FOR DEBIASED REPRESENTATION IN NEURAL NETWORKS

Jeonghwan Cheon[1], Se-Bum Paik[1]; [1]Korea Advanced Institute of Science and Technology

Infants have limited sensory capacities, including low visual acuity and color sensitivity. These limitations are often considered inevitable due to early sensory and cortical immaturity, though the functional role of this process remains unclear. In contrast, machine learning models typically maintain high sensory acuity throughout the learning process, often becoming trapped in local optima biased by peripheral properties in the training data. As a result, these models fail to generalize to unbiased data, where peripheral bias attributes have been removed. In this study, we demonstrate that initial sensory degradation can effectively guide neural networks to learn robust representations without being biased toward spuriously correlated peripheral attributes. Inspired by the developmental process of sensory systems, we trained neural networks on an initially grayscaled and blurred dataset, gradually increasing color sensitivity and spatial resolution during training, as proposed previously (Vogelsang et al., 2024). Specifically, we used spuriously correlated image benchmarks for training, such as Colored MNIST and Corrupted CIFAR-10, where biased attributes (i.e., color or corruption type) are strongly correlated with intrinsic attributes (i.e., shape or object). Both the conventional and initial degradation training schemes achieved comparable performance on bias-aligned samples. However, the initial degradation approach significantly outperformed the conventional scheme on bias-conflicting samples, where spuriously correlated peripheral biases are removed. We found that neural representations learned through conventional training often capture biased attributes, while those learned with initial degradation focus more on intrinsic attributes. Our findings suggest that initial sensory degradation is a key biological strategy that enables networks to learn debiased, robust neural representations without overfitting to biased attributes. This highlights a simple yet powerful biological strategy for learning robust neural representations and offers a potential solution to debiasing challenges in machine learning.

This work was supported by the National Research Foundation of Korea (NRF-2022R1A2C3008991 to S.P.) and by the Singularity Professor Research Project of KAIST (to S.P.).

36.418 DOMAIN-GENERAL DIFFERENCES IN HIGH-LEVEL VISUAL PROCESSING IN CHILDREN WITH

DYSLEXIA AND YOUNG READERS: IMPACT OF READING SKILL DEVELOPMENT ON STIMULUS AND CATEGORICAL REPRESENTATIONS?

Florence Bouhali[1,2], Ting Qi[1,3], Maria Luisa Mandelli[1], Cheng Wang[1,4], Rian Bogley[1], Jocelyn Caballero[1], Rayburn Tang[1], Ellie Carpenter[1], Maria Cecilia Ferrer Ladao[1], Christa Pereira Watson[1], Kevin Weiner[5], Maria Luisa Gorno Tempini[1], Fumiko Hoeft[6]; [1]University of California San Francisco, [2]Aix-Marseille University, France, [3]Beijing University of Posts and Telecommunications, China, [4]Fujian Medical University, China, [5]University of California Berkeley, [6]University of Connecticut

Developmental dyslexia is a specific learning disorder characterized by challenges in fluent reading, despite conventional instruction and average to superior intelligence. Interestingly, many studies also report domain-general challenges in visual processing in individuals with vs. without dyslexia, with less fine-grained exemplar discriminations of visual stimuli (e.g., faces). Moreover, less neural adaptation to repeated stimuli, including images of words, faces and objects, has been reported in dyslexia. Several theories therefore propose that these characteristics present before literacy instruction may be causally implicated in the emergence of dyslexia. Yet, less adaptation to various visual categories has also been observed in illiterate adults, together with differences in the functional organization of category-selective regions in ventral occipito-temporal cortex (vOTC), thereby questioning the causality between reading proficiency and general visual processing. Here, we test the impact of typical and atypical reading development on visual processing in 39 children with dyslexia and 36 without, aged 7 to 17, using a standard repetition detection (one-back) localizer task in the fMRI. We find altered stimulus-level representations in children with dyslexia or young readers with low raw reading scores, with significantly lower hit rates across all categories (words, numbers, limbs, faces, objects and places), regardless of ADHD and other confounding variables, and neural differences in repetition effects. Visual processing differences were also found at the categorical level: dyslexic/young readers showed lower categorical distinctiveness (i.e., lower differences in multivoxel activation pattern similarity computed within a category than across categories) in the left and right vOTC for bodies, faces and words – suggesting noisier neural representations. Our results therefore confirm domain-general visual challenges in dyslexia, linking behavioral and neural effects. Together with results in illiterate adults, our work further suggests that these visual challenges in dyslexia may be secondary to reading difficulties, with implications for existing theories of dyslexia.

36.419 DEVELOPMENTAL VARIABILITY IN LOCATION AND FUNCTION OF HIGH-LEVEL VISUAL AREAS

Kelly J. Hiersche[1] (hiersche.1@buckeyemail.osu.edu), Zeynep M. Saygin[1]; [1]Department of Psychology, The Ohio State University

The ventral temporal cortex contains regions of the brain specialized for recognizing high-level visual categories; the visual word form area (VWFA), the fusiform face area (FFA), and the object-selective posterior fusiform sulcus (PFS). While each region falls in the same general location across individuals, there is substantial individual variability in the exact location of the selective cortex, and we know little about the differences in variability across these regions or the development of variability. Here, we scan over 50 adults and over 100 children (ages 3-12) on a visual localizer to characterize individual variability in word, face, and object fROIs in the mature brain, and examine the development of this variability in children. Variability was measured as the squared distance from the adult group mean (of either location or selectivity of the fROIs) for both children and adults. In adults, variability in selectivity is similar for faces, objects, and words. Children show higher variability in selectivity for words than faces and objects, and higher variability than adults for word- and object-selectivity. Children converge towards adult-like levels of variability in word-selectivity (i.e., variability decreases with age), but we do not observe age-related changes in variability of object selectivity. Interestingly, children and adults show similar variability in fROI location, including for the VWFA, which is highly experience-driven. These results suggest that despite ongoing functional tuning of e.g. VWFA, the location of category-selective visual cortex is pre-constrained, perhaps by innately determined mechanisms (as suggested by the connectivity hypothesis). Ongoing work compares these visual regions to other cortices including higher-order language and social cognition. Understanding the developmental trajectory of individual variability can help us better understand the emergence of knowledge domains in the brain.

NIH R01 HD110401-01 and Alfred P Sloan: FG-2018-10994 (to Z.M.S.); NSF GRFP: DGE-1343012 (to K.J.H.)

36.420 SINGLE NEURON DIVERSITY SUPPORTS AREA FUNCTIONAL SPECIALIZATION ALONG THE VISUAL CORTICAL PATHWAYS

Julio Martinez[1] (julio.martinez@robarts.ca), Michael Feyerabend[1], Stefan Pommer[2], Andreas Neef[2], Jochen Staiger[2], Wataru Inoue[1]; [1]Schulich School of Medicine and Dentistry, Western University, London, Ontario, [2]Institute for Neuroanatomy, University Medical Center Göttingen, Georg-August University, Göttingen, Germany

Primates have specialized visual pathways composed of interconnected cortical areas. The input area V1 contains neurons that encode basic visual features, whereas downstream in the lateral prefrontal cortex (LPFC), neurons can acquire tuning for novel complex feature associations. It has been tacitly assumed that each visual cortical area is composed of repeatable neuronal subtypes (i.e., cell types follow the principle of serial homology), and that variations in synaptic strength and connectivity patterns underlie an area's functional specialization. Here, we test the hypothesis that diversity in the intrinsic makeup of single neurons contributes to area specialization along the visual pathways. We recorded intracellularly from 463 neurons in slices obtained from 42 marmosets (Callithrix jacchus). Of these, 363 neurons passed a quality control pipeline: 107 V1 and 256 LPFC cells (BA 8/46). A total of 144 neurons were examined histologically and classified as aspiny/inhibitory (29%) or spiny/excitatory (71%). Additionally, 32 cells were reconstructed. We measured morphological and electrophysiological features of single neurons and trained a random forest classifier using these features to separate neurons into three main subtypes: excitatory neurons, fast-spiking interneurons, and non-fast-spiking interneurons. We focus on fast-spiking inhibitory interneurons and excitatory cells because the classifier reliably identified these categories. Excitatory neurons in the LPFC were larger, less excitable, and fired broader spikes than V1

neurons. Some inhibitory fast-spiking interneurons in the LPFC had longer axons and fired spikes with longer latencies and a more depolarized action potential trough than those in V1. Intrinsic bursting was found in subpopulations of both excitatory and inhibitory LPFC neurons, but not in V1 neurons. The latter may favor temporal summation of spikes and, therefore, enhanced synaptic plasticity in the LPFC relative to V1. Our results show that specialization within the primate visual system permeates the most basic processing level: the single neuron.

36.421 DO RETINAL BIASES IN CATEGORY-SELECTIVE REGIONS OF THE VENTRAL TEMPORAL CORTEX DEPEND ON EXPERIENCE?

Tiago Mesquita[1] (tiago.mesquita@uni-hamburg.de), Gargi Majumdar[1], Waqar Khan[1,2], Anuhya Nalluri[2], Sunitha Lingareddy[3], Ramesh Kekunnaya[2], Brigitte Röder[1,2]; [1]Biological Psychology and Neuropsychology, University of Hamburg, [2]Jasti V Ramanamma Children's Eye Center, Child Sight Institute, LV Prasad Eye Institute, Hyderabad, India, [3]Lucid Medical Diagnostics, Hyderabad, India

Recent studies have demonstrated that face-selective regions in the ventral occipital temporal cortex (VOTC) receive visual information mainly from central vision whereas place-selective regions show a preference for the peripheral visual field. While the basic retinotopy in early visual cortex seems to be established by mid-childhood or earlier, this foveal bias in face-selective regions seems to emerge with a longer developmental trajectory. While category-selective regions in the VOTC have been observed in infants, the role of experience in the emergence of these retinal biases in the VOTC is still unknown. A first step in addressing this question is to demonstrate the foveal bias for face-selective regions in VOTC in a different cultural settings (ethnicity). We here adapted a specialized fMRI mapping procedure tailored to map category-selective regions in VTC using Indian faces. Images of faces and places were presented at different visual field locations. Regions of interest (ROI) (fusiform face area (FFA) and parahippocampal place area (PPA)) were identified using a typical localizer paradigm contrasting five different visual categories presented. Next, we computed the population receptive fields (pRFs) within each ROI. We successfully replicated the foveal bias for face processing in the FFA. For place processing, PPA featured a more distributed activation across the visual field. We concluded that retinal biases in FFA are a universal organizing principle. Whether they require early exposure to faces in order to emerge needs to be tested in people who lack such experience, e.g. in people recovering from a transient phase of congenital blindness.

36.422 EXAMINING THE NEURAL MECHANISMS SUPPORTING MENTAL ROTATION IN YOUNG CHILDREN

Lauren Aulet[1] (laulet@andrew.cmu.edu), Jessica Cantlon[1]; [1]Carnegie Mellon University

Spatial skills are critical predictors of educational success, STEM career trajectories, and future income potential. Despite their significance, the neural foundations of spatial cognitive development remain poorly understood. In the present study, we addressed this gap by examining the neural mechanisms underlying spatial cognition in young children (N = 19; mean age = 8.16 years). Participants perform a mental rotation task during functional magnetic resonance imaging (fMRI). The task consisted of blocks of 0°, 50°, and 100° rotation trials, where participants made same/different judgments about pairs of objects. 'Different' trials were always comprised of an object and a mirror-reversed version of the same object. To measure the effect of angularity disparity, behavioral performance was quantified as the slope of mean accuracy on rotation angle. Children exhibited slopes significantly different from zero, indicating robust angular disparity effects, such that mean accuracy decreased as rotation angle increased. Univariate fMRI analyses comparing rotation versus non-rotation blocks revealed significant activation in multiple regions: bilateral parietal cortex, bilateral motor cortex, anterior cingulate cortex, and thalamus. Analysis of brain-behavior relationships demonstrated significant positive correlations between angular disparity and activation in parietal cortex, indicating that greater angular disparity corresponded to increased neural activity in these regions, suggesting these regions may be involved in representing spatial orientation of the objects. Conversely, significant negative correlations were observed in motor cortex and thalamus, where greater neural activity was associated with reduced angular disparity, suggesting these regions may be involved in the speed or efficiency of mental rotation. Importantly, these correlations remained significant after controlling for age and general intelligence. This study represents the first investigation of the neural bases of mental rotation in young children and suggests that similar neural mechanisms may support mental rotation across development from childhood to adulthood.

36.423 INVESTIGATING VISUAL AND AUDITORY LANGUAGE REPONSES IN THE VENTRAL TEMPORAL CORTEX OF PRE-READING CHILDREN: A MULTIVARIATE ANALYSIS

Lauren A. Rydel[1] (rydel.2@osu.edu), Kelly J. Hiersche[1], Zeynep M. Saygin[1]; [1]Ohio State University

The visual word form area (VWFA) is a region within the ventral temporal cortex (VTC) that shows category selectivity (i.e. univariate selectivity) for visual words and letter strings but only after an individual learns how to read. What precedes this response? Previous work demonstrates that children lacking category selectivity for face, bodies, or scenes in VTC display mature representational patterns (i.e. multivariate selectivity), suggesting distributed patterns of responses may scaffold later category selectivity. Does the VWFA also emerge from initially multivariate representations in pre-readers? Given that the VWFA is a 'bridge' between language and vision, does it show linguistic multivariate representation? In this study, we investigate whether pre-reading children exhibit a multivariate response to written words or auditory language across the VTC. Next, we examine if the voxels that decode written words can also distinguish between auditory language and control conditions. Children (N=74): 35 pre-readers ages 2-6 and 39 readers ages 4-13, completed two runs each of a high-level visual fMRI localizer (word, scrambled word, object, and face conditions), and auditory language task (linguistically meaningful sentences, matched nonword sentences, and texturized sound conditions). Multi-voxel pattern analysis (MVPA) identified voxels within the VTC with higher within-condition correlations than between conditions. Despite lacking a univariate word response, pre-readers show the start of a multivariate pattern for decoding words within the VWFA, similar to readers (who have a multivariate and univariate word response). Further, these multivariate word voxels (but not other voxels) also show a strong within-condition correlation for auditory sentences, but not auditory control conditions, in both pre-readers and readers, suggesting initial distributed linguistic representations in a visual area. Overall, distributed responses may indeed scaffold later-developing category selectivity even for new knowledge domains. Ongoing work investigates the shift to categorical selectivity in longitudinal development, and contribution of connectivity and other anatomical constraints.

NIH R01 HD110401-01 and Alfred P Sloan: FG-2018-10994 (to Z.M.S.); NSF GRFP: DGE-1343012 (to K.J.H.)

36.424 A NOVEL APPROACH FOR ACCURATE MRI-GUIDED STEREOTAXIC TARGETING IN NONHUMAN PRIMATES

Peyton Harmon[1], Bruno Averbeck[1], Reza Azadi[1]; [1]Laboratory of Neuropsychology, National Institute of Mental Health, Bethesda, MD

Traditional MRI-guided neurosurgical targeting in nonhuman primates (NHPs) often relies on MRI-visible earbars in the stereotaxic devices. However, variability in repositioning can compromise the accuracy of these methods. To address this, we developed a reliable approach using fiducial markers implanted into the skull, enabling direct coregistration of MRI scans with stereotaxic coordinates based on skull position. Human fiducial markers are often unsuitable for NHPs due to their large size. Therefore, we developed a novel method using MRI-compatible cranial screws made for NHP surgeries. These screws are detectable on T1-weighted MRI scans due to the contrast with surrounding soft tissue. To evaluate this method, we first implanted 28 ceramic screws into a 3D-printed macaque skull surrounded by a contrast agent gel. The coordinates of each fiducial were manually measured on T1-weighted MRI and in a stereotaxic device, both with and without altering the skull's position. These coordinates were cooregistered using a 3D affine transformation. The results showed reliable and accurate alignment even after repositioning the 3D printed skull within the stereotaxic device. The average error was 0.383 mm using all 28 fiducial markers. Moreover, a resampled analysis showed that using as few as six screws provided nearly equivalent reliability, with an average error of 0.481 mm. Additionally, altering the 3D-printed skull position in the stereotaxic frame did not cause any significant change in the error. This novel approach eliminates dependence on earbar coordinates and their associated variability, providing more accurate and reproducible reference points for brain regions. Furthermore, similar methods can be used for coregistering MRI and CT scans.

This work was supported by the Intramural Research Program of the National Institute of Mental Health (ZIA MH002928) to B.A.

Development: Amblyopia, binocular

Sunday, May 18, 2:45 – 6:45 pm, Pavilion

36.425 ARE RETINAL VASCULATURE FEATURES ALTERED IN AMBLYOPIA?

Rijul Saurabh Soans[1], Susana T. L. Chung[1]; [1]University of California, Berkeley

Amblyopia is a developmental disorder characterized by impaired vision in one eye and is a leading cause of visual impairment in children. Early intervention in life improves the chances of restoring normal visual functions. Evidence exists suggesting that retinal vasculature may differ in individuals with amblyopia, potentially enabling automated retinal image-based detection of the disorder. This study used computer vision to investigate whether there are structural differences in retinal blood vessels between individuals with amblyopia and neurotypical controls. We first performed retinal vessel segmentation using the Spatial Attention-UNet network on OCT fundus images obtained from 22 adults with amblyopia (9 males; 37.68±18.33 years; 11 anisometropic; 11 strabismic) and 40 neurotypical (control) participants (13 males; 23.45±3.46 years). Four features were then quantified from the segmented vasculature images: Vascular Area (VA): number of pixels within segmented vessels; Fractal Dimension (FD): complexity of vascular network by measuring blood vessel density distribution patterns through regression analysis of vessel-block counts across scaled grid sizes; Vascular Skeleton Length (VSL): pixel count of the skeletonized vessel tree; and Vascular Bifurcation Points (VBP): branching complexity of the vessel skeleton indicating morphological richness. Within the amblyopia group, no significant differences were found between amblyopic and fellow eyes of the amblyopes, or between the anisometropic and strabismic groups, for the four vasculature features. Comparing amblyopic eyes to control eyes, significant differences emerged for VA (p=0.01; Cohen's d=-0.66), FD (p=0.04; d=-0.56), VSL (p=0.006; d=-0.76), but not for VBP (p=0.83). Fellow eyes also differed from control eyes for VA (p=0.03; d=-0.58), FD (p=0.03; d=-0.58), VSL (p=0.02; d=-0.59), but not for VBP (p=0.93). These results indicate significant structural differences in the retinal vasculature between the eyes (both amblyopic and fellow eyes) of amblyopes and neurotypical participants, suggesting that certain structural features of retinal blood vessels might be useful as biomarkers for detecting amblyopia.

36.426 FIXATION INSTABILITY IS ASSOCIATED WITH SLOW READING DURING RAPID SERIAL VISUAL PRESENTATION (RSVP) IN CHILDREN WITH AMBLYOPIA

Krista Kelly[1] (krista.kelly@uwaterloo.ca), Dorsa Mir Norouzi[2], YiZhong Wang[2]; [1]University of Waterloo, [2]Retina Foundation of the Southwest

Introduction: Children with amblyopia silently read paragraphs 28% slower than their peers during binocular viewing, which is associated with ocular motor dysfunction (i.e., fixation instability, increased forward saccades). We previously reported slow reading even when the need for inter-word saccades is removed using rapid serial visual presentation (RSVP) (Mir Norouzi VSS 2022). Yet, this does not necessarily rule out a role for ocular motor dysfunction in slow reading. Here, we report on fixation instability during RSVP reading in amblyopic children compared to controls. Methods: Amblyopic (n=33) and control (n=30) children ages 8-12 years silently read grade-appropriate sentences presented in RSVP (single word presentation at screen center) during binocular viewing. Exposure time per sentence changed using a descending adaptive 2 down-1 up staircase method. The child's reading speed threshold in log words per minute (WPM) was obtained. Eye movements were tracked simultaneously with the EyeLink 1000 binocular eye tracker to determine fellow eye (FE) and amblyopic eye (AE) fixation stability (log BCEA) during reading. Results: Compared to controls, amblyopic children read slower (2.8±0.5 log WPM vs 3.1±0.4 log WPM, p=0.009), and had increased AE fixation instability (0.21±0.39 log BCEA vs −0.20±0.18 log BCEA, p<0.001) and increased FE fixation instability (−0.03±0.34 log BCEA vs −0.20±0.15 log BCEA, p=0.012) during RSVP reading. Reading rate in amblyopic children with good FE stability (n=11) did not differ from controls (p=0.63) and was faster than those with poor FE stability (p=0.026). Amblyopic children with poor FE stability read slower than controls (p=0.001). Conclusions: Amblyopic children read slower than control children, even when the need for inter-word saccades is removed (i.e., RSVP reading). The direct relationship of slow RSVP reading with poor FE fixation stability during binocular viewing provides support for fixation instability as a source of slow natural reading in children with amblyopia.

National Institutes of Health – National Eye Institute EY028224

36.427 A COMPUTATIONAL MODEL OF THE DEVELOPMENT OF VERGENCE AND ACCOMMODATION CONTROL AND THEIR INTERACTION

Francisco M. López[1,2] (lopez@fias.uni-frankfurt.de), Theresa Lundbeck[1], Bertram E. Shi[3], Jochen Triesch[1]; [1]Frankfurt Institute for Advanced Studies, Germany, [2]Xidian-FIAS International Joint Research Center, Germany, [3]Hong Kong University of Science and Technology, Hong Kong

Fixating objects involves two kinds of eye movements: vergence movements to align the two optical axes on the same point and accommodation control to ensure sharp images. It has long been known that vergence and accommodation control are coupled, but it is presently unclear how they are calibrated during development and how their coupling emerges. Here we propose a computational model based on the Active Efficient Coding framework to explain the joint learning of vergence and accommodation control using intrinsically-motivated reinforcement learning to maximize the mutual information between binocular images and their internal encoding. Training proceeds in a realistic virtual environment with binocular vision and longitudinal chromatic aberration. We compare different model architectures permitting different degrees of interaction between accommodation and vergence control, allowing us to evaluate to what extent these are necessary to replicate experimental results. Our findings suggest that vergence and accommodation can act independently but that early access to visual features used by one controller is enough to replicate the development of a coupling as measured in laboratory experiments.

This research was supported in part by “The Adaptive Mind” funded by the Hessian Ministry of Higher Education, Research, Science and the Arts, Germany, and by the Hong Kong Research Grant Council (Grant 16213322). JT was supported by the Johanna Quandt foundation.

36.428 GAIT DEVELOPMENT IN CHILDREN WITH IMPAIRED BINOCULARITY

Lauren Hoare[1], Krista Kelly[1]; [1]University of Waterloo

Introduction: Coordination between the eyes and body is important for navigating the environment. Children with disrupted binocular vision from strabismus and amblyopia score lower for walking on a standardized test of motor ability (Kelly et al., 2018 IOVS). However, standardized tests do not assess gait kinematics during walking. Here, we investigated the development of gait kinematics during walking in children with impaired binocularity compared to controls. Methods: 9 children ages 7-13 years with impaired binocularity (stereoacuity, 3.0±0.8 log arcsecs) due to strabismus or amblyopia and an age-similar group of 18 controls were enrolled. Children walked the length of a GAITRite pressure-sensitive walkway and completed 3 conditions of varying complexity: 1) Straight Walk (W): walk on mat, 2) Isolated Target Walk (I): walk and step on two-dimensional targets, and 3) Distractor Target Walk (D): walk and step on two-dimensional targets while avoiding two-dimensional distractors. Gait kinematic outcomes were normalized velocity (leg lengths/second), step time (msecs), and accuracy (%) of stepping on targets or avoiding distractors. Results: Controls were faster during straight walk compared to the other two conditions (normalized velocity, W:1.49±0.23, I:1.42±0.24, D:1.38±0.25 leg lengths/sec; longer step time, W:508±52, I:522±48, D:525±51 msecs, ps≤0.011). Children with impaired binocularity showed the same pattern (normalized velocity, W:1.67±0.35, I:1.58±0.29, D:1.50±0.27 leg lengths/sec; step time, W:468±63, I:494±61, D:491±56 msecs, ps≤0.06). While there were no group differences for velocity or step time (all ps≥0.089), children with impaired binocularity were less accurate than controls when stepping on targets, especially when paired with distractors (90.0±12.5% vs 98.2±3.0%, p=0.015). Conclusions: Impaired binocularity in children did not influence walking speed but impacted accuracy, especially with a more complex task. Impaired binocularity early in life may impact gait by causing visual uncertainty as children walk or may result in the development of compensatory strategies that are less effective during navigation.

NSERC RGPIN-2024-03992

36.429 DEVELOPMENT OF DISPARITY TUNING IN FERRET PRIMARY VISUAL CORTEX

Allison Murphy[1], Kristina Nielsen[1]; [1]Johns Hopkins University

Binocular disparity, the difference in position of an image on each retina, is an important cue the brain uses to generate perception of depth. Several areas of the brain, including primary visual cortex (V1), contain neurons tuned for binocular disparity. While the encoding of disparity in V1 has been studied in adults, less is known about how it develops in younger animals. However, this is important for understanding the development of stereoscopic vision as well as how it may be disrupted in disorders such as amblyopia. The ferret is a highly visual animal and undergoes much of its development postnatally, making it an ideal model for studying the development of the structure and function of the visual brain. Despite the prevalence of the ferret as a model of visual system development, little is known about its capacity for stereoscopic vision compared to other animal models. To address this, we performed electrophysiological recordings in primary visual cortex (V1) in ferrets aged postnatal day 30 to adult. Utilizing a mirror stereoscope, we presented binocular and monocular stimuli and systematically varied the direction and binocular disparity of drifting gratings. We analyzed single-unit responses to assess the joint development of ocular dominance, direction tuning, and binocular disparity tuning across several developmental timepoints. We find that disparity-tuned neurons are present in V1 as early as postnatal day 30, although the proportion of tuned neurons and quality of tuning varies over time. These results are a first step in establishing the ferret as a developmental model for stereoscopic vision.

1R01EY035807

36.430 INCREASED POPULATION RECEPTIVE FIELD SIZE IN EARLY VISUAL CORTEX FOLLOWING THE LOSS OF ONE EYE

Matthew C. Macdonald-Dale[1], Stefania S. Moro[1,2], Diana J. Gorbet[1], Remy Cohan[1], Peter J. Kohler[1], Jennifer K. E. Steeves[1,2]; [1]Centre for Vision Research, York University, Toronto, Canada, [2]The Hospital for Sick Children, Toronto, Canada

The surgical removal of one eye (monocular enucleation) early in life results in partial visual deprivation and offers a unique model for examining the consequences of the loss of binocularity on the maturation of the visual system. Adults who have undergone early monocular enucleation exhibit several long-term visual, auditory, and audiovisual behavioural adaptations as well as structural and functional adaptations within and between cortical and subcortical regions. In the current study, we investigated whether topographically organized visual cortex is influenced by the reduction in input to the visual system following monocular enucleation. We used functional Magnetic Resonance Imaging (fMRI) to map population Receptive Fields (pRFs) in individuals who have had one eye removed, as well as binocularly intact controls. pRF mapping was performed using a standard checkerboard bar swept across the visual field while participants performed a central fixation task. For each location on the cortical surface, we determined the size and visual field location of the pRF that accounted for the maximum variance in the observed fMRI signal. In early visual areas (V1, V2, and V3) of both hemispheres, pRFs were systematically larger for people with one eye compared to controls. Additionally, there was a larger rate of pRF expansion with eccentricity in the hemisphere contralateral to the remaining eye. No differences were observed for other parameters, including overall model fit, eccentricity, or gain. In summary, monocular enucleation may induce sensory plasticity where neurons in early visual cortex respond to a larger portion of the visual world. This is consistent with previous reports of reorganization within the LGN contralateral to the remaining eye and increased surface area in the primary visual cortex following monocular enucleation. These findings, in conjunction with behavioural and neuroimaging studies, contribute to the broader understanding of the effects of sensory deprivation from eye enucleation.

This research was supported by grants from Canada First Research Excellence Fund: Vision Science to Application (VISTA) (CFREF-2015-00013); Natural Sciences and Engineering Research Council of Canada (NSERC) (327588); Canada Foundation for Innovation (CFI) (12807).

36.431 EXPLORING THE INFLUENCE OF BINOCULAR ANTI-COACTIVATION ON OCULAR DOMINANCE COLUMN FORMATION BEYOND PRIMARY VISUAL CORTEX

Sohrab najafian[1], Vladimir K. Berezovskii[1], Michael J. Arcaro[2], Margaret S. Livingstone[1]; [1]Department of Neurobiology, Harvard Medical School, Boston, MA, USA, [2]Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

Inputs from the lateral geniculate nucleus segregate into ocular dominance columns in layer 4 of the macaque primary visual cortex (V1) via Hebbian mechanisms. Beyond V1 layer 4, most visual neurons are binocular. Research on developing frogs with transplanted third eyes revealed abnormal ocular dominance patterns despite lacking binocular overlap (Constantine-Paton et al., 1978). Inspired by this, we manipulated early visual experience in infant macaques to anti-correlate eye activity. Two macaques wore helmets with liquid crystal shutter lenses alternated at 0.1 Hz for the first postnatal year to test if anti-correlated activity induces ocular dominance columns beyond V1, providing insights into feature domain development in higher areas like the inferior temporal cortex. After rearing, one eye was injected with tetrodotoxin for 3 weeks, and ~500 sagittal sections (70 µm thick) were processed for cytochrome oxidase staining. We developed a pipeline to analyze ocular dominance columns in 2D slices and reconstruct 3D volumes of CO staining at different laminar depths (Oishi et al. 2024). Slices were co-registered and combined to construct a 3D histology volume, aligned to a reference MRI, revealing clear ocular dominance columns across the cortical surface extending beyond layer 4 of V1. To quantify column periodicity, visual areas in slices were segmented and flattened into 1D vectors, capturing staining intensity variations. Flattened 1D vectors were assembled into a cortical sheet, revealing the 2D column structure beyond V1. FFT analysis confirmed accurate reconstruction of columns in V1 and their presence beyond V1. The average column periodicity was 506.06±110µm in V1, 550.59±145µm in V2 ventral, and 554.98±137µm in V2 dorsal. The larger domains beyond V1 indicate that 1) Hebbian mechanisms act during development to segregate domains according to patterns of neuronal activity, even beyond primary sensory areas; and 2) hierarchical convergence creates larger domains (Nasr et al., 2016).

Development: Infants, children

Sunday, May 18, 2:45 – 6:45 pm, Pavilion

36.432 THE BUILDING BLOCKS OF VISION: CORTICAL AND SUBCORTICAL ORGANIZATION OF THE NEWBORN VISUAL SYSTEM

Vladislav Ayzenberg[1,2], Michael Arcaro[1]; [1]University of Pennsylvania, [2]Temple University

Understanding the anatomical and functional organization of the visual system at birth provides critical insights into the mechanisms that support early developing perceptual and cognitive abilities. Using resting-state fMRI and diffusion tensor imaging, we examined the cortical and subcortical organization of the visual system in neonates (37-42 weeks gestation). We used an adult probabilistic atlas of retinotopic maps to identify putative visual areas in neonates. Functional correlations between cortical areas revealed that the infant visual system exhibits an adult-like hierarchical organization, with distinct clusters for regions of the occipital cortex, as well as ventral, lateral, and dorsal visual pathways. Direct comparisons between neonates and adults revealed that correlation patterns within occipital and dorsal areas were more adult-like than ventral and lateral areas. We then examined the extent to which this cortical organization was mirrored in the maturity of structural connections between each cortical area and the pulvinar – a subcortical structure that is extensively interconnected with the entire visual cortex in adults and plays a crucial role in visual processing. Probabilistic tractography analyses reliably identified white matter pathways between the pulvinar and each cortical visual area. These connections showed area-level specificity and overlapped with homologous pathways of adults. However, we found developmental differences within the fine-grained connectivity patterns of the pulvinar. Although the coarse connectivity organization for all neonate networks were qualitatively similar to those of adults, the connectivity maps for ventral visual areas were immature and did not show strong specificity within the pulvinar. Altogether, our findings indicate that the large-scale anatomical and functional organization of the visual system is established by birth, but there are developmental differences in the maturity of different pathways with the dorsal pathway maturing earlier than the ventral pathway.

36.433 INFANTS ARE SENSITIVE TO SELF-INDUCED MOTION IN THE ENVIRONMENT

*Jiaye Cai[*1,2] (0919786@zju.edu.cn), Naiqi. G Xiao[*2], Xiaoqing Gao[1]; [1]Zhejiang University, [2]McMaster University [*]Equal Contribution*

Interoceptive signals, such as those from self-body movement, are essential for stabilizing and optimizing sensory processing. These signals, integrated through afferent neural pathways, help differentiate between external and self-generated motion, ensuring stable visual perception. Despite their importance, the developmental origins of this integration in visual processing remain poorly understood. The current study investigated how infants' visual attention could effectively be guided by self-induced movement signals in the environment. Specifically, we developed a novel body motion feedback system allowing us to manipulate the on-screen movement of an object with participants' own head movement to examine infants' sensitivity to self-induced movement that was projected into the external environment. Forty infants (5- to 12-month old) participated in two experiments. Infants watched a visual display of three colorful geometric shapes moving on a computer screen. One shape (Target), moved in synchrony with the infant's head movements. The remaining shapes' movements were driven by head motion data recorded from previous infant participants. Infants' looking time toward the target shape was measured as an index of their attentional sensitivity to self-induced motion cues. Experiment 1 investigated this sensitivity when the movement was presented with no delay. Experiment 2 further explored this phenomenon by introducing temporal delays (0ms, 150ms, and 400ms) between the infants' head movements and the corresponding movement of the target shape. Infants exhibited significant attentional bias toward the target shape in Experiment 1. This bias persisted with 0ms and 150ms delays in Experiment 2 but diminished in early trials of the 400ms delay, increasing in later trials, suggesting the presence of a temporal window for infants' sensitivity to the contingency between self-generated motion and visual feedback. Our finding suggests that infants are sensitive to self-induced motion signals in the environment, revealing an early-emerging mechanism for integrating self-generated actions with visual perception.

36.434 WHAT CAN 100,000 FIXATIONS TELL US ABOUT INFANTS' VISUAL EXPLORATION OF NATURALISTIC SCENES?

Brianna K. Hunter[1], Shannon M. Klotz[1], Steven J. Luck[1], Lisa M. Oakes[1]; [1]University of California, Davis

The ways in which infants visually explore their environment have cascading influences on what they can learn (Oakes, 2023). Therefore, understanding patterns of infants' visual exploration is necessary to understand how they interact with and process the world around them. Visual exploration may vary across development; infants and children initially engage with the world broadly but increasingly use top-down control to focus on more limited information with age (Blanco & Sloutsky, 2024). Additionally, infants' visual exploration may vary within-infants as they adapt their looking behaviors based on the amount of information available for learning (Kidd et al., 2012). In the present study, we analyzed the spatiotemporal dynamics of 100,742 fixations made by 382 4- to 12-month-old infants viewing 243 different naturalistic scenes. We found that across a given trial, fixations progressively shifted outward from the center of the image ($p < .001$). However, this effect was moderated by the overall visual clutter in the image ($p = .041$). Early fixations during scenes with high levels of visual clutter, indexed via subband entropy (Rosenholtz, 2007), were closer to the center of the image compared to early fixations during trials with less clutter. This suggests that infants constrict early visual exploration under circumstances with high levels of perceptual competition. Moreover, the extent of infants' overall exploratory looking, measured as scan path distance—the summed distances between all fixations—varied as a function of both the visual clutter of the scenes and the age of the infants ($p < .001$). Whereas younger infants' showed similar scan path distances regardless of visual clutter, older infants reduced their exploratory looking on scenes with high visual clutter. These results indicate that infants' visual exploration varies with the visual context: in highly cluttered scenes, early exploration is limited to the central visual field and overall exploration is reduced, particularly among older infants.

We greatly appreciate the students and staff in the Infant Cognition Laboratory at the University of California, Davis, for their help with data collection and discussions of this work. This research and preparation of this work was made possible by NIH grants 1F32EY034017 and R01EY030127.

36.435 VISUAL INFORMATION GUIDING HAND REPRESENTATION IN INFANTS: THE ROLE OF CONFIGURAL CUES AND ANATOMICAL COHERENCE

Valentina Silvestri[1], Chiara Dondena[1], Hermann Bulf[1]; [1]University of Milano-Bicocca, Department of Psychology

Beyond faces, hands carry significant social information and are among the most expressive components of the human body, providing insight into actions, gestures, emotions, and intentions. The ontogenetic necessity to tune into hands is revealed by evidence showing infants' notable sensitivity to hands, reflecting their rich visual and tactile experience with this socially meaningful body part (Aslin, 2009; Deák et al., 2014). More recently, it has been shown that infants are sensitive to the structure of hands, preferring scrambled over intact hands (Jubran et al., 2019). However, in the study of Jubran and colleagues (2019), scrambled hands were created by altering the location of one or more fingers within the hand (e.g., the pinky repositioned between the thumb and the index finger), thus breaking the overall gestalt. Here, we explore whether infants' representation of hands is guided by anatomical coherence when the overall gestalt of the hand is maintained intact and the anatomical plausibility is disrupted (i.e., anatomically implausible but configurally recognizable hands). Six-month-old infants were presented with intact and scrambled hands, where all fingers were relocated to positions inconsistent with their natural anatomical origin (e.g., placing a thumb in the position of the pinky) while preserving the overall hand configuration. Infants' eye gaze to the right and left hands, presented in both dorsal and palm view, was automatically recorded using an eye tracker within a preferential-looking task where intact and scrambled hands were simultaneously presented. Preliminary results reveal that infants' preference for scrambled over intact hands depends on hand view (palm vs. dorsal) and laterality (right vs. left hand) ($p = .03$). Our findings shed light on the nature of the visual information underlying infants' sensitivity to hands, advancing our understanding of the mechanisms driving perceptual tuning to hands.

36.436 EMERGING OBJECT REPRESENTATIONS DRIVE COMPETITION FOR LIMITED CAPACITY IN THE INFANT BRAIN

Maeve R. Boylan[1], Anna-Lena Tebbe[1], Jessica Sanches Braga Figueira[1], Andreas Keil[1], Lisa S. Scott[1]; [1]University of Florida

Learning to detect and recognize a broad range of visual objects is a crucial developmental task during the first year of life. However, many of the neurophysiological changes underlying the emergence of this cognitive ability remain poorly understood. The current study tested the hypothesis that training infants to recognize novel objects leads to selectively enhanced visuocortical responses and creates competitive advantages that prioritize the processing of trained relative to untrained objects. Parent-infant dyads at 6- (n = 15), 9- (n = 17), and 12-months of age (n = 19) read books in which novel objects were associated with either (1) individual-level names, (2) a group-level name, or (3) no associated label. The next day, EEG was recorded while infants were concurrently presented with trained objects (i.e., from the book) and untrained objects (i.e., novel objects not in the book). Trained and untrained objects flickered at distinct frequencies (i.e., 5 Hz, 6 Hz) to evoke frequency-tagged steady-state visual evoked potentials (ssVEPs). Analyses of this visuocortical response showed training-related competition effects in the visual cortex increased with age. Specifically, responses to trained stimuli increased while responses to untrained stimuli decreased with age. At 6 months, infants showed no visuocortical bias for trained objects, but by 9 and 12 months, visuocortical responses favored trained objects. This pattern suggests that competitive neural interactions between trained and untrained stimuli may support the development of object recognition and that experience with objects guides attentional prioritization in the infant brain.

This research project was supported by a National Institute of Child Health and Human Development (NICHD; R21HD102715) which was awarded to Lisa S. Scott and Andreas Keil. Maeve R. Boylan was supported by the UF Substance Abuse Training Center in Public Health from NIDA of the NIH (T32DA035167)

36.437 MEASURING PERIPHERAL VISION DURING EARLY INFANCY

Aishwarya Ravi[1] (aishravi@iu.edu), Stephanie A Biehn[1], Fuseina Safianu[1], T Rowan Candy[1]; [1]Indiana University School of Optometry

Purpose: Peripheral vision is key to infants' developing interactions with the world, and in clinical assessment of visual pathology. Previous studies of detection performance have typically focused on the extent of the visual field during infancy. Here we developed an efficient approach to measuring the sensitivity of peripheral vision across the visual field in young infants. Methods: A Maltese cross (3 degrees) was presented on a rear-projection screen (PROPixx projector, VPixx). An EyeLink 1000 Plus (SR Research) was used to record eye movements with high spatial resolution. A 9-point calibration routine was performed before starting the stimulus at the center of the screen. It then moved in an interleaved combination of 5 smooth pursuit movements to demonstrate tracking and 4 saccadic jumps (5, 10, 15, and 5 degrees from fixation) in a continuous trial of less than 2 minutes. The timing of the stimuli was under the experimenter's control and the infant's attention could be redirected using a bell as necessary. Results: Adults and infants (aged two to five months) provided accurate stimulus-driven tracking and saccadic responses to all three eccentricities (5, 10, and 15 degrees). Consistent with the literature, infants' tracking responses demonstrated more saccades than adults' smooth pursuit but were accurate to the tracking stimulus. This, therefore, permitted controlled presentation of the peripheral stimuli at specific retinal eccentricities. The use of the tracking approach significantly improved infants' engagement with the task and confidence in the data. Conclusion: This preliminary study provides a rapid approach for assessing both the peripheral vision of young infants and their ocular motor performance. Modification of the spatio-temporal characteristics of the stimuli will permit evaluation of a wide range of visual functions.

NEI R01 EY014460

36.438 IMAGE STATISTICS AROUND THE POINT OF FIXATION DURING INFANTS' NATURAL ACTIVITIES

T Rowan Candy[1], Jaswanth Boppana[1], Zachary Petroff[1], Philip McAdams[1], Stephanie Biehn[1], Sarah Freeman[1], Bhagya Marella[1], Victoria Tellez[1]; [1]Indiana University

Purpose: Human infants begin to learn about and interact with their environment during the first months after birth. Experience across the visual field during this period contributes to activity-dependent maturation throughout the visual system. The goal of this study was to characterize the distribution of image statistics around the fixation point during infants' head-free natural experience. 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 ≥100ms were identified in the eye movement recordings using a dispersion algorithm, and then RMS luminance and chromatic contrast, color saturation and entropy statistics were averaged around these fixation points in the calibrated scene images for each infant and age group (2-3 (n=25), 5-6(n=31), 8-9(n=29) & 11-12(n=12) months). These statistics were also computed for the same locations in randomly selected images from the same age

group for baseline comparison. Results: A total of more than 2 hours of video were recorded for each age group. Comparison between fixated and randomly paired images indicated that the image statistics all decreased in mean magnitude monotonically with eccentricity from the point of fixation for all age groups, across the analysis radius of 20deg. These distributions were relatively consistent across age despite the dramatic changes in motor behavior in free play activities with age. Conclusions: These data suggest that young infants demonstrate selective fixation of the more visible regions of their natural environment and that fixation behavior becomes more complex over the first months after birth, although not to the extent that might be predicted by their gross motor development. The co-location of the studied image statistics in the scene likely contributes to consistent behavior.

NEI R01 EY032897

36.439 IN CHILDREN, PERIPHERAL CROWDING DEVELOPS TWO YEARS LATER THAN FOVEAL CROWDING

Denis Pelli[1,2], Emma Martindale[3], Najib Majaj[1], Marialuisa Martelli[2], Sarah Waugh[3]; [1]New York University, [2]University of Rome (Sapienza), [3]University of Huddersfield

Waugh et al. (2018) reported the foveal crowding and acuity of 201 students, ages 3 to 11, in a primary school in Cambridge, UK. Crowding distance dropped threefold from age 3 to 8. More recently, in a primary school in Huddersfield UK, we tested 70 students, age 5 to 10, in fovea and periphery (0 and ±5 deg on the horizontal meridian). Over this age range, we replicate our 2018 foveal result, and discover a similar 3-fold drop in peripheral crowding, but 2 years later, from age 5 to 10. Note that the duration of the Sloan-letter triplet used to measure these thresholds was unlimited in the fovea, and only 200 ms in the periphery. We don’t know how much, if any, of the difference in results is due to the difference in stimulus duration. If the periphery’s two-year developmental delay remains with brief presentation in both places, then this represents a factor of three violation of the Bouma law, that crowding distance is linearly related to eccentricity.

URN020-01 from University of Huddersfield to Waugh.

36.440 WEAK ACTION PREDICTIONS IN AUTISM

Mario Costantino[1] (mariolc@yorku.ca), Bat-Sheva Hadad[2], Maryam Vaziri-Pashkam[3], Erez Freud[1]; [1]Department of Psychology, York University, [2]Department of Special Education, University of Haifa, [3]Department of Psychological and Brain Sciences, University of Delaware

Humans are adept at predicting the actions of others by interpreting subtle preparatory movements, a skill crucial for successful social interactions. This study investigated whether autistic individuals, who face challenges in social interactions, exhibit reduced efficiency in predicting others' actions. Additionally, we examined whether autistic individuals demonstrate improved understanding of the actions of other autistic individuals. To address these questions, we used a competitive-reaching task in which an "attacker" was directed by an auditory cue to move toward one of two possible targets. A "blocker" anticipates the attacker’s trajectory to reach the same target as quickly as possible. Using motion-tracking technology, we measured the blockers' finger reaction time (fRT), the time interval between the attacker’s movement initiation and the blocker’s response, as well as the blockers' movement velocity. Results reveal that autistic blockers have longer fRTs and slower velocities compared to non-autistic blockers, indicating a deficit in action prediction. This impairment is consistent regardless of the attacker’s identity, challenging the hypothesis that autistic individuals are better at predicting the actions of others autistic individuals. Importantly, a control experiment measured the blockers' reaction time in response to an auditory cue rather than to another person's action. This experiment revealed no differences between autistic and non-autistic blockers, suggesting that the observed differences in the main task were specific to action prediction. Taken together, these findings provide novel evidence of reduced efficiency in integrating perceptual cues critical for predicting others' actions, offering insights into the mechanisms underlying social interaction challenges in autism.

36.441 SYNTHESIZING EVIDENCE ABOUT DEVELOPMENTAL PATTERNS IN HUMAN VISUAL ACUITY AS MEASURED BY TELLER ACUITY CARDS

Rick Gilmore[1], Julia DiFulvio[1], Brianna Beamer[1], Nicole Cruz[1]; [1]The Pennsylvania State University

Replication is a cornerstone of scientific rigor and a prerequisite for cumulative science. This project synthesized evidence from published research that employed a widely used measure of grating visual acuity (VA), Teller Acuity Cards (TAC). We sought to capture findings about the development of VA in early childhood into an aggregated dataset and share the dataset openly. Online literature searches identified papers that mentioned “teller acuity cards”, “visual acuity cards”, or “teller cards”. We found n=745 papers published from 1974-2024. Next, we identified empirical papers that used TAC to measure VA and which reported VA in an extractable tabular form. To-date, n=250 of 316 papers with available PDF versions have been evaluated and n=14 have been identified that present extractable data meeting our screening criteria. Available datasets represent more than n=3,991 participants and 7 countries (Australia, Brazil, Canada, China, Italy, Mexico, and the U.S.). As expected, group VA increases from birth to 36-months, with faster rates of change among children tested binocularly (0.47 cyc/deg per month) than those tested monocularly (0.35 cyc/deg per month). Group VA values at similar ages vary substantially across studies, especially in children older than 12 months. Our synthesis of published TAC VA data confirms anticipated age-related trends and points to avenues for future research, particularly regarding what factors account for cross-study and by-country differences in rates of development. We hope our soon-to-be openly shared dataset contributes toward a more cumulative science of visual development.

36.442 ACTION VIDEO GAMES AND POSTERIOR PARIETAL CORTEX NEUROMODULATION IMPROVE BOTH ATTENTION AND READING IN ADULTS WITH DEVELOPMENTAL DYSLEXIA

Simone Gori[1], Sara Bertoni[1], Sandro Franceschini[2], Martina Mancarella, Giovanna Puccio[2], Luca Ronconi, Gianluca Marsicano, *Gianluca Campana[2], Andrea Facoetti[2]; [1]University of Bergamo, [2]University of Padua*

The impact of action video games on reading performance has been already demonstrated in individuals with and without neurodevelopmental disorders. The combination of action video games and posterior parietal cortex neuromodulation by a transcranial random noise stimulation could enhance brain plasticity, improving attentional control and reading skills also in adults with developmental dyslexia. In a double blind randomized controlled trial, 20 young adult nonaction video game players with developmental dyslexia were trained for 15 hours with action video games. Half of the participants were stimulated with bilateral transcranial random noise stimulation on the posterior parietal cortex during the action video game training, whereas the others were in the placebo (i.e. sham) condition. Word text reading, pseudowords decoding, and temporal attention (attentional blink), as well as electroencephalographic activity during the attentional blink, were measured before and after the training. The action video game plus transcranial random noise stimulation group showed temporal attention, word text reading, and pseudoword decoding enhancements and P300 amplitude brain potential changes. The enhancement in temporal attention performance was related with the efficiency in pseudoword decoding improvement. Our results demonstrate that the combination of action video game training with parietal neuromodulation increases the efficiency of visual attention deployment, probably reshaping goal-directed and stimulus-driven fronto-parietal attentional networks interplay in young adults with neurodevelopmental conditions.

This work was funded by MUR PRIN Project CUP: F53D23004610006 – ID MUR: 2022772HTJ_01 for S.G.

Eye Movements: Pupillometry

Sunday, May 18, 2:45 – 6:45 pm, Pavilion

36.443 PUPIL FLUCTUATIONS SIGNAL ACTIVE FORGETTING OF NATURAL SCENES

Huiyu Ding[1], Jonathon Whitlock[2], Lili Sahakyan[1]; [1]University of Illinois Urbana-Champaign, [2]Mississippi State University

Forgetting is an adaptive mechanism and the desired cognitive outcome, especially active forgetting, when information is unwanted, irrelevant, or even traumatic. Importantly, successful active forgetting is associated with distinct neural recruitment and viewing patterns compared to passive forgetting. The current study explored the physiological marker for active forgetting by monitoring pupil fluctuation, a measurement linked to long-term memory retrieval, in a modified item-method directed forgetting procedure. The current study explored the physiological markers for active forgetting by monitoring pupil fluctuations in an item-method directed forgetting study. At learning, participants studied naturalistic scene images, each followed by a Remember or Forget instruction. At test, participants were presented with either the identical images or the mirrored version of a studied image and had to determine whether the image was old or new. Active forgetting is defined as when forget-cue images are subsequently forgotten. Passive forgetting is defined as when remember-cued images are subsequently forgotten. The findings revealed that active forgetting elicited greater pupil dilation than passive forgetting. In a follow-up experiment, we strengthened a subset of items by repetition to determine whether the observed pattern in Experiment 1 was driven solely by differences in memory strength. In Experiment 2, we did not observe similar patterns, as weakly encoded images produced pupil fluctuations of similar magnitude to strongly encoded images. Together, these results suggest that the success of intentional forgetting is reflected in pupil fluctuations during memory retrieval, potentially indicating the active cognitive processes engaged by participants in response to forget instructions.

36.444 TYPOGRAPHY AND PUPILOMETRY: EXPLORING THE SPEED, COMPREHENSION, AND CONCENTRATION TRADE-OFF IN READING

Borano Llana[1] (borano.llana@uri.edu), Alisa Baron[1], Kushas Khadka[1], Yusra Suhail[1], Shaun Wallace[1]; [1]University of Rhode Island

Readers constantly balance their ability to read and comprehend vast amounts of information quickly while maintaining concentration over time. Our in-lab study investigates the correlation between these variables using eye trackers and pupillometry, replicating results from our remote readability studies. Recent research shows that increased pupil size and dilation indicate higher concentration. Participants read passages in four versions of Roboto (regular, slab, monotype, and serif) in three sizes (14px, 16px, and 18px). They read one passage and answered five comprehension questions per text design. To measure reading speed, we split each passage into two screens. Our preliminary results feature three participants; we aim to recruit a mix of adult readers with and without ESL and ADHD to discover new relationships between these diverse participants and our variables. Our preliminary findings reveal positive relationships between all variables: reading speed, comprehension, and pupil size across all text designs. Reading speed and mean pupil size per screen were moderately correlated ($r(22) = 0.47$, $p < .03$). We found moderate correlations and borderline significant effects for reading comprehension and mean pupil size ($r(10) = 0.56$, $p = .058$); and reading comprehension and reading speed, ($r(10) = .561$, $p = .058$). We found no relationship between font size and pupil size, indicating initially that small manipulations in size might not affect concentration. We also found no relationship between the font style variations within a typeface and concentration. Thus, pupil size and reading performance might remain stable across everyday font variations and size manipulations. Adding to our prior remote readability studies, participants' pupil size decreased on average 14% throughout the study. Their pupil size also decreased 83% of the time during the second screen per text design. Our research suggests that the reader's concentration is essential to aid future personalization of text designs across populations.

36.445 EYE PUPIL DILATION IN RESPONSE TO SALIENT IMAGES AND CONSTRICTION IN RESPONSE TO PREFERRED IMAGES IN DIFFERENT TIME WINDOWS

Hsin-I Liao[1], Yuta Suzuki[1], Shigeki Nakauchi[2]; [1]NTT Communication Science Laboratories, NTT Corporation, Japan, [2]Department of

Computer Science and Engineering, Toyohashi University of Technology, Japan

The classic literature indicates that the pupil is dilated when seeing attractive and emotionally arousing pictures, such as nude male and female portraits (Hess, 1965; Hess & Polt, 1960), presumably via sympathetic nervous system activation. By contrast, Liao et al. (2020) showed that the pupil is constricted when seeing attractive faces, presumably reflecting a positive loop between seeing and liking (preferring), as pupil constriction improves visual acuity. Emotionally arousing images with high intrinsic salience might surprise experimental participants, in addition to merely attracting them, thus affecting pupil response differently from images that are simply preferred. The current study aimed to explain these seemingly contradictory findings by investigating how pupil response is affected by the salience and attractive aspects of an image. We selected 942 images from the International Affective Picture System (IAPS) and converted them into grayscale with an equal mean luminance. We first conducted an online experiment in which 166 participants rated the salience and attractiveness of each image. Based on the rating results, we identified three sets of images with matched average salience or attractiveness while orthogonally controlling the other rating aspect. We then conducted a pupillometry experiment where another 30 participants viewed each image for eight seconds and rated its salience and attractiveness while their pupillary response was recorded. Results showed that initial pupil constriction at around the 1 s interval was more prominent for the highly attractive images, consistent with Liao et al. (2020). By contrast, late pupil dilation (after 2 s) was larger for high-salience than low-salience images. The results reveal an early pupil constriction response to attractiveness and a late pupil dilation response to salience, presumably underlaid by different mechanisms.

36.446 PUPIL-SIZE FLUCTUATIONS DURING RECOGNITION TESTING REFLECTS A RAPID MATCH SIGNAL

Jonathon Whitlock[1], Ryan Hubbard[2], Lili Sahakyan[3]; [1]Mississippi State University, [2]University at Albany, SUNY, [3]University of Illinois Urbana Champaign

In episodic memory, pupil size increases are greater for correctly recognized old items compared to correctly rejected new items—a phenomenon known as the pupil old-new effect. This effect is often explained by the memory strength account, which posits that stronger memory traces, formed through deeper encoding or memories that involve enriched contextual details, elicit greater pupil dilation. However, challenges to this account arise with the subjective pupil old-new effect, where falsely recognized new items also elicit larger pupil sizes than missed old items. To disentangle these effects, we investigated pupil responses across the entire time course of retrieval during single-item and associative recognition tasks. In a single-item recognition task, participants studied and were tested on faces, providing recognition and confidence judgments. We obtained both the traditional (objective) and the subjective pupil old-new effects and found that confidence judgments dissociated the two effects by moderating the strength of the objective effect but not the subjective effect. Temporal principal component analyses further dissociated these two components across the time course of retrieval, demonstrating the dynamic unfolding of these effects over time. In an associative recognition task, participants studied object-scene pairs and were tested on original pairs, recombined pairs, and novel pairs. Early in the time course, pupil sizes differentiated between original and recombined pairs, despite both involving familiar (studied) information, demonstrating that the early pupil signal reflects the extent to which test stimuli matches the memory representation for studied information. Together, these findings suggest that pupil responses during recognition testing reflects a dynamic rapid match signal between the memory representation formed during learning and test stimuli. By leveraging pupil dilation as a physiological marker, this research advances our understanding of the mechanisms underlying episodic memory retrieval and highlights the complexity of distinguishing between objective and subjective recognition processes.

Eye Movements: Perception, fixational eye movements

SUNDAY, MAY 18, 2:45 – 6:45 PM, PAVILION

36.447 VISUAL MASKING OF AMPLITUDE AND VELOCITY OF SACCADE-LIKE LARGE-FIELD MOTION

Wiebke Nörenberg[1,2] (wiebke.noerenberg@hu-berlin.de), Richard Schweitzer[4], Martin Rolfs[1,2,3]; [1]Humboldt-Universität zu Berlin, [2]Berlin School of Mind and Brain, [3]Cluster of Excellence 'Science of Intelligence', TU Berlin, [4]Centro Interdipartimentale di Mente e Cervello, Università degli studi di Trento

Saccadic omission—the lack of conscious perception of visual information during saccadic eye movements—contributes to perceptual stability despite rapid retinal image shifts. Motion masking, the reduction of perceived motion amplitude through the presentation of static images before and after high-speed motion (Duyck et al., 2018), may support saccadic omission by reducing motion signals. Early visual motion processing relies on speed-sensitive neurons encoding its direction and speed, whereas amplitude perception likely involves both early and higher-order visual pathways. Speed and amplitude may thus be differentially affected by motion masking. We hypothesized that masking reduces both speed and amplitude perception and predicted that it impairs participants' ability to accurately track the decrease in their perceptual performance, reflecting limited metacognitive access. Participants viewed a repetitive background pattern (constructed from 1/f pink noise) moving left or right, with motion varying in amplitude (6, 12, 18 dva), duration (39.2, 55.4, 71.6 ms), and velocity profile (constant vs. saccadic-like). The pattern was identical before and after motion, eliminating endpoint cues. Masking intervals — static presentations of the same background before and after motion — ranged from 0 to 320 ms. In two sessions, participants adjusted on-screen stimuli to report perceived amplitude (using an arrow) or speed (using brief stimulus presentations) respectively, and rated confidence in their judgments on a four-point scale. Both perceived motion amplitude and speed strongly decreased with increasing masking durations across all motion amplitudes. Saccade-like velocity profiles induced masking effects in the absence of explicit masking intervals, suggesting naturalistic motion kinematics contribute to masking in natural vision. Confidence ratings revealed a dissociation between perceptual accuracy and metacognitive access, as participants confidence ratings did not reflect the decline in performance. These results demonstrate that purely visual mechanisms robustly reduce amplitude and velocity perception of saccade-like motion signals, likely supporting perceptual stability during eye movements.

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]).

36.448 CHARACTERISTICS AND POSSIBLE VISUAL CONSEQUENCES OF LENS MOTION DURING EYE MOVEMENTS

Bin Yang[1], Michele Rucci[1]; [1]University of Rochester

While eye movements and their visual functions have been extensively studied, less attention has been paid to the motion of the lens inside the eye. Here, we used digital dual Purkinje imaging to measure lens motion in the human eye and investigate its possible consequences for vision. Our results show that lens motion alternates between saccade-driven wobbles and fixational high-frequency (~75 Hz) oscillations. We observed a decreased oscillation frequency of the lens when subjects (N=5) accommodated to near distances, presumably due to the lower tension in the suspensory ligaments. We also found a strong negative correlation between peak lens oscillation frequency and age (N=19), consistent with the lens losing flexibility and yielding a lower tension with aging. A model of the lens as a damped harmonic oscillator driven by either saccades or fixational eye jitter indicates that lens oscillations amplify fixational motion by approximately a factor of three and that the resonant frequency of lens wobble is about half the fixation frequency (N=12). An optical model of the eye and analysis of data previously collected with an Adaptive Optics Scanning Laser Ophthalmoscope (N=6) confirm that the fixational lens oscillations cause minimal image motion on the retina. However, saccade-driven lens wobbles result in nonnegligible retinal motion, yielding luminance modulations with approximately uniform power across a broad range of spatial frequencies during viewing of natural scenes. In sum, our results suggest that, driven by the unsteady eye, the lens is in constant motion, oscillating at high frequency unless interrupted by saccade-associated sharp wobbles, which affect visual input signals to the retina.

NIH EY18363

36.449 EARLY STAGE EYE-FIXATIONS REVEAL BELIEF-DRIVEN BIAS IN CORRELATION PERCEPTION

Minsuk Chang[1], Adam Malitek[2], Keisuke Fukuda[2,3], Cindy Xiong Bearfield[1]; [1]Georgia Institute of Technology, [2]University of Toronto Mississauga, [3]University of Toronto

The human visual system excels at extracting ensemble statistics, facilitating the interpretation of complex visual information (Alvarez, 2011; Whitney, Haberman, & Sweeny, 2014; Utochkin, Choi, & Chong, 2024). However, this remarkable capability is not immune to bias. Our findings reveal that even seemingly unambiguous visual properties—such as estimating correlations from scatterplots, a task people generally perform with reasonable accuracy (Rensink, 2022)—can be influenced by belief-driven biases (Wolfe & Utochkin, 2019). We conducted an eye-tracking experiment where participants viewed scatterplots depicting meaningful variable pairs (e.g., number of environmental regulations and air quality) and estimated their correlations. They also viewed the same scatterplots with generic axes ('X' and 'Y'). We analyzed the correlation derived from participants' eye fixation points and discovered that it approximated the true correlation, slightly overshooting in the generic baseline condition (MD =0.149, SD = 0.329). For both plausible and implausible variable pairs, gaze-derived correlations also approximated the true correlation but consistently fell below those in the baseline condition, with the implausible condition showing a larger deviation. More interestingly, the dynamic analysis revealed a time-dependent impact of plausibility on gaze-derived correlations. The gaze-derived correlations significantly differed between conditions only in the first two seconds, and they plateaued near the true correlation values across all conditions within five seconds. This suggests that the plausibility of the scatterplot influences eye-gaze patterns most prominently during the initial viewing stages, suggesting that early engagement is critical for detecting belief-driven differences in perception. These results illustrate that prior beliefs can influence the perception of unambiguous visual properties like scatterplot correlation. This phenomenon, a form of belief-driven "motivated perception" (Geisler & Kersten, 2002), underscores the challenges scientists face when presenting data to persuade—our perceptions are often biased by our beliefs, even when viewing objective data points.

NSF IIS-2237585 and IIS-2311575

36.450 THE IMPACT OF VARIABLE CURSOR FEEDBACK DELAY ON VISUOMOTOR TRACKING

Lucy Turner[1] (lucy.turner@adelaide.edu.au), Steven Wiederman[2], Jessica O'Rielly[1], Anna Ma-Wyatt[3,4]; [1]School of Psychology, The University of Adelaide, [2]School of Biomedicine, The University of Adelaide, [3]Andy Thomas Centre for Space Resources, The University of Adelaide, [4]CNRS IRL Crossing, The University of Adelaide

The visuomotor system compensates for latencies associated with sensory feedback to enable people to integrate visual information with motor commands and track moving objects with remarkable precision. However, when latency associated with a movement outcome becomes variable or noisy, this ability to compensate is tested. We aimed to identify the thresholds at which these delays disrupt tracking to understand the limits of visuomotor prediction and error correction. In these studies, participants performed a visually guided pursuit task using a mouse, while their eye and cursor movements were recorded under various feedback latency conditions. In the first study, cursor latency was constant across each trial. We systematically tested each of six latency magnitudes (0ms, 100ms, 150ms, 200ms, 250ms and 300ms) and analysed its impact on tracking performance (error and velocity fluctuations). Spectral analysis of cursor velocity showed that pursuit sub movements shifted to lower frequencies interspersed with sporadic high-frequency corrective actions for high-latency conditions. In contrast, smoother, high-frequency tracking dominated under low-latency conditions. In the second study, abrupt changes in cursor latency were introduced mid-trial to examine how sudden variability can influence visuomotor adaptation. Increases in latency resulted in larger positional errors and upregulated corrective sub movements to realign the cursor with the target. Conversely, decreases in latency led to a rapid reduction in positional error and a return to smoother tracking. These rapid transitions in movement highlight the visuomotor system's ability to recalibrate quickly in response to variability in latency. Together these results show that prediction and error correction mechanisms can differentially respond to changes in latency and this response is dependent on the magnitude of the latency. These results will help inform the design of assistive display technologies that can maintain performance under challenging conditions and increase safety in high-risk environments.

36.451 ON SEMI-ON AXIS DPI + PUPIL SIZE MEASUREMENT

Boris Greenberg[1], Ward van der Tempel; [1]VoxelSensors

Classical Dual Purkinje Imaging (DPI) faces two major challenges: lateral alignment of the measurement system and accurate detection of P1 and P4 for scale and game factor. Traditionally, these systems rely on a lateral translation mechanism that requires careful alignment and depends on the initial gaze direction. Additionally, P1 and P4 detection relies on maintaining a small angle for proper optical alignment, typically achieved with two separate channels for illumination and signal collection. We propose a novel approach to a DPI system that addresses both challenges. Firstly, our system uses a novel Single Photon Active Event Sensor (SPAES) to simultaneously detect P1 and P4 relative positions on the same matrix with less than a 5-degree angle on the axis, improving accuracy and alignment. Then, to compensate for lateral device misalignment, we use an array of light sources that cover various users' interpupillary distances (IPD). Additionally, we leverage residual retinal scattering to back-illuminate the pupil and measure its dimensions whenever required. This can be done as a separate process with significantly lower sampling rates since pupil dynamics are slower than eye movements (saccades and fixation). We use a MEMS-based scanning element that tracks gaze direction, and our initial gain and offset calibration process allows us to define the system's operating parameters. Therefore, we can use the control response function to determine the eye position in 5D coordinate space. For this purpose, we can periodically resonate the MEMS and determine the deviation from the initial parameters. The same array used for lateral compensation can be used for slippage compensation of the device as any deviation in Z-distance manifests itself in a combination of lateral and angular motion.

36.452 THE INFLUENCE OF PHOTOLUMINESCENT ROAD MARKINGS ON DRIVER'S STATIONARY GAZE ENTROPY: A DRIVING SIMULATION STUDY

Leandro L. Di Stasi[1,2] (distasi@ugr.es), Francesco Angioi[2], Juan de Oña[3], Rocío de Oña[3], Carolina Diaz-Piedra[1]; [1]Mind, Brain, and Behavior Research Center, University of Granada, Spain, [2]Joint Center University of Granada - Spanish Army Training and Doctrine Command, Spain, [3]TRYSE Research Group, University of Granada, Spain

Background: Nighttime driving conditions can make it more challenging for drivers to accurately predict the upcoming road geometry and effectively control their vehicles, especially on curves. While standard retroreflective road markings have limited success in enhancing overall roadway visibility, new photoluminescent (or glow-in-the-dark) road markings (PRMs) might overcome these limitations. However, despite PRMs already being tested on conventional roads, no studies have investigated their influence on driver visual exploration strategies. Here, we examined how different PRM colors (white (unlit) vs. green vs. red), widths (conventional vs. wide), and curve directions (left vs. right) affect stationary gaze entropy during virtual driving sessions. Methods: Thirty adult drivers (mean age ± SD = 25.90 ± 6.73 years; 16 women) drove a semi-dynamic driving simulator for, approximately, one hour Participants underwent six different virtual scenarios (the order was balanced across participants). We recorded stationary gaze entropy using a 4-camera remote eye-tracker (120Hz). Additionally, we collected driving performance data and subjective ratings of complexity. Results: Stationary gaze entropy was influenced by both the PRM color and the curve direction, but not by the width of the marking. Interaction effects were not significant. Drivers exhibited lower gaze entropy on right-hand curves, indicating less exploration or scanning behavior compared to left-hand ones. Post hoc comparisons showed that green PRMs led to more stereotyped (i.e., less random) visual exploration compared to the unlit markings. Red PRMs did not have such effect. These gaze entropy results were corroborated by performance data. Perceived complexity was similar across the scenarios, however. Conclusions: Our results, consistent with the concept of self-explaining roads, confirm the benefits of introducing PRMs as a complementary passive measure to prevent roadway departure events. Furthermore, PRMs appear to be a potentially valuable visual guidance technology for enhancing road safety in low-visibility conditions by increasing driver situational awareness.

This work is part of the Research Project SARAH (PID2022-141935OB-I00; MICIU/AEI/10.13039/501100011033 & ERDF/UE). FA is supported by a Spanish national pre-doctoral FPI research fellowship (PREP2022-000307; MICIU/AEI/10.13039/501100011033 & ESF Investing in your future).

36.453 THE OPTIMAL RETINAL LOCUS FOR HIGH-RESOLUTION VISION IN SPACE AND TIME

Martin S. Banks[1] (martybanks@berkeley.edu), Josselin Gautier[2], Norick R. Bowers[3], Austin Roorda[4]; [1]UC Berkeley, [2]University of Renne, [3]Justus Lubig University, [4]UC Berkeley

The anatomical region that humans use to fixate and project a target onto the retina is the preferred retinal locus (PRL). Although the PRL is constant across task and time, we do not know whether and how fixations on the PRL, as opposed to nearby locations such as the foveola, affect visual performance. To investigate this, we employed the retinal-image-based eye tracking, and stimulus delivery capabilities of an adaptive-optics scanning-light ophthalmoscope (AOSLO) to measure how stimulus location on the retina -- the where -- and oculomotor timing and behavior -- the when and how -- affect visual acuity. Participants performed a Vernier-acuity task by reporting the direction of offset between two small horizontal bars. Stimuli were flashed very briefly every 2sec in a cadence so participants could adopt a consistent fixation strategy. As the time of presentation approached, participants made microsaccades to position the upcoming target near the PRL, and they started to make fewer microsaccades as stimulus presentation was imminent. When the stimulus fell on or near the PRL, acuity was greater than when it fell on other locations. A few hundred milliseconds after the stimulus was extinguished, drifts and microsaccades reoccurred moving the eye away from where the stimulus had been. Our results show that people exhibit machine-like fixation behavior in space and time to position the upcoming stimulus on a specific retinal location: the PRL. But the results also show that performance is not measurably better or worse in the region of peak cone density, which is roughly 5minarc away.

NIH R01EY023591; Center for Innovation in Vision & Optics, UC Berkeley

36.454 A REVIEW OF FIXATIONAL INSTABILITY IN OPHTHALMIC DISEASE

Sophia Castro[1], Nancy O. Mahfouz[2], Doris Prela[2], Ayesha Mulla[2], Simrat Renu[2], Robert G. Alexander[2]; [1]CUNY School of Medicine, [2]Department of Psychology & Counseling, New York Institute of Technology, USA

Fixational eye movements—microsaccades, drifts, and tremors—play a vital role in maintaining visual stability and supporting normal vision. These small, involuntary motions are essential for everyday visual function. However, disruptions in the dynamics of these movements are often linked to visual impairments associated with various ophthalmic disorders. This review explores the latest research on fixational stability and its connection to disease, emphasizing advancements in eye-tracking technologies' emerging diagnostic and therapeutic uses. Methods: We conducted a thorough literature review of articles exploring the relationship between fixational eye movements and ophthalmologic conditions. We identified articles through keyword searches and examined references cited by relevant publications. Results: Substantial evidence suggests that many ophthalmic conditions influence fixational behavior in specific ways, which may provide new insights into disease mechanisms and progression. Behçet uveitis, cataracts, amblyopia, macular disease, glaucoma, and other conditions have unique fixational eye movement characteristics—which also sometimes vary as a function of how the condition progresses. For example, different muscular dystrophies have varying effects on fixational stability (with less stable fixation in Stargardt disease than in the best vitelliform macular dystrophy). Fixational stability is also impacted differently by abrupt vs gradual progressions of scotomas. Conclusion: Fixational eye movement analysis can aid us with diagnosis, treatment efficacy and ultimately guide clinical intervention. Eye-tracking tools can even be integrated into routine telehealth platforms. The development of accessible and non-invasive eye-tracking tools holds promise for enhancing personalized patient care.

This research did not receive any specific grant from funding agencies.

36.455 HUMAN EYE MOVEMENTS HAVE A DOWNWARD DRIFT BIAS DURING VISUAL EXPLORATION

Stephanie Reeves[1], Jorge Otero-Millan[1,2]; [1]Herbert Wertheim School of Optometry & Vision Science, University of California Berkeley, [2]Department of Neurology, Johns Hopkins University

Humans make eye movements to explore the world around them. While saccades occur in all directions, horizontal saccades are more common than vertical saccades, and upward saccades more than downward saccades. We hypothesize that the higher frequency of upward saccades could be related to a downward bias in fixational drift, as previously shown for fixational eye movements. We also hypothesize that drift velocity depends on eye position by drawing the eye back to primary position. We analyzed four publicly available datasets (Reeves 2023, Reeves 2022, Chen 2021, and van der Linde 2009) to quantify the velocity of ocular drift as a function of eye position. After processing the sample data, there were 510,491 fixation periods across 73 subjects (more than 1,000 fixations per subject). For each subject, we used linear regression to measure the influence of horizontal eye position on horizontal drift velocity and vertical eye position on vertical drift velocity, respectively. The intercept served as a measure of overall bias and the slope as a measure of the effect of eye position on drift. Results revealed two main findings. First, the average slopes for both horizontal and vertical components were negative (horizontal: -0.013°/s per degree, $p<0.001$; vertical: -0.012°/s per degree, $p<0.001$), indicating that the eye tends to drift back to primary position during eccentric free viewing. Second, the average intercepts revealed an overall downward drift bias of -0.15°/s such that regardless of where participants looked vertically, the eye tended to drift down ($p<0.001$). There was also a rightward drift bias (0.08°/s, $p<0.001$) that was about 50% of the magnitude of the downward drift bias and only present in one of the four datasets ($p>0.39$). Taken together, these preliminary results show that the ocular motor system produces drift during fixation that is not completely random and is instead systematically biased.

R00 EY027846

36.456 SPATIAL ENCODING VIA VISUOMOTOR INTEGRATION

Y. Howard Li[1], Jonathan D. Victor[2], Michele Rucci[1]; [1]University of Rochester, [2]Weill Cornell Medical College

It is commonly assumed that the visual system encodes spatial information by the relative locations of neuronal receptive fields, beginning in the retina. The eyes, however, are never stationary; even when attending to a single point, incessant small movements (fixational drifts) also occur between saccades. A growing body of evidence indicates that spatial information is also extracted from the luminance modulations caused by eye movements. Furthermore, recent evidence suggests that the visual system can infer spatial relations from precise extraretinal knowledge of how the eye moves during fixation. How are these different sources of spatial information integrated? To investigate this question, here we decoupled eye movements from their normal luminance modulations via high-resolution eye-tracking and gaze-contingent display control. In a 2AFC orientation discrimination task, subjects were asked to report the orientation of a static grating (±45-deg) while maintaining fixation. Stimuli were presented with a directional inconsistency between fixational drifts and the expected temporal modulations. This was accomplished by moving the stimulus on the display according to eye movements, so that, on the retina, the stimulus moved at every moment as if the eye had drifted in a different (rotated) direction, but with the same velocity. All subjects' performance was greatly impaired when the retinal motion and the eye trajectory differed by 90o, a manipulation that drastically alters the relationship between grating orientation and the spatiotemporal structure of the luminance fluctuations resulting from eye movements. In contrast, performance recovered approximately to normal when retinal and drift trajectories

were rotated by 180o, a condition that restores consistency between luminance fluctuations resulting from eye movements and grating orientation. These results suggest that knowledge of drift direction provides critical spatial information, and a violation of the contingency between the expected and the actual retinal luminance flows is detrimental to visual processing.

This work was supported by National Institutes of Health grants EY018363, EY07977 and P30 EY001319.

36.457 FIXATIONAL EYE DRIFT ADAPTS TO THE HISTORY OF EYE POSITIONS

Terence L. Tyson[1,2] (terencetyson@berkeley.edu), Leland S. Stone[2], Jorge Otero-Millan[1,3]; [1]University of California, Berkeley, [2]NASA Ames Research Center, [3]Johns Hopkins University

Oculomotor behavior is a powerful tool to probe brain function and disease. Previous studies have shown how fixational drift velocities depend on eccentricity (Bertolini et al., 2013; Romano et al., 2017) and how fixational drift can adapt after prolonged eccentric fixation (Otero-Millan et al., 2019). Here, eighteen subjects completed two experiments where we measured drift velocity while fixating a briefly flashed target, to remove retinal slip cues. In the first experiment, we measured their baseline relationship between fixational drift velocity and eye position. In the second experiment, we investigated the adaptation of that relationship after prolonged fixation at a very eccentric position. This adaptation usually manifests by the presence of rebound nystagmus (RN) upon returning to central gaze. Our results replicate the finding that drift increases at central fixation in the direction consistent with rebound nystagmus, that is, towards the previously held eccentric position, highlighting that drift velocity depends on the history of eye positions. When examining the entire relationship between drift velocity and eye position we find a mostly linear relationship but with increased slope relative to the baseline condition ($p < 0.00003$). This slope (leftward gaze holding: -0.017 +/- 0.002 deg/sec/deg; rightward gaze holding: -0.014 +/- 0.002 deg/sec/deg) reflects a possible increase in the leak rate of the gaze holding neural integrator (leftward gaze holding time constant: 59 sec; rightward gaze holding time constant: 69 sec). However, the adaptation of the integrator cannot be modeled by a simple change in leakiness, as that would not result in increased drift at central gaze. Other mechanisms of adaptation such as shifts in the position or velocity set-points for gaze holding could be present but are indistinguishable with the current data.

This work was supported by the NASA Ames Graduate Fellowship Program.

36.458 **PATHOLOGISTS' ROUTINE FIXATIONS CAN BE** USED TO SUPERVISE LYMPH NODE DEEP LEARNING MODELS

Meng Ling[1], Veronica Thai[1], Shuning Jiang[1], Rui Li[1], Jeremy Wolfe[2], Wei-Lun Chao[1], Yan Hu[1], Anil Parwani[1], Raghu Machiraju[1], Srinivasan Parthsarathy[1], Zaibo Li[1], Jian Chen[1]; [1]Ohio State University, [2]Harvard University

Locating cancerous tissues from large high-spatial resolution whole-slide images (WSIs) is hindered by a lack of training data to supervise deep convolutional neural network (DCNN) algorithms. Patch-based human annotation is time- and labor-intensive. Additionally, training DCNNs would be improved by seeing data from diverse stimuli from routine clinical settings. Expert pathologists are trained professionals who know where to look to locate cancerous tissue from giga-pixel WSIs. Thus, we could acquire theoretically unlimited training samples by harvesting pathologists' routine examinations of WSIs. To validate the reliability of this idea, we collected eye-tracking data from 10 pathologists, each viewing 60 slides from the CAMELYON16 dataset. These data were entered into DeepPFNet: our automated human-intelligence-based data preparation pipeline used to supervise AI to identify tumors. Specifically, we computed a pathologist's fixation-map (PFMap) over each WSI and trained a DCNN using tumor tiles sampled from these maps and benign tiles sampled from benign slides' tissue area. We validated DeepPFNet in experiments that examined effectiveness and scalability. Our experiments show that: models trained using DeepPFNet can achieve accuracy significantly higher than random sampling (F1 = 0.84, AUC = 0.91), and increasing the number of slides sampled leads to significant improvement (ΔF1 = 0.08, ΔAUC = 0.13). DeepPFNet models have better accuracy (F1 = 0.84, AUC = 0.93) than those using clustering (F1 = 0.70, AUC = 0.82) or viewport (F1 = 0.68, AUC = 0.88) approaches. We used the DeepPFNet model to classify tiles from the training WSIs and expanded the sampling maps, significantly improving the pipeline (ΔF1 = 0.04, ΔAUC = 0.03). Fixation-based fine-tuning of weakly supervised learning improved slide-level classification accuracy (ΔF1 = 0.01, ΔAUC = 0.01). Finally, applying PFMap on benign slides to sample benign tiles can improve sensitivity (Δsensitivity = 0.04), but decrease accuracy (ΔF1 = -0.06, ΔAUC = -0.03).

OSU Translational Data Analytics Institute (TDAI) Research Pilot Award

Motion: Local, higher-order, in-depth

Sunday, May 18, 2:45 – 6:45 pm, Pavilion

36.459 GLOBAL STIMULUS CONFIGURATION MODULATES BOLD RESPONSES TO VISUAL MOTION IN HUMAN V1 AND MT

Rania Ezzo[1,2] (rje257@nyu.edu), Marisa Carrasco[2], Jonathan Winawer[2], Bas Rokers[1,2]; [1]New York University Abu Dhabi, [2]New York University

Background: Discrimination of visual motion is direction-dependent: thresholds are lower for cardinal than oblique and for radial than tangential motion (Ezzo et al., 2024, JoV). Using fMRI, we asked whether there are corresponding asymmetries in BOLD amplitude, and whether they are amplified or reduced when local direction matches the global stimulus configuration. Methods: We recorded fMRI BOLD responses to stationary and moving gratings (12º radius) within two global configurations: Cartesian (horizontal, vertical, and left/right-leaning obliques) and Polar (pinwheel, annular, and clockwise/counterclockwise spirals). For each global configuration, responses were measured for the 4 stationary patterns and the 8 motion directions perpendicular to the orientations. Stimuli were

approximately matched locally across the two configurations; e.g., for a receptive field centered at +6º on the vertical meridian, the rightward moving vertical grating and clockwise rotating pinwheel matched locally. Thus we could assess the two local motion asymmetries, cardinal vs. oblique and radial vs. tangential, in each global configuration. Using retinotopy, we defined population receptive fields in V1 and MT ("hMT+"), and analyzed data within a 4–8º eccentricity ring. Motion-specific responses were defined as BOLD amplitude for the difference between the moving and matched stationary patterns. Cardinal vs oblique motion: In the Cartesian configuration, BOLD responses to cardinal motion were larger than to oblique motion, by 10% in V1 and 15% in MT. In the Polar configuration, this effect diminished in V1 and approached zero in MT. Radial vs tangential motion: In the Polar configuration, radial motion evoked greater responses than tangential motion, 104% larger in V1 and 24% in MT. In the Cartesian configuration, this radial > tangential effect weakened in V1 and disappeared in MT. Conclusion: Neural responses to motion vary with direction, and these asymmetries are amplified when the local directions are congruent with the global configuration.

(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

36.460 THE EFFECTS OF AUTOMATION LEVEL AND TAKEOVER LEAD TIME ON TIME-TO-COLLISION JUDGMENTS

Evelyn Wilson[1], Quan Lei[2]; [1]Wichita State University, [2]Wichita State University

Automation has been widely used in all sectors of human activities yet how automation affects basic visual processing involved in human performance is understudied. The current study examined how the level of automation and the takeover lead time when automation fails affect time-to-collision (TTC) judgments in a simplified scenario simulating human-automation interaction. Participants (n=17) watched a series of two objects (2°) moving horizontally on a screen that either collided or passed each other. In the high automation condition, automation was implemented on 2/3 of the trials where participants passively watched the objects moving with no response required, while on the remaining 1/3 of trials, motion of the objects was occluded midway, signaling the failure of automation, and participants had to take over and make a TTC judgment. In the low automation condition, automation was implemented on 1/3 of the trials while participants had to take over on the other 2/3 of trials. There was also a no-automation condition where participants had to respond on all trials. To manipulate takeover lead time, objects moved at either a faster or a slower speed. TTC judgment accuracy was measured as the deviation of estimated TTC from the actual TTC. Unlike the typical underestimation reported in previous research involving collisions with the observer, TTC judgments in this study were overestimated overall. There was a significant effect of speed (or takeover lead time) on TTC accuracy, the judgments being less accurate with shorter lead time. However, no effect of automation level was found nor any interaction between automation level and takeover lead time, despite a non-significant trend for TTC accuracy to improve with higher automation level. Our results suggest that takeover lead time plays a major role in TTC judgment in scenarios involving human-automation interaction.

36.461 DOES PHYSICAL EXERCISE OR CAFFEINE MODULATE DORSAL AND VENTRAL STREAM PROCESSING?

Priyanka Roy[1] (proy7@uic.edu), Amy Nguyen[2,3], Nicholas Vo[2,4], Arijit Chakraborty[2,5,6]; [1]University of Illinois Chicago, [2]Midwestern University, [3]Pacific University, [4]Levine Eye Care Center, [5]University of Waterloo, [6]Rush University Medical Center

Everyday activities such as physical exercise and drinking caffeinated beverages are known to modulate neuroplasticity, although the interplay between the two is unclear for central nervous system functions, such as visual perception. Previously, we reported (Nguyen et al., 2023) that caffeine selectively enhances dorsal stream function (global motion perception) but does not alter ventral stream function (global form perception) or cortical excitability. This study investigated the interaction between physical exercise and caffeine on dorsal and ventral visual stream functions. Thirty two participants (26±5y) biked for 120 mins at 60% of their maximal heart rate. Participants were randomly assigned to receive either 200 mg of caffeine or a placebo pill an hour through the physical exercise. Motion coherence threshold (MCT) for random-dot-kinematogram and form coherence threshold (FCT) for Glass pattern stimuli were measured sequentially and concurrently within 2-down-1-up adaptive staircases, pre- and post-physical. There was a main effect ($F_{1,31}=11.571$, $p=0.03$) of exercise-induced improvement that was modulated by caffeine and the testing method (sequence/concurrent). When measured in sequence, post-exercise, the MCT was improved in both caffeine (pre: 22±7%, post: 11±8%, p=0.01) and non-caffeine (pre: 20±8%, post: 16±8%, p=0.035) groups. There was no change in FCT in either group. When measured concurrently, both MCT (pre: 27±11%, post: 15±7%, p=0.001) and FCT (pre: 31±12%, post: 25±11%, p = 0.02) exhibited a larger improvement in the caffeine group and a slight improvement (MCT, pre: 25±9%, post: 19±7%, p = 0.04; and FCT, pre: 29±10%, post: 25±11%, p=0.043) in the non-caffeine group. Overall, our results indicate an isolated and cumulative beneficial effect of physical exercise and caffeine on dorsal visual stream function. When measured concurrently with the dorsal stream, the additional positive impact on the ventral stream might suggest an influence of the dorsal stream in global form perception through cross-talks between the two streams.

36.462 SEEING THE OPPOSITE DIRECTION: A SEX DIFFERENCE IN GLOBAL MOTION PERCEPTION

Quan Lei[1] (quan.lei@wichita.edu), Evelyn Wilson[1]; [1]Wichita State University

Previous research reported a sex difference in visual motion processing, with males generally having lower thresholds than females (Murray et al., 2018). Viewing random dot kinematograms (RDK), participants were found to show a tendency to perceive the opposite direction of global motion (Bae & Luck, 2022). The current study investigated if a sex difference exists in perceiving the opposite direction of global motion and if such difference depends on central vs. peripheral viewing. 14 female and 14 male participants made judgements about the global motion direction of RDK stimuli consisting of 120 dots moving in a 10° circular field (contrast: 60%, speed: 6°/s, lifetime: 50 ms, duration: 300 ms). There were four directions: left, right, up and down, and seven coherence levels varying between 5% and 95%. The stimuli were either viewed centrally at the fixation or peripherally at 10° eccentricity to the left or right of fixation. A sex difference of similar size was found in direction discrimination accuracy for both central and peripheral viewing, with males being more accurate than females ($\eta 2$ =.198). Inspecting the confusion matrix, there was a strong tendency for participants of both sexes to report the opposite direction (compared with orthogonal directions) even at high coherence levels. However, the tendency was significantly stronger for females, such that if motion axis (instead of direction) was considered, there was no longer a sex difference in accuracy. Additionally, with peripheral viewing, females showed a pronounced tendency to perceive downward motion for stimuli moving along the vertical axis, no such asymmetry existing for males. These results indicate that the sex difference in direction judgement of global motion can be entirely explained by a differential process of disambiguating the two opposite directions along the same motion axis, presumably reflecting a sex difference in spatiotemporal filtering of motion signals.

36.463 SPATIAL SUPPRESSION OF MOTION AND MOTION SEGMENTATION IN PERIPHERAL VISION

Pooja Nandagopal[1], Andrew J Anderson[1], Allison M Mckendrick[1,2]; [1]The University of Melbourne, [2]Lions Eye Institute, University of Western Australia

Motion segmentation (distinguishing a moving foreground object from its background) is thought to benefit from spatial suppression of large, high-contrast backgrounds, and has been studied in central vision. This study examined motion segmentation and spatial suppression of motion in central and peripheral vision, their correlation, and whether scaling peripheral stimuli based on cortical magnification eliminates perceptual differences between eccentricities. Ten healthy adults (mean 27.5 years, range:19–35) completed computer-based tasks at five contrast levels (12%–92%), three eccentricities: (0°, 10°, 20°) and two stimulus conditions (scaled and unscaled). In the motion segmentation task, participants identified the tilt (45° right/left) of a motion defined black-and-white textured ellipse within an oppositely moving black-and-white textured background. Segmentation threshold was the minimum stimulus exposure duration for accurate tilt discrimination. To measure motion segmentation efficiency, participants were shown the ellipse in isolation and asked to identify its motion direction (up or down). Motion segmentation efficiency was then calculated as the difference in log10 threshold values: log10(ellipse motion threshold) – log10(motion segmentation threshold). Suppression strength was measured requiring participants to identify the motion direction of textured background patch (radius = 5.3°) presented alone. Duration thresholds were measured, with suppression index calculated as the difference between log10 duration thresholds of the highest and lowest contrast stimuli for each participant. In the unscaled condition, motion segmentation thresholds increased (RM-ANOVA, main effect of eccentricity, $p<0.01$) and suppression index ($p=0.007$) decreased with eccentricity. The correlation between segmentation efficiency and suppression index weakened systematically (Pearson's $r = -0.87$ at 0°, $r = -0.54$ at 10°, $r = -0.33$ at 20°). Scaling reduced segmentation thresholds (RM-ANOVA, main effect of scaling, $p<0.01$) and increased the suppression index ($p<0.01$) but did not eliminate regional differences. Therefore, our ability to segment moving foreground objects deceases with eccentricity, even when the stimulus in the periphery is enlarged.

36.464 PERCEIVING ANIMACY 'RIGHT NOW': ONLINE PERCEPTION AS A CRITICAL FOUNDATION FOR REAL-TIME AGENCY UNDERSTANDING

Ning Tang[1] (ningtangcog@gmail.com), Enjie Xu[2], Jifan Zhou[2], Mowei Shen[2], Tao Gao[3,4,5]; [1]Department of Psychology, Soochow University, [2]Department of Psychology and Behavioral Sciences, Zhejiang University, [3]Department of Communication, UCLA, [4]Department of Psychology, UCLA, [5]Department of Statistics, UCLA

The Heider-Simmel display illustrates how we transition from low-level motion of geometric shapes to high-level social storytelling, engaging both perception and cognition. While it is accepted that perception serves as the visual root for social inference, it remains unclear which aspects of Heider-Simmel experience are attributed solely to perception. In this study, we propose a new paradigm to investigate the online perceptual component of this experience, inspired by research on time perception and working memory. Studies show that our experience of "the present" is not infinitesimally small, but rather a temporal window of about 3 seconds, which corresponds to the duration of working memory. This window also aligns with the time for generating and understanding intentional actions, such as waving to a friend. We hypothesize that perception and cognition can be dissociated by slowing down the display to extend critical motion beyond the 3-second window. Disrupting this window would likely impair online perception, while offline cognition would be more immune to it. In two experiments, we adapted two types of visual displays: one resembling the Heider-Simmel display, described by action verbs like 'stalk' and 'argue,' heavily relying on perceptual processing; the other using a food truck task, where participants engaged in a sophisticated inference about the agent's desire preference among three trucks, requiring more cognitive reasoning. In both experiments, participants focused on a visual search task while viewing display clips at normal speed or 0.5x speed. Afterward, they were asked a surprise question to interpret the social interaction or preference. Results showed a drop in accuracy in the action verb task at slower speeds, but no difference in the desire preference task. Our study demonstrates that agency understanding is more closely tied to perception when captured through verb semantics, while inferring desire preference may rely more on domain-general cognitive inference.

This work was performed with the support of grants from Natural Science Foundation of Jiangsu Province (BK20240791 to Ning Tang)

36.465 STUDYING MOTION PERCEPTION IN THE REAL WORLD WITH AUGMENTED REALITY HEADSETS

Anthony LoPrete[1] (aloprete@seas.upenn.edu), Johannes Burge[1]; [1]University of Pennsylvania

Vision research has historically been dominated by experiments presented with flat-panel displays. The recent development of a large number of advanced and commercially available display technologies has widened the scope of possibilities. Here, we present our experimental development efforts on the Magic Leap 2 Augmented Reality (AR) headset. The initial aim is to investigate the Pulfrich effect

in the rich visual environment produced by real-world scenes. The Pulfrich effect is a visual illusion wherein an interocular discrepancy in processing speed (commonly introduced with a darkening filter) results in depth misperceptions of moving targets. We have developed and conducted trial-randomized experiments via the Unity game engine. The stimulus is a vertical bar moving laterally (X-motion only) across the observer's field of view alongside a stationary reference bar with the same physical dimensions. The bars were displayed at maximum luminance (i.e. Unity's maximum machine-unit values) to minimize light bleed-through from the real 3D scene. Observers are then tasked with judging whether the moving bar was "nearer" or "farther" than the stationary reference. Responses were recorded with the Magic Leap 2 controller. Recorded data was retrieved from the headset using the Magic Leap Hub's device bridge, and analyzed using Matlab. The psychophysical results are consistent with previous findings. Future development will enable experiments that incorporate the Magic Leap's ability to graphically render objects consistent with the depth structure of the 3D environment, and manipulate key viewing variables like the visibility of contact points with the ground. Depth percepts in real physical units will be recorded by indicating the perceived distance of the target with the Magic Leap controller. The results of our experiments and development progress suggest that the Magic Leap 2 will be a useful experimental tool for studying the human visual system.

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. and an Ashton Fellowship from the University of Pennsylvania to A.L.

36.466 CONSCIOUS PERCEPTION TRAVELS WITH EARLY VISUAL PATHS

Caroline Myers[1] (cmyers60@jhu.edu), Peter Mazalik[1], Jonathan Flombaum[1], Justin Halberda[1]; [1]Johns Hopkins University

A staggering example of apparent motion is the Ternus illusion: when two discs are shown side by side, and then shifted horizontally, the temporal gap between the successive presentations determines whether observers perceive element-wise or group motion. At long (20ms) intervals, group motion is perceived; at brief (0ms) intervals, element-wise motion is perceived. Will an object-file travel with our conscious percept, or will assignment prioritize spatiotemporal proximity? To test these possibilities, we exploited a well-known paradigm – the object-specific preview benefit (OSPB) – in which "previewing" information presented on one object results in faster and more accurate retrieval if the information later reappears on the "same" object relative to another object. If moving objects are defined by spatiotemporal proximity, an OSPB should not follow an object that appears to "jump" when element-wise motion is perceived. However, if the object file system tracks conscious perception, we should expect to see benefits at the "perceived" final position of the object. Observers viewed an apparent motion display consisting of three horizontally adjacent discs containing letters. Following a variable blank inter-stimulus interval, (0ms or 20ms), the discs reappeared shifted one position rightward, with a probe letter reappearing on either the same or different discs. Observers reported whether the probe contained the same letter shown at the beginning of the trial. Strikingly, observers' reaction times and accuracy reflected a characteristic OSPB in accordance with the induced percept: participants retrieved the feature faster and more accurately both when the probed object appeared to translate (20ms intervals), as well as when it appeared to jump (0ms intervals). These results suggest that object identity is mapped across the perceived object motion path induced by the inter-stimulus interval, even if the induced motion path is not the shortest. Thus, in the Ternus display, object files mirror our conscious percepts.

36.467 INVESTIGATING NON-RIGID MOTION PERCEPTION WITH MULTISENSORY INPUTS: A PILOT STUDY

Danica Barron[1], Erin Conway, Troy Smith, Ralph Hale; [1]University of North Georgia

Koerfer et al. (2024) found that non-rigid motion perception generates a motion signal that can be utilized by selective visual systems for perception and saccadic tracking but not for prediction or smooth pursuit. Our pilot study builds on this prior work to explore the impact of multisensory inputs on the perceptual stability of non-rigid motion. Using their novel, dynamic vortex stimulus, we examined whether auditory feedback influences visual perception and pursuit, as it does with rigid motion. The study involved undergraduate students who completed two counterbalanced blocks: a visual-only condition and a visual-audio condition. The vortex was composed of 8000 white dots on a black background displayed at a visual angle of 97.5° (width) x 55.9° (height). The vortex rotated at 45°/s, driven by an underlying mathematical model that simulated fluid-like motion. The stimulus was designed without distinct predictive cues, such as a clear leading edge, consistent directional markers, or static reference points. In the visual-audio condition, auditory feedback was delivered through noise-canceling headphones, with a panning, low-pitched tone synchronized to the vortex's motion, designed to add spatial depth cues. Preliminary results indicate that auditory feedback may impact perceptual stability under complex motion conditions, though smooth pursuit mechanisms remain unengaged. This pilot study demonstrates the feasibility of the paradigm and highlights areas for refinement in preparation for subsequent studies. These findings contribute to a broader understanding of how the brain processes multisensory inputs in complex visual environments.

36.468 NONRIGID VORTEX MOTION AND EYE TRACKING

Erin Conway[1], Danica Barron, Troy Smith, Ralph Hale; [1]University of North Georgia

Koerfer et al. (2024) demonstrated that nonrigid vortex motion can create a dissociation between perception of motion and smooth pursuit. In their paradigm, participants view a random dot field that is perturbed by a moving vortex. Because the dots are not moved along with the vortex, smooth pursuit mechanisms fail; however, catch-up saccades compensate for the lack of smooth pursuit gain, allowing participants to perceive and track the vortex's motion, although the motion may appear unstable and jumpy. This occurs because the system incorrectly assumes static placement of the vortex in space during this process. As a result, a catch-up saccade is initiated when the vortex appears to have jumped from its expected position relative to the eye. In the present study, this effect is tested as a necessary precursor to a larger, ongoing replication-plus-extension design. Participants viewed a black background covered with static white dots and a single dynamic vortex also comprised of white dots that moved at a rate of 10° per second laterally back and forth across the screen.

Eye tracker data measured eye position relative to the vortex position, recording eye movements to detect and measure smooth pursuit gains and catch-up saccades. In the absence of visual cues, smooth pursuit was not achieved in the tracking of the vortex. However, ongoing research aims to provide nonvisual multimodal cues that offset the lack of visual cues and facilitate smooth pursuit tracking. This study is the first vital step in improving our understanding of the interaction between various mechanisms that contribute to motion tracking and integration of multisensory stimulation.

Motion: Models, neural mechanisms

Sunday, May 18, 2:45 – 6:45 pm, Pavilion

36.469 MOVING STIMULI REVEAL LOCALIZED, NONLINEAR RECEPTIVE FIELDS

Felix Franke[1,2], Marc Büttner[1,2], Matej Znidaric[2,3], Roland Diggelmann[2,3], Federica B. Rosselli[2], Annalisa Bucci[2,3], Andreas Hierlemann[3]; [1]University of Basel, [2]Institute of Molecular and Clinical Ophthalmology Basel (IOB), [3]Eidgenössische Technische Hochschule Zürich (ETH)

Along the processing hierarchy of sensory systems, neurons become progressively more functionally specialized, i.e., their responses become increasingly selective to specific stimulus features. In the visual system, early stages encode changes in local contrast, whereas later stages selectively respond to complex stimulus features such as textures and object identities. Traditional approaches to study stimulus encoding in sensory systems are effective in revealing localized receptive fields for many neurons in early stages of the processing hierarchy. However, as functional specialization increases, and using traditional methods, the fraction of neurons showing strong and predictive receptive fields decreases. We demonstrate that this effect is already evident in the mouse retina where a significant fraction of retinal ganglion cells exhibits only weak and unpredictive receptive fields. The situation worsens one synapse downstream, in the nucleus of the optic tract (NOT), preventing receptive field identification for most neurons. We developed a method using structured, parameterized stimuli that are white in a nonlinear reparameterization of the stimulus space. Our approach efficiently identifies receptive fields, for nearly all recorded neurons, in both mouse and primate retinae, as well as in the mouse NOT. Most of the newly identified receptive fields are localized but nonlinear, i.e., the cells respond strongest to moving stimuli and not to unstructured contrast changes. Consequently, we describe the receptive fields in a movement space, rather than classical light intensity space. We use the approach to highlight differences in functional specialization between the primate and the mouse retina, as well as between the mouse retina and its direct projection target, the NOT.

Funded by SNSF (CRSII5_173728, CRSII5_216632; PCEFP3_187001; CRSK-3_220987, CRSK-3_221257; 310030_220209); ERC Advanced Grant (neuroXscales, 694829); Sedinum Foundation.

36.470 ANODAL TDCS ON VISUAL CORTEX ENHANCES MOTION AFTEREFFECT

Rebecca Weinberg[1], Yosun Yoon[1,2], Sang Wook Hong[1,2]; [1]Florida Atlantic University, [2]Stiles-Nicholson Brain Institute

Visual adaption is a critical tool for probing neural mechanisms of visual information processing. The current study investigates how enhanced neural excitability due to transcranial direct current stimulation (tDCS) affects motion perception using motion aftereffect (MAE). Participants completed 10 trials of MAE duration measurements in four conditions: (1) low contrast, (2) high contrast, (3) attention directed to the adapting stimulus, and (4) low contrast paired with either anodal or sham tDCS (randomized across participants) to the right visual cortex. Each trial began with participants viewing an adapting stimulus (rotating radial pattern) for 10 seconds. Following the adaptation phase, the same but static radial pattern was presented on the screen and participants were instructed to press and hold down the "z" key on the keyboard for as long as they continued to experience the MAE. The duration of the MAE was recorded as the time elapsed between the initial keypress and its release. We found that anodal tDCS significantly increased MAE duration compared to sham in the low-contrast condition, indicating enhanced neural adaptation under weak stimulus conditions. This lengthened MAE duration for the low-contrast with anodal tDCS was comparable to the MAE duration for the high-contrast adapting stimulus. Our results demonstrate that enhanced neural excitability by tDCS strengthens visual adaptation, suggesting that tDCS may elevate effective contrast of the adapting stimulus. Future research will investigate the role of tDCS across low to high contrast to reveal whether the modulational effect of tDCS on MAE is mediated by contrast gain or response gain.

36.471 HEMISPHERIC LATERALIZATION IN MOTION PROCESSING: A STUDY OF DEAF SIGNERS AND HEARING NON-SIGNERS

Carolyn D. Gershman[1] (cdg60@georgetown.edu), Ted Supalla[1,2], Anna Seydell-Greenwald[1,2], Barbara Landau[1,3], Elissa L. Newport[1,2]; [1]Center for Brain Plasticity and Recovery, Georgetown University Medical Center, [2]MedStar National Rehabilitation Hospital, [3]Johns Hopkins University, Department of Cognitive Science

Native signers have life-long expertise with a visual-gestural language that is produced by movements of the hands through space and constrained by spatial features, such as orientation, movement path, location, and handshape. Because language is typically lateralized to the brain's left hemisphere, lifelong processing of visual-spatial linguistic features might also result in left-lateralized processing of non-linguistic visual-spatial materials in native signers. Consistent with this hypothesis, behavioral studies have reported enhanced motion, velocity, and orientation discrimination among signers when stimuli are presented in the right visual field, and one fMRI study found a leftward shift in activation of motion processing area MT-MST for native signers viewing moving flow fields. However, few other imaging studies have investigated MT-MST lateralization in signers, and heterogeneous subject samples make interpretation difficult. Here we used fMRI to investigate lateralization for processing coherent motion in 19 deaf native signers and 19 hearing non-signers, contrasting a coherent motion direction judgment task with viewing of static dots (control). Group-level activation maps revealed bilateral MT-MST activation in both deaf and hearing participants and no significant group differences in lateralization (p=0.80). Compared to hearing participants, deaf subjects showed significant activation in auditory motion regions (p=0.002), which was lateralized to the right hemisphere. These results align with auditory deprivation research showing cross-modal changes in the deaf, with RH dominance in auditory motion regions for processing coherent motion. However, the findings do not support previous reports on LH lateralization for visual processing due to life-long sign exposure.

36.472 THE RELATIONSHIP BETWEEN RECOVERY FROM ADAPTATION AND PERCEIVING MOTION AFTEREFFECTS

Guang Yang[1], David Yu[1], Chris Paffen[2], Frans Verstraten[1]; [1]The University of Sydney, [2]Utrecht University

The motion aftereffect has intrigued many for ages. Although reported already by Aristotle, it still has outstanding questions that need answers. Following Barlow & Hill's famous 1963 Nature paper (see their figure 1), it is often assumed that the response characteristics of direction selective neurons reflect the percept of illusory motion in humans. It led for example Sekuler & Pantle (1967) to suggest that "the duration of the MAE is a function of the time to recover baseline level" as a valid model assumption. The question we address here is whether the percept of illusory motion reflects this recovery from adaptation process. In a series of experiments, we tested whether the recovery from the adaptation period reflects the perceived illusory motion. We focus on unidirectional motion. Individual MAE durations were measured assuming that subsequent adaptation to an orthogonal motion direction combines with subthreshold activity of previous presented motion, if present. That is, in the case the recovery process is still ongoing, the part below perceptual threshold can be made visible again by integrating the residual activity with a supra threshold MAE signal of a second adaptation pattern, which then is reflected in the direction of the combined aftereffect (see Verstraten et al., VisRes, 1994). This direction is in between the aftereffect directions of both adapting patterns, if there is subthreshold residual adaptation. The results show that the perceived illusory motion reflects only about half of the recovery from adaptation period. We conclude that the motion aftereffect is better defined as 'the visible part of the recovery from adaptation period'. This means that part of this process is below perceptual threshold and should be taken into consideration when using the MAE as the psychophysicist's micro-electrode (John Frisby, 1979). Not perceiving the aftereffect doesn't mean there is no recovery from adaptation anymore.

36.473 MULTICOMPONENT MOTION IS PERCEPTUALLY ORGANISED ALONG DISTINCT DIMENSIONS REVEALED BY DISTANCE SCALING AND FORCED CHOICE PARADIGMS

Andrew Isaac Meso[1], Jonathan Vacher[2], Nikos Gekas[3], Pascal Mamassian[4], Guillaume S. Masson[5]; [1]King's College London, [2]Université Paris Cité, [3]Edinburgh Napier University, [4]Ecole Normale Superieure, [5]CNRS & Aix-Marseille Université

Human participants presented with composite stimuli containing two motion vectors, e.g. with upwards and rightwards directions, or very fast and very slow speeds, typically report transparent motion. The transition from coherent to transparent relies on visual processing mechanisms which disambiguate inputs to infer perceptual organisation. To study these mechanisms we use dynamic textures, Motion Clouds, which span a range of mean spatial frequency and bandwidth in seven test conditions, and for each case generate composites of paired stimuli in which the mean speed is the same, but the speed difference is discretely varied from very low to high. We use two experimental approaches. First, with Maximum Likelihood Distance Scaling (MLDS), on each trial we present four stimuli asking which two are more different from each other avoiding direct mention of transparency or coherence. Participants' data from thousands of trials show differences between stimuli which generally increase for larger speed differences. The shape of the perceptual function and the sensitivity depend on spatial frequency and bandwidth. When participants do a second task, a Forced Choice detection in which they report whether they perceive one or two components in each presented trial, we find that they are only able to reliably report two components for stimuli with mid-spatial frequencies of 1cyc/deg, not 0.25 or 4.0 cyc/deg. Participants also show low sensitivity for broad bandwidths of 2.0 octaves. In a direct comparison of the Forced Choice and MLDS results, participants reliably perceive stimulus differences across speed differences even when they cannot report perceiving two components. The shapes of the MLDS curves vary across individuals being either convex or concave relative to a linear prediction. Our findings reveal that beyond perceiving one or two components, there are additional cues for example rigid/non-rigid on which participants reliably decompose multicomponent motion stimuli when distinguishing them.

36.474 THE CONTRIBUTION OF MOTION DETECTORS DURING MULTIPLE-OBJECT TRACKING

Maryam Rezaei[1], Remy Allard[1]; [1]Universite de Montreal

Motion perception, an essential skill for human beings, relies on two motion processing systems: a low-level system mediated by early direction-selective neurons known as motion detectors, and a high-level system that attentively tracks the position of objects. The current study aimed to investigate the contribution of motion detectors during multiple-object tracking through two experimental manipulations: reverse-phi and stroboscopic motion. By reversing the contrast polarity at each object displacement, reverse-phi inverses the direction response of motion detectors. By introducing a temporal gap between each displacement (i.e., object disappears briefly before reappearing at a new position), stroboscopic motion can weaken the contribution of motion detectors. Five young participants were asked to track and identify four balls among eight identical ones bouncing around within a virtual three-dimensional cube. The maximum speed threshold at which the 4 target balls were successfully tracked were determined under fourteen conditions: 2 contrast polarities (i.e., reversed or not) X 7 temporal gaps (i.e., 0, 17, 33, 50, 67, 83, or 100 msec). A significant interaction between contrast polarity and temporal gap on maximum speed threshold was found ($F(6, 24)=9.194$, $p<.001$). Post hoc analyses revealed that with temporal gaps of 0 and 17 msec, the maximum speed threshold was significantly lower when the contrast polarity of the balls was reversed compared to when it was not ($p<.05$). With longer temporal gaps, however, no significant difference was observed between the two contrast polarity conditions. The results suggest that motion detectors considerably contribute to multiple-object tracking when the temporal gaps between object displacements are short (<33 msec). But for longer temporal gaps, no considerable contribution of motion detectors was observed suggesting that multiple-object tracking relied mainly on the higher-level, attention-based motion processing system.

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Fond du Recherche du Québec-Nature et Technologie (FRQNT) to RA.

36.475 TESTING MODEL PREDICTIONS FOR A CHASE DETECTION TASK

Maria Kon[1] (maria.r.kon.civ@us.navy.mil), Sangeet Khemlani[1], Andrew Lovett[1]; [1]U.S. Naval Research Laboratory

Chase detection is a demanding task in which participants view a scene with moving objects and indicate whether one is chasing another. While this task offers insight into intention attribution, the mechanisms of chase detection remain poorly understood. Recent work (Kon et al., 2024) introduced a new chase detection paradigm that measured response time and accuracy while varying set size and whether a chase was present or absent. The authors developed an attention-based computational model that made specific predictions about performance. Here we present the results of an experiment that used the same chase detection paradigm but also varied a cue introduced by Gao et al. (2009), i.e., chase subtlety (the degree to which the chasing object deviates from a direct, "heat-seeking" path towards the chased object). Prior to collecting data, we conducted model simulations and made preregistered predictions about performance. Those simulations correlated well with the data, with two notable discrepancies: (1) human response times for chase-absent trials were much lower compared to the model; (2) for chase-present trials with the highest chasing subtlety, the model had lower accuracy and faster response times compared to humans. In light of (1), we refined the model's stopping rules. (2)'s mismatches may be the result of human participants learning to classify videos with higher chasing subtlety as chases due to feedback received after each trial, and they may have waited longer to determine whether there was such a chase. We tested these hypotheses by conducting a follow-up experiment that provided no feedback on experimental trials. Thus, the model both makes testable predictions about task performance and suggests the need to reassess the role of chasing subtlety in such tasks. Further, the experimental results provide an opportunity to refine the model, leading to a deeper understanding of how humans detect chases.

This research was supported by a National Research Council Research Associateship awarded to MK and by a grant from the Naval Research Laboratory awarded to SK.

36.476 TESTING THE "CAUSAL" IN CAUSAL INFERENCE DURING MOTION PERCEPTION

Xinyi Yuan[1], Sabyasachi Shivkumar[2], Maeve Silverman[1], Ralf M. Haefner[1]; [1]University of Rochester, [2]Columbia University

Retinal motion is a key cue for inferring object motion, scene parsing, and object relationships. Recent work has modeled this as hierarchical causal inference (Shivkumar et al., 2024), assuming the brain infers causal links between moving elements in a scene in order to construct the reference frame within which to represent motion. For example, a person's motion should be interpreted in the frame of a moving escalator if they are walking on it, but not otherwise. We tested this by manipulating the causal structure of a scene using a center-surround motion stimulus. Typically, perceived center motion is biased towards the surround's direction for small relative differences (integration) and away for larger differences (segmentation). If causal inference drives this bias, it should weaken when evidence against a causal link is provided and strengthen when evidence supports it. Four human observers reported their perceived motion direction using a dial under four conditions each (within-observer comparisons): (1) both center and surround reversed direction in a way consistent with a causal link, (2) the center did not reverse while the surround did, weakening causal evidence, and (3)&(4) observers controlled the reversal point using a button press, further modulating causal evidence. Results showed biases that clearly differed between causal conditions. Interestingly, direct intervention by observers provided little additional causal information compared to the observation of common motion alone. Notably, we found substantial inter-observer variability: two observers showed a reduction in surround influence in conditions 2&4 compared to 1&3, reflecting the differences in causal information, while the effect was less clear for the other two. Furthermore, even in the acausal conditions 2&4 a bias remained. These findings highlight unexpected variability in how observers interpret common motion and suggest that retinal motion provides robust causal information, with limited additional benefit from direct stimulus control.

We acknowledge funding support from NIH/U19 NS118246, and NSF/CAREER IIS-2143440.

36.477 A DEEP LEARNING MODEL OF SELF-MOTION USING GAZE-CENTERED IMAGE SEQUENCES

Nathaniel Powell[1], Youjin Oh[1], Mary Hayhoe[1]; [1]University of Texas at Austin

Retinal flow patterns while walking depend not only on gaze location but also on the oscillations of the head (Matthis et. al., 2022). When humans fixate a location while walking, the eyes counter-rotate in the orbit to stabilize the retinal image. Additionally, normal gait is accompanied by rhythmic translations and rotations of the head in space, meaning the momentary heading direction varies greatly over the gait cycle, creating a constantly changing, complex pattern of motion on the retina. These patterns are ubiquitous during development, however most stimuli used to investigate the coding of self-motion do not use the stimuli that reflect the complexity of the actual retinal motion patterns. Therefore, how might the visual system encode the visual motion? Neural recordings from visual areas MT and MST show responses to translational, expanding, and rotating patterns of motion. Recent work by Mineault et. al., (2021) used deep learning and found that predicting the translation and rotation of a camera moving through space based on a stacked-sequence of images was sufficient for producing MT and MST-like receptive fields. The sequences of images used in training, however, were different from the natural stimuli people experience while walking. A similar deep learning model was trained on gaze-centered images sequences recorded from a subject walking through virtual environments with different depth structures. The model was trained to predict the translation and rotation of the head. It learned the motion parameters well and developed speed and direction tunings similar to receptive fields in MT and MST. This suggests that the information present in the retinal flow field is enough to extract the momentary direction of heading. Surprisingly, this model learned the heading parameters despite the presence of saccades in the input sequences, suggesting that the model is robust to discontinuities caused by saccadic eye movements.

36.478 COMPARING A DUAL-STREAM ARCHITECTURE WITH SINGLE-STREAM CNNS TO SIMULATE VISION IN LOCOMOTOR CONTROL

Zhenyu Zhu[1], Thomas Serre[1], William H. Warren[1]; [1]Brown University

Zhu and Warren (2022) asked participants to follow a group of textured objects, whose heading direction or speed of motion was briefly perturbed. Overall, locomotor responses were consistent with boundary motion (feature tracking). When the object texture and boundaries moved in Opposite directions (the reverse-phi illusion), responses decreased with increasing boundary blur, but this did not occur in the Same direction condition. This widening gap between the two conditions indicates that visually-guided locomotion depends on a weighted combination of feature-tracking and motion energy. Here, we leverage deep neural networks to investigate what network architectures can replicate these effects. We evaluated several model architectures, including single-stream 3D convolutional networks (MC3, R(2+1)D, R3D, DorsalNet), a dual-stream network (SlowFast), and a benchmark neurophysiological motion energy model (Nishimoto and Gallant 2011). The dual network includes a low temporal-frequency/high spatial-frequency stream, and a high temporal/low spatial stream. These models were fine-tuned to estimate the heading and speed of a group of objects with attached surface textures moving across various backgrounds from 12-frame video sequences, as in ecological contexts. The fine-tuned models were then tested on Zhu and Warren's (2022) reverse-phi stimuli. The mean heading estimates of the single-stream and motion energy models decreased significantly in both Same and Opposite conditions as boundary blur increased (range of $F(2, 594)$: 16.45 - 96.00, all $p < 0.0001$), deviating from human responses. These findings suggest that single-stream 3D convolutional networks function similarly to motion energy detectors, without the feature-tracking observed in humans. However, the SlowFast network failed to replicate the increasing gap between Same and Opposite conditions with boundary blur. We conclude that the SlowFast model does not capture human-like feature tracking. This indicates the need for further architectural improvements, such as incorporating recurrent connections to support feature-tracking.

Funding: NIH R01EY029745, NIH 1S10OD025181, NIH T32MH115895

Multisensory Processing: Perception, neural, clinical

SUNDAY, MAY 18, 2:45 – 6:45 PM, PAVILION

36.479 CROSS-MODAL SERIAL DEPENDENCE

Adam Zaidel[1] (ajzaidel@gmail.com), Baolin Li[2], Biyao Wang[2], Yael Shamir-Bercovich[1], Benjamin Menashe[1], Shir Shalom-Sperber[1], Aihua Chen[3]; [1]Bar-Ilan University, Ramat Gan, Israel, [2]Shaanxi Normal University, Xi'an, China, [3]East China Normal University, Shanghai, China

Serial dependence is often studied within a uni-modal context – primarily, vision. However, our experience of the world is naturally multisensory. In a series of four recent studies, we tested cross-modal serial dependence between vision and other modalities (visual-vestibular self-motion perception, and visual-auditory timing perception). To dissociate the differential effects of previous stimuli and previous choices on current decisions, we fit the data from uni-modal and cross-modal conditions, with a logistic regression model. In general, previous stimuli led to negative (repulsive) effects, while previous choices led to positive (attractive) effects. Cross-modal serial dependence was seen for visual-vestibular self-motion perception. Specifically, previous stimuli elicited a repulsive effect on the subsequent perceptual decision, in both uni-modal and cross-modal conditions. In a follow up experiment, we found that these effects remain even when attention is diverted away from the previous stimuli (using a distractor task). By contrast, in two studies of audio-visual timing perception (perception of duration, and of changes in tempo) repulsive effects of previous stimuli were seen only in uni-modal, but not cross-modal, conditions. Here, I will present the various results from these studies, and discuss what can be learned from the presence or absence, and different types, of cross-modal serial dependence.

36.480 HOW DO AUGMENTED REALITY OBJECTS, WHICH CAN ONLY BE PERCEIVED VISUALLY, INFLUENCE OUR SPATIAL BEHAVIOR?

Shachar Maidenbaum[1] (shachar.maidenbaum@gmail.com), Tevel Katzir[1], Ilan Vol[1]; [1]Ben Gurion University, Israel

Navigating around objects is a common spatial activity for humans. However, what happens when these objects are virtual - e.g., added to our environment via augmented reality tools - and can be perceived only visually? Do humans view and treat them as real obstacles? We first explore this question using inanimate objects (e.g. chairs). Participants (n=36) walked down real hallways while wearing augmented reality displays, with matching physical and virtual objects blocking their path. We found that most participants (~85%) treated the virtual and real obstacles similarly, avoiding both with equal distances, while a smaller group consistently treated them differently, keeping significantly less distance from virtual obstacles. In a separate study we tested these questions on augmented reality avatars including humans and robots, replicating results found in both fully virtual and fully real environments demonstrating that participants viewed and treated them as having a real presence - and here too with a similar subgroup who behaved significantly differently. These results strengthen the baseline of human spatial behavior in mixed reality, offer a useful tool for testing the basic science of spatial interaction and multisensory interaction, and offer practical potential both for rehabilitation and for the design of mixed reality interfaces. Finally, our findings demonstrate the importance of individual differences in user reactions to augmented content, and the way it affects their spatial performance.

ISF 1322/22

36.481 THE INFLUENCE OF VISION ON TACTILE STROKE PERCEPTION: EVIDENCE FROM A NOVEL MIRROR BOX ILLUSION

Stephanie Dietz[1] (stephanie.dietz@emory.edu), Anupama Nair[2], Jared Medina[3]; [1]Emory University, [2]University of Delaware

Information from different senses is weighted based on their relative precision. For instance, in the mirror box illusion, the greater spatial precision of vision compared to proprioception causes observers to feel their hand in its seen location rather than its actual location. How does the brain resolve conflicting visual and tactile information? To address this question, we ran three experiments using a novel mirror box illusion. Both hands (one hand in front of the mirror, one hand behind the mirror) were stroked simultaneously across the entire hand dorsum in congruent or incongruent directions (stroke disparity ranging from 45-180°) for 30 seconds with the mirror viewed or covered. Participants then reported the perceived stroke direction on their hidden hand. We found evidence for a visuotactile illusion, such that participants frequently perceived touch in the seen direction rather than the actual direction. Interestingly, when the angular disparity was 180° (e.g., viewed stroke towards the fingers, felt stroke towards the wrist), we observed an all-or-nothing effect, such that participants reported either the actual touch or complete visual capture. To investigate whether the effect of vision is modulated by hand realism, we examined performance with a real or rubber hand. Although we predicted greater illusion for the real than rubber hand, we found no difference between the two conditions, suggesting that body ownership may not modulate the illusion. Lastly, to determine whether vision influences the perception of tactile stroke direction in a hand-centered or external reference frame, we varied the posture of the viewed hand relative to the hidden hand (0-90°). When the viewed hand was rotated and stroke direction was somatotopically congruent, participants reported a shift in perceived stroke direction towards the visual stroke direction in external space. Our results suggest that visuotactile temporal congruence can override spatial incongruence, leading to illusory percepts.

36.482 MULTISENSORY OVERWEIGHTING OF PERCEPTUAL PRIORS RELATES TO POSITIVE SYMPTOMS IN INDIVIDUALS AT CLINICAL HIGH RISK OF PSYCHOSIS

Victor Pokorny[1] (vpokorny123@gmail.com), Lauren Ellman[2], Gregory Strauss[3], Elaine Walker[4], Scott Woods[5], Albert Powers[5], Philip Corlett[5], Steven Silverstein[6], James Waltz[7], James Gold[7], Jason Schiffman[8], Vijay Mittal[1]; [1]Northwestern University, [2]Temple University, [3]University of Georgia, [4]Emory University, [5]Yale University, [6]University of Rochester, [7]University of Maryland, [8]University of California, Irvine

Positive symptoms, such as hallucinations and delusions, are core features of psychotic disorders and are elevated in individuals at risk for developing psychotic disorders. We, and others, have hypothesized that positive symptoms, especially hallucinations and perceptual distortions, are caused by overweighting of perceptual priors. We administered two tasks that are hypothesized to reflect overweighting of perceptual priors, a visual Mooney Faces task and an auditory Sine Wave Speech task, to over 800 adolescents and young adults oversampled for psychosis risk. Both tasks required participants to indicate whether they recognized naturalistic stimuli (faces or speech) that were heavily filtered to create ambiguous perceptual experiences. We observed a significant correlation between recognition rates produced by the Mooney Faces task and the Sine Wave Speech task ($r(765)=0.32$, $p<.001$, 95% CI [0.26,0.38]). Thus, these tasks appear to jointly reflect overweighting of perceptual priors, despite differing with respect to perceptual modality. Recognition rates from both tasks independently predicted positive symptom severity (Mooney faces: $r(784)=0.19$, $p<.001$, 95% CI [0.12,0.26]; Sine Wave Speech: $r(818)=0.16$, $p<.001$, 95% CI [0.1, 0.23]). However, a composite score (combining indices from both tasks) exhibited a stronger association than either task individually ($r(760)=0.23$, $p<.001$, 95% CI [0.16,0.29]). Our results suggest that multisensory overweighting of perceptual priors may, in part, explain the provenance of positive symptoms in those at risk for developing psychosis. Furthermore, we argue that tasks that measure such overweighting may be useful for better identifying individuals at risk for developing such disorders.

This research was funded by National Institutes of Health (R01MH120090, R01MH112613, R01MH120091, R01MH120092, R01MH116039, R21MH119438, R01MH112545, R01MH1120088, U01MH081988, R01MH112612, R01MH120089).

36.483 SPATIALLY UNINFORMATIVE SOUNDS MODULATE MIDBRAIN VISUAL ACTIVITY WITH AND WITHOUT PRIMARY VISUAL CORTICAL INPUT

Tatiana Malevich[1,2], Matthias P. Baumann[1,2], Yue Yu[1,2], Tong Zhang[1,2], Ziad M. Hafed[1,2]; [1]Werner Reichardt Centre for Integrative Neuroscience, [2]Hertie Institute for Clinical Brain Research, [3]University of Tuebingen

We recently discovered that spatially uninformative sounds can activate otherwise dormant visual-motor pathways bypassing the primary visual cortex (V1). Here, we aimed to better understand how this might happen. We recorded from superior colliculus (SC) neurons (two monkeys) with either intact or focally inactivated V1 (muscimol microinjection; 1.5-2.5 µL; 10 mg/1mL). We presented a 0.2 deg radius disc within the neurons' receptive fields (RF's) and randomly interleaved trials in which we paired the visual stimulus onset with a bilateral sound pulse (50 ms; 1 kHz). This sound pulse was neither informative about the visual stimulus location nor spatially aligned with RF locations. With intact V1, SC neurons showed little, if any, responses to the sound alone. Nonetheless, visual response strength and latency were diversely affected (sometimes being stronger and earlier for the vision+sound trials, other times being weaker and later, and yet other times being unaffected). Such multisensory integration was also evident in local field potentials (LFP's), with evoked responses to multisensory stimuli being enhanced and distinct from those elicited by unimodal stimuli in either modality. With inactivated V1, SC visual responses were much sparser, both at the single unit and LFP levels. However, adding spatially uninformative sounds unmasked a relatively weak visually-evoked LFP response that was not explained by sound-only responses. There were also few single units that exhibited clear multisensory integration in their spiking. Next, we sampled some inferior colliculus (IC) neurons. With intact V1, when IC neurons exhibited visual responses, they showed clear multisensory integration, even without having sound-only responses;

however, LFP visually-evoked responses were predominantly driven by sound. With inactivated V1, IC single-unit and LFP responses were abolished. These results underscore the distinct SC and IC roles in multisensory integration, and they support a potential SC involvement in visually-guided behavior when V1 is compromised.

36.484 LATERALITY OF THE VISUAL CLAUSTRUM

David Linhardt[1], Dominik Zuschlag[1], Adam Coates[2,3], Natalia Zaretskaya[2,3], Christian Windischberger[1]; [1]Medical University of Vienna, Austria, [2]University of Graz, Austria, [3]BioTechMed-Graz, Austria

The claustrum is an interconnected sheet-like subcortical structure with a potentially important role in sensory integration and attention. Recent work provided the first fMRI evidence of a visual zone within the human claustrum (Coates et al., 2024). Structural connectivity studies have further demonstrated white matter pathways linking the claustrum to the ipsilateral visual cortex in both sheep (Pirone et al., 2021) and humans (Milardi et al., 2015), suggesting its importance in sensory processing. Since the visual cortex of one hemisphere contains a representation of the contralateral hemifield, this suggests that a similar functional organization may exist in the visual claustrum zone. To investigate this, we analyzed data from the Human Connectome Project retinotopy dataset (Benson et al., 2018). Using fMRI data collected from 181 participants during visual stimulation, we designed two orthogonal regressors corresponding to periods when a wedge stimulus traversed the left or the right visual hemifield. These regressors were applied within a general linear model (GLM) to analyze BOLD signals in the claustrum and early visual cortex. This allowed us to probe the lateralization of visual responses. Our results reveal a lateralized activation pattern: the claustrum in each hemisphere shows activation together with its ipsilateral visual cortex during contralateral visual field stimulation. This organization mirrors the cortical contralateral representation of visual space and provides the first evidence of functional lateralization in the human claustrum's visual zone. These findings highlight the claustrum's potential role in hemifield-specific visual processing.

This research was funded in whole or in part by the Austrian Science Fund (FWF) [10.55776/P35583; 10.55776/PAT8722623]. AC and NZ were supported by the BioTechMed-Graz Young Researcher Group Grant.

MONDAY MORNING POSTERS IN BANYAN BREEZEWAY

Visual Memory: Imagery, long-term

Monday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

43.301 IT'S ALL ABOUT SEMANTICS: HOW SEMANTIC CATEGORIES SHAPE MEMORY PARTITIONING IN HYBRID VISUAL SEARCH

Nurit Gronau[1], Makaela Nartker[2], Sharon Yakim[1,3], Igor Utochkin[4], Jeremy Wolfe[5,6]; [1]The Open University of Israel, Israel, [2]University of Texas at Austin, USA, [3]The Hebrew University, Israel, [4]University of Chicago, USA, [5]Brigham and Women's Hospital, [6]Harvard Medical School, USA

In many everyday situations, we search our visual surroundings for one of many types of possible targets held in memory, a process known as hybrid search (e.g., searching for items on a shopping list). In some cases, only a portion of the memorized list is relevant to a specific visual context, so restricting the memory search to the relevant subset would be beneficial (e.g. there's no need to search for carrots in the dairy section). Previous research has shown that participants often fail to "partition" memory into distinct, useful subsets on a trial-by-trial basis. However, given the known role of semantic content in long-term memory organization, we hypothesized that clearly semantically-defined subsets could facilitate flexible memory partitioning in dynamic hybrid search situations. Experiment 1 revealed that, indeed, semantic characteristics (i.e., object category), but not perceptual features (e.g., arbitrary color), can serve as a strong basis for flexible memory partitioning. Experiments 2 and 3 further demonstrated that this memory partitioning is cost-free and independent of the nature of surrounding visual distractors (i.e., whether the distractors are categorically homogeneous or heterogeneous). These findings demonstrate that confining memory search to a relevant subset of items can be highly effective when the subsets are defined by clear semantic categories. The results underscore the importance of conceptual information in the organization of activated long-term memory (aLTM) – the portion of LTM relevant to the current task - and to aLTM's role in enabling flexible, trial-by-trial memory selection. Additionally, our findings highlight the relationship between visual search and memory search: Despite the activation of multiple sets or categories that could interfere with each other at the attentional (visual search) level, category-based memory partitioning seems to remain relatively immune to interference from other categories during this type of memory search.

Israeli Science Foundation (ISF) grant 1622/15 (to NG) and by the National Institutes of Health (NIH) grant EY017001 (to JMW).

43.302 ANTICIPATING ATTENTIONAL GUIDANCE ENHANCES RELIANCE ON LONG-TERM MEMORY

Duygu Yücel[1], Nursena Ataseven[1], Lara Todorova[1], Berna Güler[1], Keisuke Fukuda[2,3], Eren Günseli[1]; [1]Department of Psychology, Sabanci University, Istanbul, Turkey, [2]Department of Psychology, University of Toronto, Toronto, Canada, [3]Department of Psychology, University of Toronto Mississauga, Mississauga, Canada

Many attention models propose that working memory (WM) is crucial for biasing perception towards attentional goals. However, in daily life, we often search for information stored in long-term memory (LTM). We hypothesized that if WM plays a critical role in attentional guidance, it should reactivate attentional goals stored in LTM. To test this, we used the contralateral delay activity (CDA), an electrophysiological marker of WM storage, to compare the recruitment of WM in storing LTM goals during preparation for a visual search task versus a recognition task. CDA amplitude was higher when storing the memory item for recognition than for visual search, suggesting that humans rely more on LTM than WM to store attentional goals. This finding reveals an adaptive strategy: humans offload task goals to LTM during preparation for more demanding tasks to preserve WM resources. This finding challenges models that claim WM is essential for guiding attention, highlighting greater reliance on LTM for attentional guidance.

This project is funded by The Scientific and Technological Research Institution of Turkey (TUBITAK - 118C248) provided to Eren Günseli.

43.303 CHALLENGING DUAL-CODING THEORY: PICTURE SUPERIORITY IS PRESERVED IN APHANTASIA

Muhan Yan[1], Brady R.T. Roberts[1], Wilma A. Bainbridge[1]; [1]University of Chicago

Dual-coding theory proposes that superior memory for pictures compared to words (picture superiority effect) and concrete words compared to abstract words (concreteness effect) stems from the ability to store information in both verbal and image codes. According to the theory, having two distinct memory codes available during retrieval increases performance because either code can independently lead to successful recall, and one code can serve to cue the other. For example, when recalling an image of a dog, one can retrieve the verbal code ('dog'), or they can first retrieve the mental imagery of the previously seen dog which then affords the verbal code to recall. The current study tests this theoretical explanation by examining recall performance in individuals with aphantasia—the inability to voluntarily generate mental imagery. Based on dual-coding theory, aphantasic individuals should show neither the picture superiority effect nor the concreteness effect due to their inability to generate mental images as retrieval cues on a recall memory test. We compared recall performance between aphantasic and typical imagers across four stimulus types: pictures, symbols, concrete words, and abstract words. Despite their imagery deficit, aphantasic individuals still demonstrated better recall performance for both pictures and symbols compared to words. Moreover, while typical populations show similar recall for pictures and symbols, aphantasic individuals displayed superior recall for symbols compared to pictures, possibly reflecting their alternative cognitive strategies that particularly benefit symbol processing. The persistence of better recall performance for both pictures and symbols in aphantasia suggests that there are memory mechanisms beyond dual-coding that drive these effects, or that dual-coding theory requires significant modification. This study provides novel insights into the relation between mental imagery and memory, challenging traditional theoretical frameworks and suggesting the need for alternative explanations of these well-established memory phenomena.

This research was supported by a Natural Sciences and Engineering Research Council (NSERC) of Canada Postdoctoral Fellowship to BRTR and National Eye Institute Grant R01-EY034432 to WAB.

43.304 MENTAL IMAGERY PREDICTS PERFORMANCE ON MEDIAL PARIETAL RELATED COGNITION

Elissa Aminoff[1] (eaminoff@fordham.edu), Anya McGoldrick[1]; [1]Fordham University

Mental imagery ability varies across individuals. How pervasive these individual differences are across different cognitions is unknown. The medial parietal region, which includes the retrosplenial cortex, has been linked to mental imagery. We predicted that performance in other cognitive abilities linked to this brain region would correlate with mental imagery ability. If so, this would suggest a more fundamental source of individual differences may lie within differences in medial parietal/retrosplenial functioning rather than in mental imagery. To test this, we examined 189 participants from the general population. Their imagery capabilities were correlated to scores in memory, scene perception, virtual navigation, and an unrelated control task of lexical decision-making. Those with better mental imagery performed significantly better on the memory test and had faster reaction times during scene perception, with similar accuracy. The virtual navigation task yielded null results, which contradicted the self-reported navigation results in the memory questionnaire and was concluded to be a design error. The control condition – the lexical decision task – did not correlate with mental imagery, strengthening the study's findings that activity in the medial parietal/retrosplenial accounts for these differences. The study suggests that differences in mental imagery ability may be related to differences in medial parietal/retrosplenial functioning more generally, manifesting in correlated differences in other cognitions that utilize the same brain region. However, more research, especially using fMRI, is needed to explore what role this brain region has in imagery and the potential role of imagery in related cognitions such as memory and scene perception.

43.305 DECODING HIERARCHICAL VISUAL FEATURE REPRESENTATIONS IN MENTAL IMAGERY OF APHANTASICS

Tomoyasu Horikawa[1] (horikawa.t@gmail.com); [1]NTT Communications Science Laboratories

Mental imagery, regarded as a quasi-perceptual phenomenon, typically generates visual experiences without external stimuli and recruits neural representations associated with visual features that overlap with those used in visual perception. While these shared neural representations are often assumed to underlie the phenomenal aspects of conscious visual experiences, their contribution to subjective visual awareness remains unclear. Aphantasia, characterized by the subjective inability to voluntarily generate visual mental imagery, offers a unique opportunity to examine how mental images are represented without subjective visual experience and to test the relationship between neural representations and conscious experience. We thus measured brain activity in aphantasics as they viewed and imagined object images using functional magnetic resonance imaging, applying deep neural network (DNN) feature decoding to comprehensively analyze hierarchical visual feature representations across multiple levels. Decoding models were trained on stimulus-induced brain activity in visual cortical areas to predict feature values of viewed images. We then applied these models to imagery-induced brain activity to decode DNN features, comparing them with category-averaged features from a large-scale image database. Feature values decoded from imagery-induced brain activity positively correlated with those of imagined objects across multiple DNN layers, enabling above-chance identification of imagined object categories by matching them to averaged candidate features. Remarkably, while the decoding performance of aphantasics was somewhat lower, it approached that of typical imagers. These results demonstrate that neural representations of hierarchical visual features, shared with visual perception, are engaged during mental imagery in aphantasia, even without subjective visual awareness. This dissociation between neural representations and visual awareness highlights significant limitations in assuming that decodable neural representations directly reflect the neural correlates of consciousness. While supporting the presence of mental imagery, decodable representations alone fall short of fully explaining conscious experience, underscoring the need for cautious interpretation as evidence of consciousness.

This research was supported by grants from JSPS KAKENHI Grant Number JP22H03910.

43.306 CHARACTERIZING THE VISUAL FEATURES ENCODED IN FEEDBACK-RELATED ALPHA RHYTHMS DURING NATURAL VISUAL IMAGERY

Rico Stecher[1], Daniel Kaiser[1]; [1]Justus-Liebig-University Gießen

Recent research suggests that when we generate a mental image, our brain employs feedback-related alpha rhythms to reactivate visual content representations. However, what visual features are encoded in these rhythms has only scarcely been examined. To answer this question, we acquired a large EEG data set through extensive sampling of individual participants. Each participant imagined 16 natural scenes according to short text prompts for a total of 10 sessions (i.e. 4,320 trials per participant, 44,320 trials in total). In order to characterize the visual feature information alpha rhythms carry during imagery, we approximated participants' mental images using generative AI. We fed the imagined scene prompts into a latent text-to-image diffusion model and generated a large set of candidate images for each prompt (100 images per prompt, 1,600 in total). We then evaluated the AI-generated candidate images with a convolutional neural network (CNN) trained on scene categorization. Finally, we generated representational dissimilarity matrices (RDMs) from both the CNN activations and the EEG responses in the alpha band and compared the CNN RDMs to the EEG RDMs to determine how CNN features extracted across the CNN hierarchy predict the neural representation of mental images. Across individual participants, we found remarkable variations in how features of different complexity (coded at different depths of the CNN) contribute to the representations of mental images. On the group level, we found that the correspondence between participants' alpha-band representations was most pronounced in intermediate CNN layers, suggesting that mental images of scenes are prominently shaped by mid-level visual features. Taken together, our results show that, when we conjure complex natural environments before our minds' eye, alpha rhythms reactivate a complex range of visual feature information.

DK is supported by the DFG (SFB/TRR135,222641018; KA4683/5-1, 518483074, KA4683/6-1, 536053998), "The Adaptive Mind", and an ERC Starting Grant (PEP, ERC-2022-STG 101076057).

43.307 VISUAL MEMORY CAPACITY IS SUSCEPTIBLE TO MULTIPLE INFLUENCES DURING MORE NATURALISTIC ENCODING

Sharon Gilaie-Dotan[1], Reout Dayan[1], Olga Kreichman[1]; [1]Bar-Ilan University

Although studies show that human visual memory for object images is phenomenal during active engaging encoding, it is still unclear if this reflects memory during naturalistic visual behavior. Recent studies show that during naturalistic encoding, image memory is much poorer than the levels reported earlier and that the knowledge that an image was not seen before is much higher than recognition memory for already-seen images. Here in a set of comprehensive experiments (n=249) we parametrically examined whether during more naturalistic-like encoding, memory for object images is robust or susceptible to multiple factors. Earlier results with face, people, indoor and outdoor images were first replicated with object images. In addition, we found that during these less engaging encoding conditions people's memory was adversely affected by the number of encoded images and by the number of novel (distractor) images the previously seen image was tested amongst (from best performance of ~90% memory in a 2AFC task for 60 encoded images down to ~50% memory in a 6AFC task for 180 encoded images). While human capacity for remembering new images can reach outstanding levels during highly engaging conditions, our results indicate that during more naturalistic-like conditions, image memory is much weaker and susceptible to multiple influences. We assume that our findings overestimate everyday visual memory capacity and that much less of the novel information we come across everyday actually registers in memory.

This study was funded by ISF Individual Research Grants 1485/18 and 1462/23 to SGD, and by CHE Lev-Zion Scholarship to OK.

43.308 WIDE LEARNING: BUILDING SEMANTIC NETWORKS AND DIFFERENTIATING BETWEEN SEMANTIC AND VISUAL REPRESENTATIONS IN LONG-TERM MEMORY

Yael Schems Maimon[1,2], Yoed Kenett[3], Adva Shoham[1,4], Ayala Allon, Galit Yovel[1,5], Roy Luria[1,5]; [1]Tel Aviv University, [2]Minducate Science of Learning Research and Innovation Center, [3]Faculty of Data & Decision Sciences, Technion, [4]Sieratzki Institute for Advances in Neuroscience, [5]Sagol School of Neuroscience, Tel Aviv University

It has long been demonstrated that representations in long-term memory (LTM) are semantically organized, with related words represented 'closer' in LTM. Importantly, previous LTM research either tapped its' episodic properties (e.g., using lists of unrelated words), or relied on the existing semantic network (by presenting a list of related words). The current research introduced Wide Learning, a novel paradigm that exposed participants to rich details regarding objects from an unfamiliar category (beekeeping or skippering), using texts, videos and images. Learning was assessed by measuring object naming, familiarity ratings, semantic similarity ratings between each word pair, visual similarity ratings between each image pair and story writing. Critically, this type of learning was found to effectively create a semantic network connecting all learned objects. Additionally, this network was connected to existing related LTM representations, while disconnected from irrelevant concepts. This semantic network was absent before learning, and for unfamiliar words that were not learned. Moreover, the results revealed a strong reliance on the emerging semantic network in the story-writing task. While a great part of the learning phase was visual (e.g., images and videos), we further demonstrated that learning was semantic rather than visual. To examine the unique contributions of visual and semantic information, we generated pure visual representations of the objects using a visual deep neural network (DNN) based on their images, and pure semantic representations using a language DNN based on their dictionary definitions, and examined their correlations with human similarity ratings before and after learning. Whereas the correlations with semantic similarity ratings increased after learning, there was no change in the correlations with perceptual similarity ratings of the objects. Overall, this study introduced a novel paradigm, Wide Learning, which revealed the formation of semantic networks in LTM after learning, further demonstrating the significance of semantic networks in natural behavior.

Funding for this research was provided by the Minducate Science of Learning Research and Innovation Center of the Sagol School of Neuroscience, Tel Aviv University, and by the Ariane de Rothschild Women Doctoral scholarship to K.T.

43.309 PUPILLARY SIGNATURES OF WORKING MEMORY CONTENT ARE FLEXIBLE TO VISUAL VS. SEMANTIC TASK DEMANDS

Yueying Dong[1], Yun-Chen Hung[1], Connie Xie[1], Anastasia Kiyonaga[1]; [1]UC San Diego

Prominent working memory (WM) models posit that transient goal-relevant information is distributed across activations in sensorimotor cortical regions. Newer evidence suggests that even peripheral physiology like the pupils can contain WM content information. For instance, darker WM items elicit greater delay-spanning pupil dilation than brighter items, even when sensory input is matched between conditions. This pupillary WM response is sensitive to the behavioral priority of the WM content and is greater for individuals with stronger self-reported visual imagery. This suggests that pupillary WM effects may be both flexible to behavioral goals and magnified when relying on more visual maintenance strategies. Here, we manipulate WM probe demands, during continuous eye-tracking, to test whether WM task strategy modulates the WM pupil effect. Participants completed a WM task for naturalistic scene images and were probed on either the precise visual features or the semantic category of the WM content. Participants encoded both daytime and nighttime WM sample images, followed by a retrocue indicating which item would be tested. The retrocue color also varied, to signal whether the visual or semantic item features would be probed. Pupil dilation was greater during the delay when a nighttime scene was prioritized (vs. daytime), corroborating prior work that used oriented Gabors of varying luminance as stimuli. Moreover, the WM-pupil effect here was larger when the task probed precise visual information. The effect diminished when the task probed higher level semantic categories, resembling the previously observed diminished response in individuals with weaker visual imagery. These findings demonstrate that peripheral sensorimotor responses can carry information not only about the visual properties of WM content, but also how it will be used. Pupillary WM responses are modulated by the upcoming WM task, pointing to a potential role for the earliest stages of visual processing in strategic and flexible WM maintenance.

43.310 REVERSE SPATIOTEMPORAL HIERARCHY DURING CROSS-MODAL MEMORY RETRIEVAL AND IMAGERY

Yu Hu[1,2], Yalda Mohsenzadeh[1,2]; [1]Western University, [2]Vector Institute for Artificial Intelligence

When we hear the sound of an object, it may trigger associated visual memory and imagery. This process involves perceptual, memory retrieval, and imagery activities. Based on previous research, this process may recruit sensory- and memory-related brain areas. However, the underlying spatiotemporal dynamics remain poorly investigated. Here, we used naturalistic videos of audiovisual events and recorded fMRI data during the tasks in which human participants (n=16) recalled visual contents when hearing associated sounds and recalled sounds when watching silent videos, after they well-memorized the video contents. With time-resolved fMRI multivariate pattern analyses, we observed reverse spatiotemporal hierarchy during the sound-induced visual memory retrieval: the neural activity started in auditory areas, then high-level visual areas, and finally reached the primary visual cortex. A similar trend was found during visual-induced auditory memory retrieval, where the activity of the high-level superior temporal gyrus was earlier than mid-level planum temporale. However, the primary auditory area was not involved, suggesting modality differences in the role of primary sensory areas in corresponding memory recall. We also observed the activity of the hippocampus, the parahippocampal cortex, the retrosplenial cortex, and the precuneus, which were involved in similar timing of imagery-related sensory areas. Overall, our study provided both spatial and temporal accounts of neural activity during the cross-modal memory retrieval and imagery.

43.311 PROBING VISUAL IMAGERY IMPLICITLY: PUPIL RESPONSES TO PREDICTED LIGHT AND DARK STIMULI

Xiaonan Li[1] (xiaonan.li@emory.edu), Moxuan Liu[1], Yaxin Liu[2], Stella F. Lourenco[1]; [1]Emory University, [2]Georgetown University

The (cognitive) pupillary light response has emerged as an index of visual imagery. When tasked with imaging dark and light scenarios (e.g., night vs. day sky), people's pupils dilate and constrict, respectively. As an assessment of visual imagery, pupillometry is advantageous because it bypasses the need for conscious reporting. Yet, even with pupillometry, explicit instructions or cued words are generally included to elicit visual imagery, which may confound imagery-related processes with other (non-imagery) processes such as metacognitive awareness. Here we develop a novel, implicit pupillometry-based paradigm, in which there are no instructions related to visual imagery. This paradigm leverages predictive learning during the occlusion of a moving object. In an initial familiarization phase, participants repetitively viewed a coin-shaped object that moved horizontally across the screen (e.g., left-to-right), while rotating between its front (bright) and back (dark) sides. In a subsequent test phase, the object rotated behind an occluder toward the end of its trajectory. On half the trials, the object rotated from dark to bright and, on the other half, it rotated from bright to dark, with luminance during occlusion matched across trials. Analyses of perception and imagery periods with 9 participants revealed robust pupillary light responses to dark and light stimuli. When tracking visible dark and light stimuli, pupil sizes exhibited dilation and constriction, respectively. Moreover, during the occlusion phase, baseline-corrected pupil sizes were significantly larger for the predicted dark stimulus than the predicted bright stimulus, suggesting that participants engaged in visual imagery consistent with the learned luminance associations. These findings support the validity of a novel implicit task of visual imagery in which imagery occurs without explicit instructions or conscious reporting. This task has the potential to provide a powerful tool for investigating visual imagery across diverse population, significantly advancing the study of visual imagery.

Visual Memory: Memorability

Monday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

43.312 EARLY VISUAL PROCESSING INFLUENCES MEMORABILITY: A CASE STUDY FROM NATURE

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

What makes certain images stick in our memory more easily than others? While memorability has been linked to high-level visual processing, the role of low-level vision is still not fully understood. In previous work, we explored vision and memory for the striking 'warning patterns' that toxic butterflies display to deter predators. Warning patterns evoked stronger and less sparse activity when exposed to a biologically inspired computational model of low-level vision, setting them apart from non-toxic species and other natural images. Here, we used our vision model to measure magnitude and sparseness of responses to natural and man-made textures from the Oxford Describable Textures Dataset. Using the same statistics that distinguish toxic from non-toxic butterflies, we selected sets of textures that evoked high vs. low activity and presented them to humans in a memory test. Observers (N = 100) viewed textures and reported on a 10-point scale their impression of how 'memorable' they appeared. Next, observers performed an old/new recognition test ('Seen before?'). We found that textures that evoked high neural activity were rated as more memorable (mean rating ± SE: high = 0.50 ± 0.03, low = 0.37 ± 0.03, p < .001) and were also more easily remembered (mean d' ± SE: high = 2.13 ± 0.19, low = 1.51 ± 0.16, p < .001) than those that evoked low activity. We compared human performance to a state-of-the-art deep learning model for memorability prediction, ResMem (Needell & Bainbridge, 2022, Comp Brain & Behav. DOI:10.1007/s42113-022-00126-5). The model was moderately predictive of both human ratings (Spearman's rho = 0.45, p < .001) and proportion correctly remembered (Speaman's rho = 0.30, p < .001). Our findings suggest that early visual computations, which distinguish toxic from non-toxic animals in nature, may also differentiate memorable images from forgettable ones in humans.

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

43.313 MODELING SUBJECTIVE AND OBJECTIVE FACE MEMORABILITY IN A COMMON LATENT SPACE

Nakwon Rim[1], Marc G. Berman[1], Stefan Uddenberg[2], Wilma A. Bainbridge[1]; [1]University of Chicago, [2]University of Illinois Urbana-Champaign

Faces are among the most important visual stimuli to remember when navigating the social world. Previous research has shown that some faces are more reliably remembered or forgotten than others, suggesting that faces have objective memorability. Interestingly, the subjective memorability of faces, or how memorable people "think" a face is, is positively correlated with objective memorability but does not capture the whole variance. This signals that there are both commonalities and dissociations between objective and subjective face memorability. To further explore the relationship between these two notions of face memorability, evaluating them in a common space is crucial. Here, we build models of objective and subjective memorability of faces using the latent dimensions of a Generative Adversarial Network (StyleGAN2) pre-trained on face images. We first collected objective memorability scores from 1,004 synthetic faces generated by StyleGAN2 via a continuous recognition task. As with past research using real-face images, the objective memorability of synthetic faces was reliable across participants. Furthermore, the subjective memorability of faces, acquired via slider ratings made by a separate group of participants, was positively correlated with objective memorability but did not account for the full variance. Building on this, we trained two models predicting objective and subjective memorability from the latent vectors corresponding to the synthetic faces. The two models showed comparable predictive power, measured by cross-validated R-squared, but differed in terms of the weights of the latent vectors. This implies that there are differences between objective and subjective face memorability that we can observe in the common latent face space. Importantly, these models can be also used to transform synthetic faces, allowing us to create a diverse array of face stimuli originating from the same starting point that can be used in further studies of both objective and subjective face memorability.

The research was supported by the National Eye Institute (R01-EY034432) to W.A.B.

43.314 SEMANTIC MEMORIES: IMAGE MEMORABILITY TRANSFERS TO SCENE DESCRIPTIONS

Diana Kollenda[1,2] (diana.kollenda@psychol.uni-giessen.de), Sophie-Luise Schwenger[1], Sophie Halstenberg[1], Benjamin de Haas[1,2]; [1]Experimental Psychology, Justus Liebig University, Giessen, Germany, [2]Center for Mind, Brain and Behavior (CMBB), Universities of Marburg, Giessen and Darmstadt, Germany

Some images are inherently more memorable than others, a quality that can be predicted with ResMem, a residual neural network sensitive to semantic image features. Large-scale human data further indicate that semantic features strongly contribute to the memorability of object images. This raises a compelling question: Is image memorability tied to visual features, or at least in part modality-independent? Here, we tested whether the memorability of an image generalizes to its verbal description. We prepared descriptions of 160 scenes from the OSIE dataset using two approaches: a) a standardized protocol used by one of the experimenters, and b) a multimodal transformer model (ChatGPT-4). We then assessed the memorability of these descriptions in independent groups of participants (n = 81 each). During the encoding phase, participants were tasked to categorize 80 descriptions as depicting either outdoor or indoor scenes. In a subsequent surprise memory task, they were presented with the same 80 descriptions, intermingled with 80 lures, and asked to indicate for each description whether they had seen it during the preceding encoding phase or not. Half of the targets and lures were descriptions of images with low ResMem scores (0.64 ± 0.04) and half descriptions of images with high scores (0.91 ± 0.02). Recognition sensitivity (d') was much higher for descriptions of high-memorability images, regardless of whether the descriptions were prepared by an experimenter (t = 10.8, p <.001, d = 1.2), or Chat-GPT (t = 8.67, p <.001, d = 0.96). Replicating the experiment with images instead of descriptions showed a comparable memorability effect (t = 11.3, p <.001, d = 1.25). The predictiveness of ResMem scores for description memorability held up when controlling for a range of image features. These results provide clear evidence that image-computable memorability is at least in part modality-independent.

This work was supported by ERC Starting Grant 852885 INDIVISUAL; BdH was further supported by DFG Project Nos. 222641018-SFB/TRR 135 TP C9 as well as "The Adaptive Mind", funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art.

43.315 IMAGE MEMORABILITY FACILITATES VISUAL WORKING MEMORY FORMATION: ELECTROPHYSIOLOGICAL EVIDENCE FROM CONTRALATERAL DELAY ACTIVITY

Chaoxiong Ye[1,2] (chaoxiong.c.ye@jyu.fi), Lijing Guo[1,2], Edward Awh[3], Edward Vogel[3], Qiang Liu[1,2,4], Weizhen Xie[5]; [1]Anyang Normal University, China, [2]University of Jyvaskyla, Finland, [3]University of Chicago, [4]Sichuan Normal University, China, [5]University of Maryland

Some images, such as faces and scenes, reliably embed themselves in memory, while others are easily forgotten. This phenomenon is often attributed to image memorability, a stimulus-driven property that predicts consistent memory responses across individuals. While this phenomenon is often studied in long-term memory, recent research suggests that image memorability also influences visual working memory (VWM). However, it remains unclear whether this influence stems from facilitated VWM formation or enhanced maintenance that extends VWM storage capacity. To address this, we recorded scalp EEG from 35 participants performing a change-detection task involving images of faces with varying memorability (high vs. low). While maintaining central fixation, participants briefly viewed two faces (100 ms) on one of the visual fields and, after a short delay, were tested for memory accuracy. Consistent with prior findings, high-memorability faces were associated with better behavioral performance. We further analyzed the contralateral delay activity (CDA), an ERP component associated with VWM formation and maintenance. Our results demonstrated that CDA amplitudes at posterior electrode sites were significantly larger for high-memorability stimuli during the early retention period (300–450 ms post-stimulus). However, this difference diminished during the later retention period (850–1000 ms post-stimulus). These findings suggest that high-memorability stimuli enhance early-stage encoding processes, resulting in more robust memory representations without altering the overall storage capacity of VWM. By identifying the specific processing stage where image memorability impacts VWM, this study highlights its role in facilitating VWM formation rather than maintenance.

This work was supported by the Research Council of Finland (former Academy of Finland) Academy Research Fellow project (#355369 to Chaoxiong Ye) and Finnish Cultural Foundation (#231373 to Chaoxiong Ye).

43.316 MEMORABILITY PREDICTS WIDESPREAD VIRALITY ON SOCIAL MEDIA

Shikang Peng[1,2] (speng@research.baycrest.org), Wilma A. Bainbridge[1]; [1]University of Chicago, [2]University of Toronto

Visual content on social media plays a pivotal role in global entertainment and information sharing, yet some images receive significantly more likes and comments than others, raising questions about perceptual features in the images driving such engagement. While emotional and moral content has been linked to viral posts, these factors vary across populations. We propose that a post's image memorability—its ability to be remembered—might predict its viral potential. Here we collected a fully naturalistic dataset of image posts from the well-known social media platform Reddit, across three separate timepoints (N=1,247 image posts). Leveraging memorability-trained ResMem, we obtained a predicted memorability score for each image. By correlating these with the virality metrics, we found that memorable images tend to attract more comments, replicated across the three timepoints. Even after accounting for the image category of the post using ImageNet pre-trained ResNet-152, we found that ResMem's predictions still explained unique variance in virality. A semantic analysis was then conducted on comments to investigate the mechanisms of this effect. We found that memorable images tend to elicit more neutral-affect comments, demonstrating memorability has a different pathway to virality from emotion. Further, a visual consistency measure using Google Vision AI and word embeddings showed that memorable image posts were associated with more diverse, externally-associated comments. To explore which characteristics drive memorability and virality, we analyzed image features through ResMem's convolutional layers. We found semantic distinctiveness was key, indicating a shared underlying mechanism between memorability and its viral potential. In sum, this study provides a detailed investigation surrounding the topic of 1) whether memorability can predict social media virality; 2) how and why memorability affects human online behavior leading to virality; and 3) what underlying visual features contribute to the viral potential of a post.

43.317 ARE IMAGES TRULY INTRINSICALLY MEMORABLE? MEMORABILITY RESULTS ARE EXPLAINED BY BETWEEN-IMAGE SIMILARITY STRUCTURE, NOT FEATURES OF INDIVIDUAL IMAGES

Dyllan Simpson[1], Benjamin Johnson[1], Timothy Brady[1]; [1]University of California San Diego

For decades, recognition memory models have successfully predicted human performance without incorporating any notion of intrinsic memorability, instead relying on similarity-based comparisons and decision criteria. Global matching models operate on the principle that memory decisions are a function of summed similarity to all stored items, plus a criterion for "old" responses. Despite this success, the notion of memorability as an intrinsic property of objects - independent of context and distinctiveness - has gained prominence in the field. In three behavioral experiments (N=352), we demonstrate how apparent memorability effects emerge from dataset composition and context. First, we manipulated the prevalence of supposedly memorable animate objects in the THINGS database, creating datasets that were either 90% or 10% animate images. When animate objects were overrepresented rather than rare, they became less memorable than inanimate objects ($p<0.001$), demonstrating that memorability reflects relative distinctiveness rather than intrinsic properties. If memorability depends on distinctiveness, it should vary with both encoding and test context. Our second experiment confirmed this by using the THINGS database but manipulating similarity during encoding: presenting "highly memorable" objects among similar items reversed established memorability patterns ($p=0.019$). Our third experiment manipulated test sequences while keeping encoding constant, showing that low-memorable objects tested among dissimilar items performed equivalently to high-memorable objects tested among similar items ($p=0.43$). Simulations using a global matching model (Shiffrin & Steyvers, 1997), captured all of these effects, as 'memorability' estimates decreased as objects' feature distributions became more homogeneous. These results suggest that reliable memorability effects emerge from the relative distinctiveness of features in experimental datasets rather than being intrinsic to images. Memory performance reflects the interaction between an observer's representational similarity space and the distributional structure of both study and test contexts.

43.318 MEMORABLE BY DESIGN: THE INTRINSIC PROPERTIES OF EFFECTIVE SYMBOLS

Brady R.T. Roberts[1], Wilma A. Bainbridge[1]; [1]University of Chicago

Recent work has begun to evaluate the memorability of everyday visual symbols (e.g., !@#$%) across individuals as a new way to understand how abstract concepts are processed in memory. Symbols are highly memorable, especially relative to words, but it remains unclear what drives this memorability. Across two experiments in the current study, we first identified which visual attributes predict memory for 80 conventional symbols, then we manipulated those features in the generation of novel symbols that were designed to be memorable or forgettable. To provide a metric of each symbol's memorability, 248 participants completed a continuous recognition memory test. This task revealed that certain symbols were reliably more memorable than others across participants. Next, participants sorted symbols spatially based on their visual similarities. A principal components analysis on the Euclidean distance between each pair of symbols revealed three visual dimensions participants used to sort symbols: thin to thick, straight to curvy, and good to poor vertical symmetry. All three of these principal components were then validated using computer vision analyses, and each uniquely predicted memory for symbols in a multiple regression. Next, generative artificial intelligence was used to create a set of novel symbols while accentuating or downplaying these predictive visual features to create memorable and forgettable symbol variants for a list of abstract words, respectively. Memory was tested across 329 participants, revealing that performance (as measured by both recognition and cued recall) was substantially improved for symbols that were designed to be memorable. Put simply, not only were symbols that were designed to be memorable easier to recognize on a memory test, but their links to associated abstract words were also improved. This work demonstrates for the first time that certain visual features drive memory for symbols and offers clear evidence that memory can be intentionally engineered.

This research was supported by a Natural Sciences and Engineering Research Council (NSERC) of Canada Postdoctoral Fellowship to BRTR and National Eye Institute Grant R01-EY034432 to WAB.

43.319 THE PROPAGATION OF MEMORABILITY IN BINDING MEMORY

Jingyao Cheng[1] (chengjingyao690@gmail.com), Yingtao Fu[1], Mowei Shen[1], Hui Chen[1]; [1]Zhejiang University

Memorability refers to an intrinsic property of stimuli that influences their likelihood of being remembered or forgotten. Previous studies have consistently replicated the memorability effect and demonstrated that this effect is closely tied to the perceptual properties of stimuli. However, memory is not merely a collection of isolated entities but a complex system organized through the relationships between individuals. The present study aims to explore the distinct memorability effect on relational memory beyond individual items, encompassing time-related, space-related, and inter-item binding memory. Participants completed a sequential working memory task in which the memorability of scene images (memorable or forgettable) varied trial by trial. They were instructed not only to remember which item appeared but also to retain the temporal binding (Experiment 1), spatial binding (Experiment 2), and image-word binding (Experiment 3), allowing for the separate assessment of item memory and binding memory. We first replicated the effect of memorability on item memory, with memorable ones being remembered better than forgettable ones. More importantly, binding memory for memorable items was consistently better at all serial positions and across all types of binding. In Experiment 3, regardless of whether the cue recognition was image-based or word-based, memorable images showed superior binding memory, while the item memory for words was unaffected by memorability. Our findings provide robust evidence that memorable items are not only better remembered in terms of item memory, but also exhibit enhanced binding memory. This suggests that memorability is not merely a perceptual property confined to the item itself, but a mnemonic property that propagates item-related information.

43.320 AN EVENT SEQUENCE TO REMEMBER: ABSTRACT TEMPORAL STRUCTURE INFLUENCES MEMORABILITY

Niels Verosky[1], Brian Scholl[1]; [1]Yale University

Much work in recent years has explored how some stimuli are intrinsically more memorable than others, often due to distinctive sensory or semantic features. But what determines memorability beyond such features, e.g. in sequences of musical tones? Here we explored the possibility that memorability also depends on abstract temporal structure, in both vision and audition. Participants were presented with short temporal sequences of three items varying along a scalar dimension--either three successive tones of different pitches (in auditory experiments) or three successive circles of different sizes (in visual experiments). Items were sampled from five fixed points along the relevant scalar dimension, creating a combinatorial space of 60 possible sequences in each modality, with a one-to-one mapping between analogous auditory and visual sequences. (Examples of possible sequences would thus be 1-2-3, 1-3-2, 1-3-5, and 5-2-4--with the items 1-5 mapping onto either the pitches of tones or the sizes of circles.) To comprehensively characterize memorability, we presented each participant with sequences that were randomly sampled from the full combinatorial space. During stimulus presentation, participants made an orthogonal judgment with no mention of memory. Subsequently, they completed a surprise recognition task. Though all sequences were constructed from the same small library of meaningless items, some were nevertheless reliably more memorable--and some particular sequences were exceptionally memorable across both modalities. For example, among all sequences in the combinatorial space, those that stood out as especially memorable were monotonically increasing sequences consisting either of three consecutive items (1-2-3, 2-3-4, 3-4-5) or of the lowest-magnitude plus the two consecutive highest-magnitude items (1-4-5). Intriguingly, this same pattern emerged independently for both vision and audition. This work thus adds a new dimension to the study of memorability: beyond distinctive sensory and semantic properties, what we incidentally remember is also shaped by abstract temporal structure.

Visual Search: Attention, clinical

Monday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

43.321 ADAPTATION OF ATTENTIONAL CONTROL: THE IMPACT OF DISTRACTOR PREVALENCE ON DISTRACTOR LOCATION LEARNING

Mustafa Zeyd Söyük[1], Anna Schubö[1]; [1]Philipps-University Marburg

Attentional control strategies adapt to visual field regularities like distractor prevalence or spatial patterns. Previous research has shown that frequent exposure to distractors reduces their interference, and spatial learning can tune down interference by deprioritizing high-probability distractor locations. So far, these effects have been studied separately. Here, we investigated how the extent of distractor practice influences distractor location learning in an additional singleton search task. We expected that performing a search task with high distractor prevalence will result in efficient distractor handling, which might reduce the need for additional distractor location learning. Participants completed a training and test phase, searching for a diamond-shaped target among homogenous non-target items, with a color distractor appearing in some trials. During training, distractor prevalence was 20% for the low-prevalence group and 80% for the high-prevalence group. In the test phase, distractor prevalence was equalized (66%), and distractors were more likely to appear at a high-probability location in both groups. As expected, the low-prevalence group experienced larger distractor interference during training than the high-prevalence group. In the test phase, both groups showed less distractor interference for distractors at the high-probability location (compared to low-probability locations). Interestingly, the high-prevalence group showed no impairment in target detection at high-probability location when the distractor was absent, whereas the low-prevalence group exhibited impaired target detection at this location under the same conditions. These differences in target processing at the high-probability location might result from the varying amount of distractor practice the groups experienced during the training phase. We assume that during training, the high-prevalence group developed a control strategy to selectively ignore the distractor whereas the low-prevalence group did not and, as a consequence, broadly suppressed any signal at the high-probability location. Taken together, our findings demonstrate that attentional control strategies are adapted to environmental demands.

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation – project number 290878970-GRK 2271, project 9).

43.322 AGE-INVARIANT BENEFITS OF PREDICTIONS IN DYNAMIC VISUAL SEARCH WITH VARYING DISTRACTION LOADS

Nir Shalev[1] (shalenir@gmail.com), Sage Boettcher[2], Anna C. Nobre[3]; [1]Haifa University, [2]University of Oxford, [3]Yale University

Visual search tasks are widely used to study how people guide their attention amidst distractions. Participants are typically asked to find a target displayed among distractors in a static display, with attentional demands varied by increasing the number and similarity of distractors. In lifespan developmental research, visual search has been used to investigate age-related differences in attentional control. Typical findings indicate a performance decline from younger to older adulthood, alongside elevated sensitivity to added distractors. Such findings are often interpreted as evidence of declining goal-driven attentional abilities. However, previous studies comparing younger and older adults rely almost exclusively on static visual scenes. Real-life scenarios, in contrast, are constantly changing and often embed regularities that can be learned and utilised to optimise performance, reducing reliance on purely top-down guidance. We investigated how people of different ages guide searches in a dynamic display under varying levels of distraction. The Dynamic Search Task is an extended variation of a visual search task, incorporating spatiotemporal regularities that afford prediction-led guidance. Critically, we manipulated the number of visual distractors presented on each trial to study whether the distraction load interacts with search efficiency and predictions. Younger adults (ages 18–22; N = 120) and older adults (ages 68–72; N = 120) searched for multiple visual targets as they continuously faded in and out of the display among distractors (either 12 or 24 per trial). Half the targets appeared at consistent times and approximate locations throughout. Results showed lower accuracy and heightened distractor interference in older adults. However, target predictability conferred behavioral benefits regardless of age. Interestingly, in both groups, the benefits of predictions were significant only during periods of high distraction. This study enhances our understanding of age-related attentional control in dynamic contexts and supports a distinction between memory-driven and goal-directed guidance.

N.S is supported by the Israeli Science Foundation Research Grant (1073/24); A.C N Is supported by a Wellcome Trust Senior Investigator Award (104571/Z/14/Z), James S. McDonnell Foundation Understanding Human Cognition Collaborative Award (220020448)

43.323 BEHAVIORAL PERFORMANCE AND TASK-EVOKED PUPILLARY RESPONSE INDICES OF ERRORS DURING VISUAL SEARCH

Sylvia Guillory[1], Hiu Mei Chow[2], Beatrix Culligan[2], Sierra Nastasi[1,3], Chad Peltier[1,3], Jeffrey Bolkhovsky[1]; [1]Naval Submarine Medical Research Laboratory, [2]St Thomas University, [3]Leidos, Inc

Cognitive demands and uncertainty in decision-making can influence task-evoked changes in pupil size (e.g., increase with working memory load). Uncertainty can arise from internal or external sources, such as lapses in attention or distractions. These factors unrelated to task instruction can influence uncertainty and lead to performance errors. A goal of this experiment was to examine the relationship between changes in pupil size and uncertainty in attentional selection. We hypothesized that reduced pupil size would be associated with greater uncertainty/task-disengagement, as indexed by poorer task performance. Observers (n=10) participated in a visual search task in which a target letter "T," presented at a 10% prevalence rate, appeared among 25 non-target offset "Ls." Eye positions and pupil diameter were sampled using a video-based eye-tracker. For each trial, observers made a response within seven seconds of stimulus onset to report whether the target was present or absent. No feedback was given. Trial type was categorized based on target presentation and observer response (hit, miss, correct rejection, false alarm). Results showed there was a significant difference in pupil diameter based on trial type (p<.05) suggesting that task-evoked changes in pupil diameter may reflect uncertainty (i.e., errors). Specifically, the average pupil diameter across the stimulus presentation period was 1.1 times larger in trials with hit vs. miss responses, and 1.05 times larger in trials with correct rejection vs. miss responses. This pupillary pattern corresponded with task performance (e.g., shorter reaction time in trials with hit vs. miss responses (p<.013). Lapses in attention can be detrimental in professions requiring extended periods of watch, such as the military operations or long-distance driving. These findings indicate that pupil diameter may serve as an index of uncertainty during attentional tasks, providing the information needed to optimize performance by identifying operational safety thresholds in terms of lapses in attention.

This work is supported by the Defense Health Program Research, Development, Test, and Evaluation Restore Core Funds (restoral) under funding work unit number F1710 awarded to JB/SG, as well as a Natural Sciences and Engineering Research Council of Canada Discovery Grant (RGPIN-2024-06028) to HMC.

43.324 ERROR-PRONE STATES IN VISUAL SEARCH

Jeremy Wolfe[1], Jeunghwan Choi[2]; [1]Brigham and Womens Hospital / Harvard Med, [2]Graduate Program in Cognitive Science, Yonsei University

In visual search, why do people miss targets that are clearly visible? In a search for a T among Ls observers will reliably miss 5-10% of targets. When "retrospectively visible" targets are missed in tasks like breast cancer screening, radiologists can be sued. For TvsL search, Li et al (2024) found errors to be largely random with respect to the specific stimulus. That is, missing the target on one trial tended not to alter the probability of missing the same target in the same display a second time. Of course, you are more likely to miss hard-to-see targets, but if some class of stimuli produces, say, 20% errors, the specific 20% seems to occur randomly. What about the state of the observer? Certainly, errors fall with learning and rise with fatigue, but within a relative steady-state, does anything modulate the probability of error? Bruno et al (2024) have proposed that there is an electrophysiologically identifiable brain state (perhaps related to 'mindwandering') that is associated with being temporarily more error prone. Introspectively, it sometimes feels like errors occur in clumps. If this is the case in visual search, then the probability of another error should be elevated after an error. To assess this, we reanalyzed several visual search datasets. We computed the distribution of lags between successive errors. Random production of errors predicts a geometric distribution of lags. If one error marks entry into an error-prone state, then the likelihood of another error should rise after the error. Chi-sq tests reveal significant over-representation of errors shortly following other errors, especially for tasks without feedback. This effect appears to be smaller or non-existent for experiments where feedback may disrupt any error prone state. Of course, in the real world, reliable error feedback is often lacking.

Grant support from NSF 2146617 & NIH-NEI EY017001

43.325 HOW CONSISTENT ARE YOU? IDIOSYNCRATIC POLAR ANGLE BIASES IN VISUAL SEARCH FOR DIFFERENT STIMULI.

Cailey Tennyson[1] (ctennyson1@bwh.harvard.edu), Injae Hong[1], Rawan Ne'meh[2], Jeremy Wolfe[1,3]; [1]Brigham and Women's Hospital, [2]Hariri High School 2, Beirut, Lebanon, [3]Harvard Medical School

In a 250-300msec fixation, you can attend to ~4-6 nearby items. If there are more candidates in the neighborhood, choices must be made. In prior work, observers moved their eyes to a fixation spot when it moved to a new location. A ring of one T and seven Ls was flashed around the spot after 300 msec. Observers made 4AFC assessments of the T's orientation. Flash duration was adjusted to produce ~25% errors. There was an average non-uniformity in the distribution of errors with more errors on the vertical meridian. More interestingly, there were often significant idiosyncratic deviations from that average. To find out if patterns of deviations were stimulus-specific, we repeated the TvsL condition and added a search for a hammer among seven different tool silhouettes. Observers ran 360 trials for each task in four, 180 trial sessions. The TvsL condition replicated prior results with 17 of 20 observers showing significant deviations from the average result. Moreover, patterns were quite consistent with an r=0.6 correlation of session 1 with session 2. Tools showed a more dramatic average deviation with ~26% of all errors at the bottom of the ring of locations. Only 10 of 20 observers showed significant idiosyncratic deviations from this pattern. Observers were strongly consistent between sessions (r=0.73). There was also a reasonable correlation between TvL and Tool tasks (r=0.39). We conclude that there are idiosyncratic variations in the deployment of attention in the vicinity of the current point of fixation. These appear to be quite different from the inhomogeneities in sensitivity to visual stimuli (e.g. Himmelberg, Winawer, & Carrasco, TINS, 2023). The choice of stimuli modulates the patterns of errors for reasons that are not entirely clear. These idiosyncratic relative attentional blindspots at different locations relative to fixation could contribute to errors in visual search.

NEI EY017001, NSF 2146617

43.326 INCREASED PHYSICAL EFFORT REDUCES MISS ERRORS IN VISUAL SEARCH

Jeff Moher[1], Caitlin Boyd[1], Alexis Henry[1], Macey Roberts[1], Dasha Zdvizhkova[1]; [1]Connecticut College

People often miss targets in visual search, even in relatively simple search tasks. Changes in the design of search interfaces that reduce these miss errors could be valuable, as errors in domains such as medical image reading or x-ray baggage screening carry high costs. In the present study, we manipulated the physical effort required to indicate target presence or absence in a simple visual search task. Prior research has demonstrated that anticipated motor effort can modify higher-level cognitive processes. In this case, we were interested in whether increasing the motor effort required to execute a target absent response would motivate participants to search more thoroughly, and thus reduce miss errors in visual search. Participants searched for a vertical line among tilted lines in a simple display. Targets were present on a randomly selected 50% (high prevalence) or 10% (low prevalence) of trials; prevalence was randomly assigned to participants for each experimental session. Across two blocks of trials, participants completed two distinct versions of the task. In one version they had to reach to the top of the display (more effort) to indicate target absence and the bottom of the display (less effort) to indicate target presence. In the second version, those response mappings were flipped. Eye and hand movements were tracked throughout the experiment. We found that when more effort was required to indicate target absence, miss rates were indeed lowered. This result suggests that a simple manipulation – increasing the effort required to indicate target absence – can modify search strategies and reduce miss errors. Data collection is ongoing, and subsequent analyses will allow us to examine the impact of target prevalence as well as the role of eye movements. These preliminary results highlight the value of integrated approaches to the study of vision and action.

This work was supported by National Science Foundation grant BCS-2218384 andNational Institute of Neurological Disorders and Stroke grant R15NS113135 to JM

43.327 SEARCHING IN SUMMARY: MEAN AND RANGE ENSEMBLE STATISTICS GUIDE ATTENTION IN VISUAL SEARCH TASKS

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

Our visual environment uses ensemble perception to summarize multiple sources of information into a single summary statistic, helping overcome capacity limitations. Our study explores how ensemble representations guide attention in a subsequent visual search task, and the precision of these representations relative to the mean orientation of a set of items. To explore this, we integrated an orientation-based ensemble task with a visual search task. Each trial began with an ensemble display of eight bars with varying orientations. Then, on a subsequent blank screen, the color of a fixation cross signalled which of two possible tasks —a search or an ensemble task— participants would perform. For the search task (75% of trials, orange cross), participants saw a display of randomly oriented and sized bars and identified whether a T or L appeared inside the shortest bar. Importantly, the target bar's orientation could either exactly match the average orientation, be within the range, or be outside the range of the initial ensemble display. In the ensemble task (25% of trials, blue cross), participants selected which of two bars matched the average orientation of the initial ensemble display. The results revealed shorter response times (RTs) in the search task when the target bar's orientation either matched the average orientation or fell within the range of the initial ensemble display, compared to the orientation that was outside of the range. These results suggest that ensemble processing guides attention a visual search task, and does so via precise representations of different summary statistical properties (i.e., the mean and range of a distribution). Together, these findings highlight the prioritization of ensemble representations in everyday behavior (i.e., visual search), and provide insight into the interaction between different perceptual and cognitive systems.

43.328 THE EFFECT OF PREVALENCE IN A CONTINUOUS MONITORING TASK WITH DYNAMIC SCENES AND TARGET EVENTS

Michael L. Paavola[1], Cathleen M. Moore[1]; [1]University of Iowa

There are situations in the real-world where people continuously monitor dynamic scenes for rare target events, such as in lifeguard surveillance. Standard visual search studies, in which participants search within static displays for targets defined by features, and which use discrete-trial designs with each ~2sec trial either containing a target or not, have shown that people tend to respond more slowly and make more errors with low-prevalence (LP) targets than higher-prevalence (HP) targets (Low-Prevalence Effect, LPE). We asked whether these findings generalize to conditions in which the participants continuously monitor dynamic displays over an extended period for target events, which requires the detection of spatiotemporal patterns. In addition, inspired by a practice in the aquatics community, we asked whether adding a probe-detection task would mitigate the LPE. Displays consisted of 8 to 20 gray disks continuously cycling between high and near-threshold contrast at random phases. Participants monitored for circles that remained at near-threshold contrast for longer than the standard time. They pressed the spacebar when a target event was detected and indicated its location using the mouse. For different groups of participants, target events occurred with a probability of .25 (LP) or .75 (HP). Additionally, a red circle occasionally flashed somewhere within the display. Half the participants were to respond to these by pressing the space bar and indicating its location. The other half were told nothing about them. Results confirmed an LPE. The probe task reduced the LPE, but did so by increasing the miss rate for HP targets rather than reducing it for LP targets. Finally, the addition of the probe task also reduced the standard effect of set size on response time. These results provide initial insight into how dynamic continuous monitoring for event targets is similar to and different from static visual search in discrete-trial designs.

43.329 THE EFFECT OF VISUAL AND VERBAL CUEING ON SEMANTIC BIAS ACTIVATION AND TARGET SELECTION

Mahsa Zafarmand[1], Xinying Zou[2], Nancy Carlisle[3]; [1]Lehigh University, [2]Lehigh University, [3]Lehigh University

Traditional models of visual attention emphasize physical features like color, orientation, and shape in guiding target selection. However, real-world scenes and memory involve more complex interactions, including semantic relationships between objects. These relationships, learned over time, significantly influence attention, often disrupting performance when semantically related distractors (SRDs) are present. Research has shown that SRDs typically prolong response times (RTs) and decrease accuracy by competing for attentional resources. While visual search tasks depend on prior knowledge of the target, optimal performance requires a clear mental representation. Research suggests that visual cues are more effective than verbal cues, reflecting the visual system's preference for direct visual inputs. While semantic studies predominantly use verbal cues and visual search research relies on visual cues, this study bridges the two domains to examine how cue specificity impacts semantic biases and attentional processes. We hypothesized that the influence of SRDs may depend on the specificity of the target cue (visual and verbal).We manipulated cue type (verbal vs. visual), target presence (target present vs. absent), and SRD presence (SRD present vs. absent) for twenty-eight participants. Results demonstrated that visual cues significantly enhanced RT and accuracy compared to verbal cues. Target presence facilitated faster RTs and improved accuracy, while SRD presence slowed RTs and reduced accuracy. SRDs had a more pronounced effect in target-absent conditions, and this effect was stronger for verbal cues than visual cues. Although the three-way interaction for RT was non-significant, it reached significance for accuracy, highlighting the larger SRD effect in verbal cue conditions.In conclusion, visual cues proved superior in guiding attention during search tasks, emphasizing the visual system's reliance on direct representations. The findings reveal that SRDs exert greater influence on accuracy in verbal cue scenarios, underscoring the role of cue type in activating semantic relationships and shaping attentional dynamics.

43.330 DO YOU REALLY NEED THAT CUP? EFFECT OF CAFFEINE ON VISUAL ATTENTION AND LEARNING

Sojung Youn[1] (syoun@tamu.edu), Jenna Glotfelty[1], Ming-Ray Liao[1], Brian Anderson[1]; [1]Texas A&M University

Caffeine is commonly used to improve alertness, and has been shown to impact memory and tasks involving the attention network. However, the influence of caffeine on visual search, attentional capture, and learning-dependent attentional biases is not known. Here, we studied the effects of caffeine on visual search and attentional learning in university-aged participants across two experiments. For both experiments, following caffeine abstention, participants were randomly assigned to receive dissolvable caffeine strips (200mg) or placebo breath strips. Experiment 1 involved a visual search task where participants were asked to identify the rotated direction of the target letter "T" among multicolored "L" distractors. The target could be rendered in a frequent or an infrequent color, and the search trials varied in set size (5 to 50 items). Overall, participants were faster to report the frequent colored target. The two experiment groups exhibited comparable search slopes and frequent color benefits, while the caffeine group exhibited a smaller intercept, indicative of faster target identification and responding. In Experiment 2, participants completed an additional singleton task in which they searched for a shape singleton target. A uniquely-colored distractor was presented on a subset of trials, which could be rendered in a frequent or infrequent color. Similar to the first experiment, participants overall benefited from learning the statistical regularities, but caffeine did not modulate this benefit. Neither did caffeine influence the magnitude of attentional capture or overall response speed. Collectively, it appears that caffeine has little influence on statistically learned attentional biases or the efficiency of distractor filtering. Our data are consistent with an effect of caffeine on the process of target identification, with Experiment 2 suggesting that this benefit is not reducible to faster motor responses.

43.331 VISUAL SEARCH AND ASD: BASIC VS. SUPERORDINATE CATEGORY SEARCH

Safaa Abassi Abu Rukab[1] (safa.abassi@gmail.com), Shaul Hochstein[1]; [1]Hebrew University, Jerusalem, [2]ELSC Edmond & Lily Safra Center for Brain Sciences, [3]Life Sciences Inst

Visual search has been classified as easy feature search, with rapid target detection and little set-size dependence, versus slower difficult search, focused attention and set-size dependent speed. Reverse Hierarchy Theory attributes these classes to rapid high-cortical-level vision at a glance versus low-level vision with scrutiny, attributing easy search to high-level representations. Accordingly, faces "pop out" of heterogeneous object photographs. Individuals with Autism Spectrum Disorder (ASD) have difficulties recognizing faces, and yet we found this disability doesn't disturb face search. We now explore visual categorization abilities in individuals with ASD, focusing on differences between basic-level and superordinate-level tasks. Participants completed a visual search task, identifying objects from superordinate categories such as animals, food, body parts and furniture, versus basic level categories such as dogs, apples, arms, chairs, and we analyzed reaction times (RT) and their set-size slopes for both categorization levels. We find that individuals with ASD encounter greater challenges when performing superordinate categorization compared to basic categorization. Superordinate-level search RTs increase more steeply with set size (92 ms/item), indicating higher cognitive demands as task complexity increases. Conversely, basic-level categorization shows less pronounced RT increase (50 ms/item), suggesting more automatic, less effortful cognitive processes. Finally, variability in individual performance, particularly at the superordinate level, highlights participant-specific influences on cognitive processing speed and suggests that different individuals with ASD may experience varying levels of cognitive challenge during complex tasks.

Israel Science Foundation (ISF)

43.332 THE LOCUS FOR EYE MOVEMENTS AND ATTENTION IN MACULAR DEGENERATION

Andrew Freedman[1] (afreedman@ski.org), Preeti Verghese[1]; [1]Smith-Kettlewell Eye Research Institute

The healthy fovea acts as an oculomotor reference for the control of eye movements and is closely coupled with attention, with the goal of an upcoming saccade showing greater sensitivity. Individuals with macular degeneration (MD) develop a preferred retinal locus (PRL) for fixation, but it is unclear to what extent the PRL adopts the control and attentional role of the fovea. Utilizing a naturalistic search task, we evaluated whether participants with MD employed their PRL in the same manner as controls with an intact fovea. Two participants with MD and 5 controls with intact vision searched for generic cylindrical targets among similar distractors, distributed throughout a realistic indoor scene, while their gaze was tracked. Participants visually searched the scene and indicated when they detected the target. Target contrast, eccentricity, and whether the border was enhanced for visibility, and number of distractors were randomized across trials. Targets and distractors were twice the minimum size necessary for participants to discriminate between them. Participants were shown a preview of the target before each trial. Our results show that fixations of MD participants using a PRL were centered on the target, with a spatial distribution of 11.82 – 22.4 degrees squared, similar to that of controls using the fovea (4.81 – 30.42 degrees squared). The temporal distribution of fixations relative to the detection response was also similar for MD and controls. For all participants, fixations immediately prior to reporting a detection were significantly more likely to be on target, compared to off target, Tx1, x2 = 3.2528 – 9.3664, all $p \leq .001$. Thus, the PRL mirrored how the fovea was used by controls, suggesting that the PRL is utilized as an oculomotor reference point during visual search in MD, adopting the close coupling with attention typically associated with the fovea.

Funding: This work was supported by the NIH grants T32EY025201 and R01EY027390

43.333 COMPARING OBJECT LOCALIZATION DEFICITS IN DESKTOP AND IMMERSIVE VIRTUAL REALITY SEARCH TASKS IN INDIVIDUALS WITH CEREBRAL VISUAL IMPAIRMENT

Madeleine Heynen[1] (mheynen@meei.harvard.edu), Kerri Walter[2], Jonathan K. Doyon[1], Peter J. Bex[2], Lotfi B. Merabet[1]; [1]Mass Eye and Ear, Harvard Medical School, [2]Northeastern University

Cerebral visual impairment (CVI) is a brain-based visual disorder associated with early neurological injury and maldevelopment of visual processing pathways and areas. Functional visual deficits related to impaired object localization are commonly observed in CVI. Virtual reality (VR), paired with eye-tracking, offers an objective method to assess the nature of these higher-order visuospatial processing deficits. We characterized gaze behavior associated with target recognition abilities during desktop and immersive VR-based versions of a visual search task using a screen-mounted eye tracker (Tobii Spark 60 Hz) and Meta Quest Pro integrated eye tracking (100 Hz), respectively. Twelve CVI (19.67 years ± 6.39 SD) and 10 control (21.60 years ± 7.59 SD) subjects were instructed to search and locate a pre-identified target within a series of rooms with low and high clutter to investigate the effect of environmental complexity on the exploration of naturalistic scenes. Comparing success rates, there was a significant effect of group, with the CVI group showing an impaired ability to correctly locate the target ($F(1,60)=23.67$, $p<0.001$). Analyzing reaction times showed significant main effects of group ($F(1,60)=45.07$, $p<0.001$), display type ($F(1,60)=87.401$, $p<0.001$), and clutter level ($F(1,60)=6.43$, $p<0.05$), with a significant interaction between group and display type ($F(1,60)=10.07$, $p<0.01$). Both groups exhibited longer reaction times during the immersive VR task and in high-clutter conditions. Under all conditions, the CVI group showed longer reaction times than controls. Finally, fixation patterns for both groups exhibited a larger search space under high clutter conditions ($F(1,54)=14.34$, $p<0.001$), with the CVI group demonstrating larger search spaces than controls ($F(1,54)=21.58$, $p<0.001$). These findings highlight visuospatial deficits in CVI individuals, who exhibit slower and less accurate object localization and broader search patterns. Greater reaction time differences between groups in the VR condition suggest that increased immersion may better capture the higher-order visual perceptual challenges faced by this population.

Face and Body Perception: Parts and wholes

Monday, May 19, 8:30 am – 12:30 pm, Banyan Breezeway

43.334 USING SIX BEHAVIORAL TESTS TO INVESTIGATE THE LEFT VISUAL FIELD ADVANTAGE (OR LACK THEREOF) IN FACE PERCEPTION

Elizabeth (Shuxuan) Li[1] (lielizabeth63@gmail.com), Marie-Luise Kieseler[1], Antonio Mello[1], Brad Duchaine[1]; [1]Dartmouth College

The conventional wisdom among face perception researchers is that right-handed, neurotypical adults perform better when faces are presented in the left visual field (LVF) than the right visual field (RVF). However, our review of studies that compared performance between the two visual fields suggests that a file drawer problem may have influenced the literature. The distribution of p values in the studies was bimodal; out of 58 experiments, 27 reported a statistically significant LVF advantage and 6 found a significant RVF advantage. Only 25 experiments had p values between the significance thresholds for an LVF advantage and an RVF advantage, and these values were not clustered close to the thresholds. To investigate visual field asymmetries in face perception in an unbiased manner, we designed six tasks using the divided visual field paradigm as part of our Registered Report: Mooney face detection, two-tone face versus face parts task, sex classification, expression classification, age classification, and famous face identification. Initial data (N = 60) demonstrated that LVF minus RVF accuracy was 0.3%, 4.3%, 0.1%, -1.9%, -1.1%, and -1.2% for the six tasks. Neither accuracy nor RT significantly differed between visual fields for any test. Moreover, we found no effect of participant sex on visual field differences. These initial results suggest that the left visual field advantage is not as strong as the literature indicates, and we will further assess this hypothesis by testing 420 participants prior to VSS, as outlined in our Registered Report. Our study is an initial step toward developing an accurate understanding of visual field asymmetries (or lack thereof) in face processing, and we are confident that this goal can be achieved via pre-registered studies from multiple labs.

The project was supported by the Stamps Family Charitable Foundation, the Presidential Scholarship from Dartmouth College, and the Benjamin J. Benner '69 Undergraduate Research Fellowship.

43.335 REVERSE CORRELATION IMAGES AND REAL FACES: INSIGHTS FROM DEEP LEARNING-BASED FACE RECOGNITION

Yoshiyuki Ueda[1]; [1]Kyoto University

Face memory is highly susceptible to distortion due to various factors, and recognition tests are often used to investigate it. Although these tests can reveal whether memories are preserved, they cannot clarify how they are distorted. The reverse correlation image classification method (Dotsch & Todorov, 2012) offers a potential solution. In this paradigm, participants memorize a face and then select the face most similar to the one they memorized from a pair of faces generated from the base face image, with inverted polarity noise added. By aggregating the noise patterns from the selected faces, it is possible to construct classification images (CIs) that reflect the features participants used in their judgments in a bottom-up manner, and these CIs are considered to represent participants' mental images. However, it is currently up to the subjective judgment of a third person to decide whether the facial features of a CI correspond to those of the actual memorized face. This study applied a deep learning-based face recognition model (Serengil & Özpınar, 2024) to CIs to assess their similarity to memorized faces. The results showed that the similarity between CIs and memorized faces became stable after completing more than 200 trials, and CIs included features of the memorized faces. However, significant individual differences and variations due to the specific faces used in the experiment were observed. Furthermore, the group CI for one group was more similar to the memorized face compared to the other group that memorized a different face, but the face model was not always able to correctly identify which face had been memorized. These findings highlight the limitations of the current method and suggest the need for more elaborate techniques to visualize our representations of faces.

43.336 EYE DOMINANCE AND THE PERCEPTION OF EYE CONTACT

Colin Palmer[1] (c.palmer@nus.edu.sg), Gwenisha Liaw[1], Shui'er Han[2,3]; [1]National University of Singapore, [2]Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore, [3]Centre for Frontier AI Research, Agency for Science, Technology and Research, Singapore

Eye contact is an essential cue in social interactions, yet the physical conditions under which eye contact is felt can vary. For example, viewers can perceive a range of gaze directions, falling across different features of their face, as making 'eye contact' with them, with the position and width of this range varying between individuals. Here, we investigate how characteristics of binocular vision influence the perception of eye contact. Participants (n = 40) made judgements about when they shared eye contact with faces that were rendered in a Virtual Reality (VR) environment that simulates left- and right-eye perspectives onto a 3D scene, in both binocular and monocular viewing conditions. Across trials, the horizontal gaze direction of the faces varied to fixate a range of locations distributed across each participant's face. The gaze directions perceived as making eye contact differed depending on whether participants viewed faces with their left eye, with their right eye, or binocularly. In monocular viewing conditions, gaze directed towards the participant's open eye was most often perceived as making eye contact, while in binocular viewing conditions, gaze directed in between the two eyes was more-often perceived as making eye contact. In general, binocular vision can be weighted towards information received from a specific eye (i.e., the viewer's dominant eye). However, in the current study, the specific gaze directions that viewers perceived as making eye contact with them did not appear to vary with individual differences in sensory eye dominance (measured using an interocular suppression technique) nor sighting dominance (measured with the card test). Future research might investigate the relationship between eye dominance and gaze perception further in populations with more profound imbalances in binocular vision, such as amblyopia.

This research was supported by an NUS FASS Start-Up Grant awarded to CP.

43.337 A COMPARISON OF IMAGE NATURALNESS PERCEPTION BETWEEN HUMANS AND IMAGE GENERATIVE MODELS

Taiki Fukiage[1] (t.fukiage@gmail.com); [1]NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, Japan

Image generative models, particularly diffusion models, have rapidly advanced, producing images increasingly indistinguishable from real photographs. By learning from large datasets of natural images, these models may capture features that humans perceive as "natural." This study examined how closely their assessments align with human perception through two experiments: one employing the Thatcher illusion and another focusing on lighting inconsistencies. In the first experiment involving the Thatcher illusion, we modified 140 face images by flipping the eyes and mouth. Human participants rated the naturalness of original and modified images in upright and inverted orientations using a five-point scale. Two diffusion models (Stable Diffusion v1.5 and Stable Diffusion XL) estimated naturalness using the variational lower bound of log-likelihood. For both humans and models, we computed "unnaturalness scores" as the difference in ratings or likelihoods between the modified and original images. As a result, both humans and models demonstrated the Thatcher illusion, showing significantly higher unnaturalness scores for upright than for inverted images. Moreover, significant correlations between human and model unnaturalness scores were found within both upright (r=0.52-0.59) and inverted conditions (r=0.29-0.49), suggesting reasonable alignment in perceived facial naturalness. In the second experiment, we rendered scenes with two or three objects using a physically based renderer. We introduced lighting inconsistencies by swapping object positions or flipping objects, resulting in 288 pairs of original and modified images with varying degree of unnaturalness. The results showed that both humans and models exhibited significantly positive unnaturalness scores, revealing their ability to detect lighting inconsistencies. However, humans were more sensitivity to these modifications, and correlations between human and model scores were modest (r=0.22-0.26), suggesting limited alignment at the individual image level. Overall, these findings demonstrate that while generative models can detect certain unnatural modifications similarly to humans, their sensitivity and alignment with human perception remain limited.

This work has been supported by JSPS Kakenhi 24H00721

43.338 HUMANS VERSUS MACHINES: DISTINGUISHING KOREAN, CHINESE, AND JAPANESE FACES VIA INTERNAL AND EXTERNAL FEATURES

Cansu Malak[1], Christian Wallraven[1]; [1]Korea University

Deep learning algorithms have shown super-human performance for face identification for a number of years now, raising the question of to

what degree these algorithms process faces similarly to humans. In order to dive deeper into comparing human versus machine face processing, here we present results from a fine-grained ethnicity categorization task, in which the goal was to distinguish three different categories of East Asian faces (Chinese, Japanese, and Korean) using internal and external facial features. For the human experiment, we showed participants 600 grayscale images of male soccer players either as cropped faces (Experiment 1, internal features only, N=53) or as full-face images (Experiment 2, including external features like hair and face outline, N=52) in a three-alternative-forced-choice task. For cropped faces, the performance was 40.08% on average, which improved to 52.38% for full faces (p<.001), showing how difficult the task was for humans. To compare with deep learning algorithms, we used the DeepFace library to extract embeddings from 10 state-of-the-art face recognition models, using a support vector machine classifier to predict ethnicity via 10-fold cross-validation. The average validation set performance for the models was significantly higher than human performance at 74.88% (p<.001) - with no differences across algorithms or image types. Interestingly, we observed that Japanese faces were easier to categorize for both humans and machines. Conversely, an item-based analysis showed only weak concurrence for accuracy (all r2<.08; p<.001) between humans and machines. Overall, our results show that internal features suffice for ethnicity categorization for deep learning algorithms, whereas humans require external information (from hairstyle and/or face outline) for this task. Whereas there are some shared performance patterns, the deep learning algorithms seem to process faces differently from humans in this task.

This study was supported by the National Research Foundation of Korea (BK21 FOUR, NRF-2022R1A2C2092118, NRF-2022R1H1A2092007) and by IITP grants funded by the Korea government (No. RS-2019-II190079, Dept. of AI, Korea University; No. RS-2021-II212068, AI Innovation Hub).

43.339 IS FACIAL ATTRACTIVENESS JUDGMENT HOLISTIC?

Cheng Hsuan Chen[1] (r09122005@ntu.edu.tw), Chien Chung Chen[1]; [1]National Taiwan University

It is known that face recognition is holistic, or that faces are processed based on the spatial relationships between facial features rather than individual components. However, it remains unclear whether facial attractiveness judgement is based on separate facial features or the configuration between them. We investigated this issue with a composite face paradigm. We created a stimulus set by morphing faces between original and averaged faces and then extrapolated to create caricatures. We then split the faces into upper and lower halves and recombined them to create composite faces with varying degrees of attractiveness in upper and lower halves of a face. Participants were to judge the faces by three criteria: solely on the upper half, the lower half, and the whole face. When the judgement was based on either upper or lower halves, ratings were only influenced by the morphing level of that half. When the judgement was based on the whole face, the attractiveness rating depended on a combination of independent upper and lower face evaluations. Using a model derived from the multidimensional signal detection theory, we revealed that there was only minor interaction between the two halves of the faces. Both lower and upper faces contributed to attractiveness judgement with a greater emphasis on the upper face. The findings demonstrate that facial attractiveness judgments are primarily analytic rather than holistic, contrasting with the holistic face recognition processes. These results help resolve conflicts in existing literature by showing that attractiveness judgments involve predominantly independent processing of facial features with limited interaction between them.

NSTC (Taiwan) 112-2423-H-002-002

43.340 RETHINKING CONFIGURATIONAL AND FEATURE-BASED PROCESSING IN VISUAL RECOGNITION

Yuxuan Zeng[1] (zeng.774@osu.edu), Ren E Hentz[1], David E Osher[1]; [1]The Ohio State University

Second-order configuration has long been considered a hallmark of face perception, unique among high-level visual categories. Behavioral phenomena, such as the difficulty in recognizing inverted faces, are frequently attributed to the unique configurational information inherent to faces. However, it remains unclear whether such information is exclusive to faces or if it can be generalized to other visual stimuli. To address this question, we developed novel, abstract stimuli that contained only second-order configural information, with no resemblance to faces, which contain identical features across the stimulus set (i.e. they can only be distinguished though configuration and not by the presence or absence of features). In parallel, we designed control stimuli that included only feature information without distinct configurations (i.e. they can only be distinguished by their features). Twenty participants were trained to categorize these stimuli and were tested on tasks traditionally associated with configurational processing, including inversion, misalignment, part-whole, and composite tasks. Reaction times during categorization were compared across tasks and conditions relative to the final training stage. Results showed significant effects in the configuration condition across all tasks (p = 3.79x10-3 inversion, 1.35x10-4 misalignment, 2.15x10-3 part-whole, 4.04x10-5 composite; Bonferroni corrected). Importantly, the feature condition revealed no significant effects for any of these manipulations. These findings align with prior research on face and object perception but challenge the prevailing notion that second-order configuration is exclusive to faces. By demonstrating that second-order configuration can exist independently of facial features, our results invite a re-evaluation of the interplay between configurational and feature-based information in broader object recognition processes, including face perception.

43.341 GRADUAL INVERSION COSTS AS A FUNCTION OF ROTATION FOR FACE RECOGNITION IN A FLICKER PARADIGM

Lauren Williams[1], Jonathan Flombaum[1], Justin Halberda[1]; [1]Johns Hopkins University

Faces are harder to recognize when they appear upside down—an inversion cost. This classic result is explained by a shift from configural & component processing to only component processing. However, this shift in encoding is not well characterized in the face literature through intermediate rotations. Is a face also harder to recognize when it is nearly upside down, e.g. rotated by 135°? How hard is it, comparatively, when the rotation is 90° or 45°? Face recognition as a function of orientation—from upright to upside down—should reveal how and why face processing mechanisms absorb costs for inversion. We therefore investigated face recognition as a function of orientation using flicker change detection. In Experiment 1, participants searched for a changing face within an array of five (drawn from the Chicago Face Database) at orientations between 0° and 180° in 45° steps. When all the faces were upside down reaction times were significantly longer than when all were upright—an inversion cost. Reaction times increased through the intermediate rotations following a sigmoidal function, with an inflection at 90°. These results are consistent with a monotonic, non-linear loss of configural signal as faces rotate from upright to inverted rather than an immediate and complete loss at any non-upright orientation. In Experiment 2, we characterized the inversion cost for other objects (scenes, animals, and vehicles). If faces are special, they should be an outlier among our tested objects. However, all object classes contributed to the positive relationship between time to detect an item change and the inversion effect – i.e., the magnitude of the inversion effect scales with the overall difficulty of identification. That is, all categories show an inversion effect that is proportional to the difficulty of processing that item type, and faces are not special in this regard.

43.342 CONTRAST-NEGATION INCREASES FACE PAREIDOLIA IN NATURAL IMAGES

Benjamin Balas[1], Emily Westrick[1], Lily Roshau[1], Molly Setchfield[1]; [1]North Dakota State University

Face pareidolia, the phenomenon of seeing face-like patterns in non-face images, has a dual nature: Pareidolic patterns are experienced as face-like, even while observers can recognize the true nature of the stimulus (Stuart et al., 2024). Though pareidolic faces seem to result largely from the canonical arrangement of eye spots and a mouth, we hypothesized that competition between veridical and face-like interpretations of pareidolic patterns may constrain face pareidolia in natural scenes and textures. Specifically, we predicted that contrast negation, which disrupts multiple aspects of mid- to high-level recognition, may increase rates of face pareidolia in complex natural textures by weakening the veridical, non-face stimulus interpretation. We presented adult participants (N=27) with a series of natural images depicting textures like grass, leaves, shells, and rocks. Each of our original 8 grayscale images was first used to create a mirror-symmetric stimulus (per Paras & Webster, 2013) by dividing pictures at the vertical midline and replacing the right half of the image with a flipped copy of the left half. These symmetric images were then rendered in positive and negative contrast for a total of 16 images, which were shown in a pseudorandomized order to each participant on an iPad. We asked participants to circle any patterns in each image that looked face-like, with no constraints on response time or pattern size, position, and orientation. We analyzed the resulting count data (number of circled patterns per image) with a generalized linear model including image contrast (positive or negative) as a fixed effect and participant ID and item number as random effects. This revealed a significant effect of image contrast, with contrast-negated images yielding more pareidolic face detections than positive images. We conclude that disrupting veridical object and texture recognition enhances pareidolia by compromising half of the dual nature of a pareidolic pattern.

This research supported by NSF-2338600 awarded to BB.

43.343 INVESTIGATING THE ROLE OF KEY FACIAL FEATURES FOR FACE DETECTION

Laurianne Côté[1] (lauriannecote31@gmail.com), Jérémy Lamontagne[1], Alexis Bellerose[1], Caroline Blais[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais

An important body of literature is dedicated to understanding how humans compute the complex information required for face recognition. Specifically, many have stated that the eyes and mouth regions play a fundamental role in this process, supported by evidence from individual differences and prosopagnosic patients (Caldara et al., 2005; Royer et al., 2018; Tardif et al., 2019). While psychophysical studies have established the importance of these regions in identification, few studies have explored the preceding step: face detection. That said, a study by Xu and Biederman (2014) suggests a link between these two processes, showing that a prosopagnosic patient experienced notable difficulties in a face detection task, suggesting that the same facial information might be used for both processes. The present study therefore aims to identify the key regions involved in face detection, a crucial step before delving deeper into this aspect in patients with acquired prosopagnosia. In this study, twenty participants completed 3,000 trials divided in two face detection tasks (1 - Does the presented stimulus contain a face? 2 - Which of the two stimuli contains a face?), with the non face stimuli being 100% wavelet decomposed faces (Koenig-Robert & VanRullen, 2013). In both tasks, stimuli were overlaid with Bubbles across five spatial frequency (SF) bands (Gosselin & Schyns, 2001). The resulting classification images reveal that face detection relies on all facial features (eyes, nose, and mouth, $p = .001$) across all spatial frequency bands, with the presence of the eyes being associated with higher Z-scores. These results suggest that face detection and face recognition processes are closely linked, as both rely on the same facial regions to accomplish the task. This opens new perspectives for understanding and diagnosing deficits associated with prosopagnosia, but also for better understanding the underlying processes of face recognition.

Natural Sciences and Engineering Research Council of Canada (Discovery) and Master's Research Scholarship (B1X) from the Fonds de recherche du Québec - Nature et technologies (FRQNT)

43.344 INTEGRATIVE PROCESSING OF INVARIANT AND CHANGEABLE FACIAL INFORMATION: EFFECTS OF RACE, GENDER, AND EXPRESSION ON FACE IDENTIFICATION

Olivia S. Cheung[1,2] (olivia.cheung@nyu.edu), Rawan Abdelaal[1], Treedom Beiyin Zhang[1]; [1]New York University Abu Dhabi, [2]Center for Brain and Health, NYUAD Research Institute

At a glance, multiple aspects of information are processed from a face. Is invariant information (e.g., identity, race, gender) integrated more strongly than changeable information (e.g., expression), despite visual salience of the information? Here, integration of invariant and changeable facial information is revealed by how race, gender, or expression may influence identification. Participants searched for a target face (e.g., a happy Black female) among distractor faces either shared identical aspects of information (baseline), or differed in one, two, or all three aspects. Integration was revealed by faster search when there was a difference involving a particular aspect in the distractors, compared with the baseline (see Zhao & Hayward, 2013). Prior to the main experiment, we first minimized potential influences that were unrelated to face processing: 1) overall luminance was adjusted among all faces using SHINE (Wilenbockel et al., 2010); 2) relative visual salience of the three aspects of facial information was assessed using inverted faces. When participants (N=20) searched for inverted target faces, we found the strongest effect when the target and distractor faces differed in race, then expression, then gender, information. Moreover, differences in two aspects of facial information on inverted faces also reflected the combination of visual differences. However, unlike for inverted faces, search performance for upright faces did not appear to be merely affected by visual salience differences. Instead, while participants (N=20) were the fastest when target and distractor faces differed in race, search speed was comparable when the target and distractor faces differed in either gender or expression. Similar results were also observed when target and distractor faces differed in two aspects of facial information. These results suggest that for upright faces, invariant information such as race and gender are more integrated with identity processing than changeable information, despite visual salience of expression over gender differences.

The research was supported by a faculty grant (AD174) and a Tamkeen NYU Abu Dhabi Research Institute grant (CG012) to O.S.C., and a NYU Abu Dhabi capstone grant to R.A.

43.345 DECOMPOSING THE PERCEPTUAL AND CONCEPTUAL BIAS OF 3D FACE MEMORY IN INDIVIDUAL PARTICIPANTS

Zigui Zhu[1], Xiaoxue Gao[2], Philippe Schyns[3], jiayu Zhan[1]; [1]Peking University, [2]East China Normal University, [3]University of Glasgow

Face identification relies on both observers' perceptual abilities and their conceptual understanding of faces. However, it's unclear how these components contribute to the identity information observers memorize to recognize familiar faces. To address this question, we applied a data-driven approach to model, at individual level, the facial features that 48 participants (29 Females and 19 males) used to identify their own faces, and for comparison the facial features they used in face discrimination and facial attractiveness rating. Specifically, we used a generative model of 3D faces to synthesized a broad range of female and male faces that varied naturally in 3D face shapes and complexions. Then, we tested these generated faces in three within-subject tasks: 1) assessing same-sex faces for similarity to their own face in memory, 2) evaluating the physical similarity of paired faces based on perception, and 3) rating opposite-sex faces from "very unattractive" to "very attractive" according to personal preference. We reverse correlated the specific features that modulated memory-based self-face similarity, perceptual-based visual similarity, and concept-based attractiveness. Through pairwise and triple-wise comparisons, we analyzed the specific fit of 3D shape and complexion features across three feature sets, and identified the memory bias linked to individuals' perceptual sensitivity and conceptual preferences. These biases were further characterized in relation to participants' face memory abilities and their attitudes toward self-attractiveness. Our findings reveal how memory biases are shaped by perceptual and conceptual processes, offering a novel framework for understanding the idiosyncratic biases in high-dimensional facial feature processing.

43.346 INVESTIGATING THE INDEPENDENCE OF FACE SHAPE AND MOTION THROUGH REVERSE CORRELATION

Raphael Tordjman[1], Emily Martin[2], Fabian Soto[3]; [1]Florida International University

Face perception involves processing both static shape and dynamic motion cues, yet the degree of independence between these dimensions remains unresolved. Using reverse correlation, we investigated whether the information used for recognition of face shape and natural motion are separable across varying contextual dimensions. Dynamic, naturalistic stimuli were generated using 3d face modeling software (MakeHuman) extended with a novel model of face motion (implemented in FaReT), with Gaussian noise applied to either shape or motion parameters. Participants classified stimuli based on shape or motion in multiple contexts. Participants in shape groups classified face shape in the context of different motions (natural vs. unnatural), and participants in motion groups classified face motion in the context of different shapes (i.e., identities). We then tested whether any information used during each task was altered by the changes in context. Our findings revealed that the information used by participants to classify a motion as natural strongly depended on the face shape accompanying that information. In other words, participants changed their expectations of how a face should move depending on the face's shape. This challenges the assumption of strict independence between these perceptual dimensions. However, information used by participants to classify faces based on their shape did not show the same strong dependence on the type of motion shown by the face. These results underscore the adaptive nature of face perception, and our novel methodology opens new avenues for the study of the context-specific nature of face perception.

MONDAY MORNING POSTERS IN PAVILION

Plasticity and Learning: Perceptual learning

Monday, May 19, 8:30 am – 12:30 pm, Pavilion

43.401 DIFFERENTIAL NETWORK METRICS AS PREDICTORS OF SPECIFICITY AND TRANSFER IN PERCEPTUAL LEARNING

Lin Zhong[1], Gong-Liang Zhang[2], Cong Yu[1,3]; [1]Zhejiang University, [2]Soochow University, [3]Peking University

Visual perceptual learning is characterized by specificity to trained conditions. Understanding the mechanisms behind this specificity remains a fundamental topic in perceptual learning research. In a previous ERP study (Zhang-et-al., 2013), half the participants (n=14) in Vernier learning exhibited location specificity (TI<0.5), while the other half demonstrated learning transfer (TI>0.5) across the visual hemifield. Here, we examined small-world network metrics to compare the changes in the topological properties of the whole-brain network

between the specificity and transfer groups. We established functional connectivity between each pair of EEG channels by calculating the phase-locking value using wavelet transform in the beta band. We then applied a threshold to obtain undirected binary small-world networks for each participant at trained and transfer locations before and after training. From these networks, we calculated the clustering coefficient, indicating local interconnectedness, and the shortest path length, reflecting global information integration. Before training, both groups exhibited similar small-world properties, characterized by high clustering coefficients and short path lengths. After training, while both groups displayed similar threshold reductions, the specificity group showed higher clustering coefficient (p=.003) and longer path length (p<.001), whereas the network metrics of the transfer group remained unchanged (ps>.111). Additionally, when the transfer effects were tested at the untrained location, the specificity group demonstrated trends (ps<.007) in network metrics similar to those at the trained location, while the transfer group displayed higher clustering coefficient (p=.035) with unchanged path length. We conclude that whether learning is location-specific is pre-determined by training-induced changes in network properties. While the transfer group maintained both local and global efficiencies, the specificity group exhibited over-optimized local efficiency at the expense of global efficiency, which likely hinders learning transfer. These results provide crucial insights into the mechanisms of specificity and transfer in perceptual learning from the perspective of brain network dynamics.

STI2030-Major Projects grant (2022ZD0204600)

43.402 POSITION SPECIFICITY OF LEARNING USING COMPLEX VISUAL STIMULI

Jamie G.E Cochrane[1] (cochrj1@mcmaster.ca), Natasha Lacku[1], Allison B. Sekuler[1,2,3], Patrick J. Bennett[1]; [1]Department of Psychology, Neuroscience & Behaviour, McMaster University, [2]Rotman Research Institute, Baycrest, [3]Department of Psychology, University of Toronto

Visual perceptual learning (PL) is characterized by long-lasting, stimulus-specific improvements in simple visual tasks. One explanation for stimulus specificity is that PL causes changes in early cortical areas; an alternative explanation is that PL is due to a higher-level process, reflecting what is learned rather than where it occurs in the visual pathway. One method to evaluate this hypothesis is to test if PL occurs in a task using complex stimuli encoded later in the visual pathway. We measured response accuracy with a 1-of-5 identification task using complex stimulus types (textures and faces) encoded by mechanisms in the inferotemporal cortex. In the training phase, participants saw one of the two stimulus types above or below a central fixation point. In the test phase, Group 1 identified the same stimuli in the same position; Group 2 identified the same stimuli in a new position; Group 3 identified new stimuli of the same type at the same position; and Group 4 identified new stimuli of the same type in a new position. For both stimulus types, we found evidence of both generalized and stimulus-specific learning: changing the stimuli and/or stimulus position significantly reduced accuracy but not to the level shown at the start of training. In addition, for textures (but not faces), accuracy was significantly lower in Group 4 than in Groups 2 and 3. Therefore, stimulus- and position-specific PL occurs in an identification task using complex stimuli.

NSERC

43.403 ROLE OF SLEEP IN GENERALIZATION AND SPECIFICITY OF VISUAL PERCEPTUAL LEARNING

Theodore LaBonte-Clark[1], Teruaki Kido[1], Nitzan Censor[2], Takeo Watanabe[1], Yuka Sasaki[1]; [1]Department of Cognitive and Psychological Sciences, Brown University, [2]School of Psychological Sciences, University of Tel Aviv

Sleep is instrumental for various types of procedural learning, including improvement on basic visual tasks, referred to as visual perceptual learning (VPL). Sleep after training on VPL tasks leads to further improvement, referred to as off-line performance gains. Given that VPL is often specific to the trained location, previous research primarily focused on off-line performance gains at the trained location. However, it remains unclear whether off-line performance gains occur at untrained locations — in other words, whether sleep contributes to location transfer in VPL. Here, we show the tendency that off-line performance gains occur at both trained and untrained locations. In our experiment, both groups underwent training and two posttests, one immediately after training and one 12 hours later. The wake group trained in the morning and retested that evening, while the sleep group trained in the evening and retested the following morning. Only participants in the sleep group slept between posttests. Both groups followed otherwise identical procedures; participants were trained to detect an orientation presented in one upper quadrant of the visual field and were tested at both the trained and untrained upper quadrants. The sleep group exhibited off-line performance gains at both the trained and untrained locations, indicating location transfer. Conversely, the wake group does not show a clear pattern of location transfer at this stage of data collection, supporting the hypothesis that location transfer occurs only in the sleep group. Our results suggest that sleep may facilitate location transfer of VPL in an orientation detection task, generalizing learning to untrained locations. Interestingly, previous research reported no off-line transfer of learning to untrained orientations in a similar detection task (Tamaki et al., 2020, Nature Neuroscience). Comparing our findings with prior work highlights the possibility that distinct mechanisms underlie specificity and generalization of VPL for features and locations.

NIH R01EY019466, R01EY027841, R01EY031705, NSF-BSF BCS2241417

43.404 HEBBIAN PLASTICITY IN A RECURRENT NETWORK MODEL EXPLAINS ENHANCED WAKEFUL CONSOLIDATION VIA REPETITIVE SENSORY STIMULATION

Xin-Yue Yang[1,2,3,4], Fang Fang[1,2,3,4]; [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]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China, [4]IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

Continuous training on a visual task improves corresponding visual abilities, but the mechanisms of post-training wakeful consolidation and how to intervene during this stage remain unclear. Previously, we demonstrated that repetitive grating stimulation at 10 Hz after training on an orientation discrimination task enhanced wakeful consolidation. This enhancement was absent without prior training or when repetitive uniform-disk stimulation was used. Additionally, augmented visual evoked potentials and increased glutamate-GABA concentration ratio in the visual cortex accompanied this enhancement, suggesting the involvement of long-term potentiation (LTP). However, the link between these findings and underlying neural circuit mechanisms remains unexplored. To address this, we developed a Habbian recurrent network model to replicate key findings through simulations. In the model, spiny stellate (SS) cells integrate inputs from thalamocortical neurons and project to superficial pyramidal (SP) neurons and parvalbumin-positive interneurons (SI). Recurrence is introduced as SPs and SIs mutually project onto both themselves and each other. All synapses are Hebbian-plastic, with LTP driven by repeated presynaptic contributions to postsynaptic firing. Orientation preferences were assigned using Gaussian initial weights, with reduced sigma values for the connections between neurons preferring the orientation used in visual training, a process known to sharpen tuning curves. Iterative orientational inputs simulated the repetitive sensory stimulation, with network properties statistically compared pre- and post-iterations. Results showed that both feedforward synaptic weights and SP tuning curves demonstrated enhanced orientation discrimination following the LTP simulation. The enhancement diminished without prior tuning-curve sharpening and was further weakened with uniform noise inputs, which simulate uniform-disk stimulation. Additionally, elevated glutamatergic presynaptic activity relative to GABAergic activity accompanied the enhancement. The model successfully reproduced behavioral and neurometabolic findings, supporting the hypothesis that LTP underlies enhanced wakeful consolidation induced by repetitive sensory stimulation, and offering insights that could guide future research into wakeful consolidation mechanisms.

43.405 PERFORMANCE DECREASES FOR UNTRAINED ORIENTATION OBSERVED IN DOMINANT COMPUTATIONAL MODELS BUT NOT HUMANS ARE MITIGATED BY DIVISIVE NORMALIZATION IN ENCODING PROCESSES OF VISUAL PERCEPTUAL LEARNING

Yu-Ang Cheng[1], Yuka Sasaki[1], Thomas Serre[1], Takeo Watanabe[1]; [1]Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, USA

Visual perceptual learning (VPL) refers to long-term performance improvements following visual experience. It is controversial whether VPL arises from plasticity at the level of neural encoding processes or downstream readout processes (Watanabe & Sasaki, 2015, Ann. Rev. Psych.; Dosher & Lu, Perceptual Learning, MIT Press, 2020). To address this issue, we compared how training on an orientation detection task alters performance in humans and in a well-established neural network model of the early visual cortex, which includes broad excitation, untuned inhibition, and readout components. We focused on performance changes in untrained orientations, which have been discussed only relative to the trained orientation. In the psychophysical experiment with humans, as predicted, the greatest performance enhancement was observed at the trained orientation. Enhancements tapered off as orientations deviated further from the trained one, until around 90 degrees, where no performance increase was observed. Performance at orientations approximately 90 degrees from the trained orientation remained unchanged. In the neural network model simulation, plasticity in the readout components always led to unexpected performance decreases at untrained orientations. However, the introduction of plasticity in untuned inhibition during neural encoding processes, leading to a divisive normalization effect, mitigated the performance decreases at untrained orientations. These findings suggest that divisive normalization plasticity may resolve the discrepancy between the results from human psychophysics and the initial computational model. Our results further suggest the involvement of neural encoding processes in VPL.

NIH R01EY019466, R01EY027841, R01EY031705, NSF-BSF BCS2241417, ONR (N00014-24-1-2026), ONR (N00014-19-1-2029), DARPA (D19AC00015), NIH/NINDS (R21 NS 112743), the ANR-3IA Artificial and Natural Intelligence Toulouse Institute (ANR-19-PI3A-0004)

43.406 CAN MINDFULNESS MEDITATION IMPROVE LEARNING ABILITIES? A CROSS-SECTIONAL STUDY

Adrien Chopin[1] (adrien.chopin@gmail.com), Lia Antico[2], Marie Angelillo[3], Maeva Briguet[3], Orane Schaefer[3], Daphne Bavelier[3]; [1]Smith-Kettlewell Eye Research Institute, San Francisco, California, USA, [2]Mindfulness Center, School of Public Health, Brown University, Providence, Rhode Island, USA, [3]Faculty of Psychology and Educational Science, University of Geneva, Geneva, Switzerland

Introduction: Several meta-analyses have shown that mindfulness meditation enhances various cognitive aspects, but the mechanisms underlying these effects and their specific impact on learning abilities remain unclear. This study aims to explore the relationship between mindfulness meditation and learning abilities, as learning may be a key mechanism through which cognition enhancements occur in a generalized manner. Methods: Using a cross-sectional design, we compared regular meditators and non-meditators in two learning tasks, one tapping audio-visual working memory and the other visual perception. Pre-existing performance differences were assessed through initial performance scores, while learning speed was measured by the rate of performance change across task repetitions. Participants were matched for age, sex at birth and videogaming proficiency. Additionally, we examine the effect of potential covariates such as sport, music practice and expectations, comparing several generalized linear mixed-effect models. Results: Regular meditators outperformed non-meditators in initial working memory performance, consistent with prior research; yet, they did not show faster learning. In the visual orientation discrimination task, initial performance was seen again to differ but here meditation interacted with music practice. Meditators with lower levels of music practice outperformed non-meditators, while those with higher levels of music practice started with lower performance. Interestingly, although meditators did not learn faster overall, exploratory analysis revealed greater learning gains in meditators, as measured by the difference between final and initial performance. Finally, neither performance nor learning variables were influenced by expectations. Conclusion: Mindfulness meditation facilitated initial task performance, albeit conditioned on music expertise for our perceptual task. In contrast, mindfulness meditation did not change learning speed in either task; yet, it may result in greater perceptual learning gain, although replication is needed. Further research is also required to explore the mechanisms underlying the effects of mindfulness meditation on perception, in particular its interactions with music practice.

43.407 TRAINING ON A VISUAL CROWDING TASK IN ECCENTRIC VISUAL LOCATIONS WITH VARYING LEVELS OF SUSTAINED ATTENTION PERFORMANCE

Elena von Perponcher[1], Renata Hechenrieder[1], Mark Greenlee[1], Tina Plank[1]; [1]Institute for Experimental Psychology, University of Regensburg

Individuals with central vision loss rely on peripheral vision, leading to the development of a "preferred retinal locus" (PRL) through oculomotor and perceptual learning. The mechanisms guiding PRL location selection remain unclear. One proposed factor influencing PRL location selection – and potentially identifying optimal sites for PRL training prior to the onset of a complete central scotoma – is the individual's ability to sustain attention at that location. To investigate the link between sustained attention and perceptual learning, we assessed sustained attention in eight peripheral locations (8° eccentricity) in 14 normally sighted participants. Participants fixated centrally and directed their attention for 2.5 – 4 seconds to a cued eccentric location before identifying the orientation of a Landolt-C gap (up, down, left, right) at that location, amidst full-ring distractors in the other seven locations, followed by 100 ms masks across all locations. Participants then underwent four training sessions in a Landolt-C gap detection crowding task at their highest (HPL) and lowest (LPL) sustained attention performance locations. 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. Following training, sustained attention was re-evaluated across all locations. Preliminary results indicate: (1) significant improvement in sustained attention at the LPL ($p < .05$), but not at the HPL (despite room for improvement), (2) stronger initial crowding effects and radial-tangential anisotropy at the LPL compared to the HPL, and (3) a trend towards greater perceptual training effects at the LPL. The findings suggest that perceptual learning training may enhance sustained attention performance in locations recommended for PRL development, particularly when these locations initially exhibit suboptimal sustained attention. Interactions between HPL and LPL placement within the visual field and crowding performance need to be determined further with larger cohorts.

This research was supported by funding programs of the German Academic Scholarship Foundation (Studienstiftung des deutschen Volkes), the Marianne Plehn Program of the Elite Network Bavaria, the Deutsche Forschungsgemeinschaft (DFG: PL641-1/1), and the ZSER e.V. Regensburg.

43.408 PLFEST; FEASIBILITY DATA FOR AN OPEN SCIENCE TOOL FOR RELIABLE PERCEPTUAL LEARNING RESEARCH

Marcello Maniglia[1] (mmanig@ucr.edu), Jaap Munneke[2], Diya Anand[2], Samyukta Jayakumar[1], C. Shawn Green[3], Aaron Seitz[2]; [1]University of California, Riverside, USA, [2]Northeastern University, Boston, USA, [3]University of Wisconsin-Madison, USA

Perceptual learning (PL), the practice-induced improvement in perceptual tasks, holds the promise of expanding our understanding of brain plasticity mechanisms and developing effective interventions for visual pathologies such as myopia and amblyopia. However, despite its promise, research in this field is often hindered by issues of reproducibility and accessibility. Concerning reproducibility, observed inconsistencies might be due to methodological or environmental differences. Regarding accessibility, despite offering a potentially cost-effective intervention for visual pathologies, PL research is often confined to university laboratories or specialized clinics. To overcome these challenges and enhance the potential of perceptual learning (PL) as a tool for studying learning-related neural processes and broadening its application for interventions, we recently launched PLfest — a UNITY-powered application that enables data collection across different sites and platforms, including computers and tablets. PLfest currently supports several PL training paradigms, as well as visual, attentional and cognitive assessments. Here, we present preliminary data from a large number of healthy individuals as a proof of concept for using PLfest as a framework to test a variety of training paradigms within a controlled and reproducible environment across multiple testing locations. Specifically, we report data from participants trained on variations of a contrast detection task, a classic PL paradigm, conducted across four experimental sites. Training types varied in the parameters of the adaptive procedure (long and short staircases and sessions) and the stimulus configuration (contrast detection with flankers, contrast detection in noise, stimulus variety). Before and after training, participants were tested on several assessment tasks, assessing perceptual, attentional and cognitive performance. The results demonstrate that PLfest is a reliable tool for cross-site data collection, highlighting its potential as a platform for large-scale, multi-site, and remote studies.

Attention: Temporal

Monday, May 19, 8:30 am – 12:30 pm, Pavilion

43.409 AN ELECTROPHYSIOLOGICAL INVESTIGATION OF RHYTHMIC ENVIRONMENTAL SAMPLING IN A CUED TEMPORAL ATTENTION PARADIGM

Travis Powell[1], Megan Wong[1], Jake Eliot[1], Ding Mingzhou[2], George R. Mangun[1]; [1]University of California, Davis, [2]Universtity of Florida

Behavioral fluctuations oscillating in the theta-band (3—8 Hz) have been observed in visuospatial attention tasks when sampled over a range of cue-target intervals. This periodic nature of performance during the deployment of attention provides evidence that attentional capture is a discrete process. Previous work has found that saccadic activity and visual sensitivity are in antiphase to each other, governed by a theta rhythm common across a control network. This mechanism thereby resolves an inherent conflict between sensory and motor processes, facilitating shifts of attention from one area of space to another. We extended this model of attentional rhythmicity beyond the visuospatial domain with a modified Posner paradigm, cueing participants to two temporally distinct sampling windows of cue-target intervals while acquiring scalp electroencephalography (EEG) and eye-tracker data. Mean reaction time (RT) was observed fluctuating in the 3—8 Hz range. Attention was found to modify event related potentials (ERPs) time-locked to the target stimulus. We also examined post-cue spectral activity and eye movements. Replication of earlier rhythmic behavioral findings, extended to temporal attention — which recruits a different control network, and different hemispheric bias — suggests a more general oscillatory mechanism than previously considered.

43.410 DO PEOPLE SEE SPATIO-TEMPORALLY PREDICTABLE VISUAL INPUTS SOONER, AND DOES THIS ALIGN WITH NEURAL PRE-PLAY FINDINGS?

Blake Saurels[1] (b.saurels@uq.edu.au), Derek Arnold; [1]The University of Queensland, [2]The University of Queensland

The neural latencies of your sensory brain are not too severe – just tens of milliseconds. Nonetheless, some have suggested that your brain tries to compensate for these brief lags to help you interact with fast moving objects. They suggest this could work via predictive processes at both early (retinal motion extrapolation) and later (neural pre-play) stages in the visual hierarchy. We are interested in how predictable motion impacts perception, and relating these changes in perception to neural predictive processes. We have approached this in two ways. The first uses the 'Twinkle-Goes' illusion (Nakayama & Holcombe, 2021, Journal of Vision), an apparent perceptual extrapolation of moving objects that suddenly disappear against dynamically updating white noise. We have: 1) refined the temporal conditions that produce the illusion, 2) showed that it is not the result of a decision level bias, and 3) showed that it impacts motor planning. Our second approach uses a temporal order judgement task to determine how predictable motion changes when objects seem to appear. We adapted the design of an EEG experiment that produced neural pre-play activity via apparent motion (Blom et al., 2020, PNAS). By testing different locations along the motion path, we can partition out the impact of exogenous attention capture (related to 'prior entry' effects). We have found that people report seeing objects sooner when they are presented along the expected motion path. But, only for a brief window (50 ms) after the motion ends, and we have found the opposite for objects presented after this window. Our latest work looks at how these changes in perception are predicted by differences in neural pre-play activity within and across individuals, as measured by EEG, and what factors might moderate these relationships.

43.411 EMOTIONAL GAZE INCREASES TARGET TEMPORAL PROCESSING

Florence Mayrand[1], Sarah McCrackin[1], Jelena Ristic[1]; [1]McGill University

Human attention is spontaneously oriented in the direction of eye gaze, with such gaze following behavior enhanced when a face shows an emotional expression. Here we investigated how emotional eye gaze affects the temporal precision of target perception. Participants viewed a face that either maintained a neutral expression or displayed an emotional reaction (fearful or happy) upon averting its gaze to the left or right. Two peripheral targets were presented, one to the left and another one to the right of the face. These targets were temporally offset by 0ms, 50ms, 100ms, or 150ms. Participants performed a temporal order judgement, reporting on which target appeared first (left or right). Data showed greater accuracy when the first target appeared at the gazed-at location, demonstrating that eye gaze facilitates temporal perception. Greater sensitivity for target timing at gazed-at locations was also supported by steeper response curves (reflecting proportion of right vs. left responses) and was further modulated by the emotional expression of the face, such that increased temporal target perception was present when faces displayed emotional expressions (fearful and happy) but not when they remained neutral. Together, these findings show that eye-gaze perception enhances the temporal processing of targets at gazed-at locations, and particularly so when the face display an emotional expression. This potentially reflects an adaptive mechanism that facilitates rapid responses to biologically relevant emotional face stimuli.

43.412 IS TEMPORAL INTEGRATION A UNITARY PROCESS?

Abi Wyllie[1], Alon Zivony[1]; [1]University of Sheffield

Temporal integration occurs when individual features presented very close together in time are combined in our perception as a single item (e.g. seeing '/' followed immediately by '\', but perceiving it as 'X'). This process allows us to combine large quantities of disparate feature information (e.g., colour, shape) into coherent objects, enabling us to detect and identify important events and navigate our environment. Research into temporal integration often uses one of two tasks: the Missing Dot Tasks (MDT) and the Rapid Serial Visual Presentation (RSVP) task. In MDT, two partially filled grids are presented and participants report the location of a single unfilled cell, which can only be achieved through successful integration of the two grids. In RSVP tasks, a series of visual stimuli are presented in rapid succession, and participants report the target item based on a selection cue. In this task, temporal integration between the target and a distractor can result in binding errors, and such errors have been associated with the deployment of spatiotemporal attention. Whilst both tasks are thought to represent temporal integration, it is currently unclear whether performance in these two tasks represents a single integration mechanism. We investigated this in 2 experiments. In the first experiment, participants completed the two tasks separately and correlations between individual's performances were measured. In the second experiment, participants completed both tasks concurrently. During the RSVP, the two MDT grids are presented alongside the target and post-target items, and participants report both the central target and the missing dot. No significant correlations were observed between the two measures of integration. These results therefore suggest that the process known as temporal integration is not caused by a single attentional mechanism. Rather, we suggest that multiple processes may determine integration.

43.413 PREDICTING PATHOLOGIST ATTENTION DURING CANCER-IMAGE READINGS

Gregory Zelinsky[1] (gregory.zelinsky@stonybrook.edu), Souradeep Chakraborty[1], Joel Saltz[1], Dimitris Samaras[1]; [1]Stony Brook University

We obtained the attention behavior of pathologists conducting cancer readings of whole-slide-images (WSIs) of prostate. WSIs are giga-pixel in size and comparable in scale to geo-spatial imagery, precluding their exhaustive inspection. The pathology task is more comparable to the expert navigation through a 3D space in targeted pursuit of information rather than standard visual search. Accordingly, we measured the x, y, and m (magnification) movement of a pathologist's viewport as they navigated through a WSI using a digital microscope, which we named their attention trajectory. We did this for

43 pathologists providing Gleason gradings for 123 prostate WSIs, from which we obtained 1,016 attention trajectories. To pre-process the data for model training, we developed and used a fixation-extraction algorithm to convert the densely sampled attention trajectories into sparser scanpaths of attention "fixations". We trained a single two-stage model on these attention fixations to predict spatio-temporal attention scanpaths from pathologists in disjoint test data. In the first stage of processing, we train a vanilla vision transformer to predict the attention heatmaps computed for multiple magnification levels and we show that this produces new state-of-the-art (SOTA) performance in attention heatmap prediction for WSI readings. In the second stage, we propose a new transformer-based model that takes the multi-magnification attention heatmaps predicted from first-stage processing and uses them as feature representations to predict the attention scanpath of the pathologist. It does this by sequentially predicting each fixation of the scanpath, starting from the WSI center, in an autoregressive manner until the entire scanpath is obtained. We show that our model outperforms baseline models, thereby establishing it as SOTA performance in the new task of pathologist attention-scanpath prediction. Tools developed from this model could assist pathology trainees in learning to allocate their attention during WSI reading like an expert.

This work was supported by NSF grants IIS-2212046 and IIS-2123920 and grants UH3-CA225021, U24-CA215109, and U24-CA180924 from the NCI and NIH.

43.414 THE INFLUENCE OF LOCAL AND GLOBAL TEMPORAL REGULARITIES ON VISUAL DETECTION

Lizzie Gao[1], Anna C Nobre[1,2]; [1]Yale University, [2]University of Oxford

The natural flux of the dynamic external environment contains embedded temporal regularities across multiple time scales. Previous studies have examined how local temporal regularities guide perception and behavior, showing that items appearing at predictable moments are processed better than unanticipated items. In parallel, studies have examined how segmenting ongoing stimulation into discrete events modulates psychological functions and temporal judgments of stimuli. In this study, we explored how temporal regularity at two timescales – intervals between stimuli and durations of events – impacted stimulus processing, aiming to uncover the mechanisms of nested hierarchical temporal structures and their putative interactions. 20 healthy participants viewed rapidly presented faces (100 ms) from the Chicago Face Database. Their task was to respond as quickly and accurately as possible every time a pre-designated target face appeared. Targets appeared sparsely, once per eight faces on average, among a stream of another repeated non-target face. The target face remained the same throughout the experiment, but the non-target face changed on occasion, demarcating separate temporal contexts. Temporal expectations were manipulated both locally and globally. Locally, the inter-stimulus interval between faces could be fixed (consistently short (500 ms) or long (1400 ms)) or variable (intermixed short and long). Globally, context durations were fixed (6 seconds) or variable (4–24 seconds). Reaction times (RT) and accuracy were measured. Consistent with previous findings, local temporal expectations significantly improved RTs for faces at predictable short intervals. Response times to targets were slower soon after the non-target face changed and decreased over time, showing an effective induction of contextual boundaries. Behavioral measures were not sensitive to the global temporal predictability of contexts. RTs to targets were similar before and after predictable or variable boundary changes. Ongoing neural studies will explore the levels of stimulus processing impacted by local and global temporal structures with greater granularity.

43.415 THE TIME COURSE OF FOVEAL AND PERIPHERAL INFORMATION INTEGRATION DURING DYNAMIC GAZE-CUEING

Srijita Karmakar[1] (srijita@ucsb.edu), Miguel P. Eckstein[1]; [1]Department of Psychological and Brain Sciences, University of California, Santa Barbara

Introduction: The gaze of others carries rich information about their intentions and future actions, and orients the observer's overt and covert attention (Bayliss & Tipper, 2007; Driver et al., 1999; Friesen, Ristic, & Kingstone., 2004; Han and Eckstein, 2023). Here, we quantify the time course of foveal and peripheral information integration influencing observers' perceptual decisions with dynamic gaze cues and overt attention. Methods: Eight observers searched for a bright spatial Gaussian target (SD = 0.93°) that could be present on the left or right of a central gazer (10.5° eccentricity). Target and distractor luminances were perturbed with independent Gaussian noise every 25 ms over a presentation time of 825 ms. In the gaze-cued condition, the central gazer head turned and cued either the left or the right location with 80% cue validity. In the neutral-cue condition, the central gazer did not turn to cue either location. Observers made free eye movements and their final perceptual decisions about target presence and location were recorded. Results: Saccade-aligned temporal classification images were computed to find time windows critical for foveal and peripheral information integration. In the neutral-cue condition, observers integrated pre-saccadic peripheral stimulus information presented 80-120 ms before the first saccade onset and post-saccadic foveal information, presented at subsequently fixated locations, 100-250 ms after saccade execution. Notably, dynamic gaze cues dampened the weighting of the pre-first-saccade peripheral information presented before the completion of the gazer's head movement (80-150 ms before saccade onset). The present results differ from a similar gaze-cueing paradigm but with covert attention which showed increased weighting of peripheral stimulus information before the completion of the gazer's head movement. Conclusion: Together, we find that observers' perceptual decisions are guided both by pre-saccadic peripheral and post-saccadic foveal information, the time course of which is delayed in the presence of dynamic gaze cues.

43.416 VOLUNTARY TEMPORAL ATTENTION IMPROVES PERCEPTION EVEN IN THE ABSENCE OF TEMPORAL COMPETITION

Jennifer Motzer[1], Karen Tian[1], Rachel Denison[1]; [1]Boston University

Motivation: Stimuli that appear close together in time compete for representational resources, impairing perception in a phenomenon called temporal competition. Temporal competition can be biased by directing voluntary temporal attention to improve perception at behaviorally relevant moments while diminishing perception at irrelevant moments. It is unknown if this trade-off arises because temporal attention selects among actively competing stimulus representations, such as within working memory, or if temporal attention biases stimulus representations prior to a competitive stage. Here we used a temporal cueing task that independently manipulated temporal attention and temporal competition to test whether attention affects perception even without competition. Methods: Human participants (n=15) performed a challenging perceptual task where we manipulated temporal attention using a precue and temporal competition by varying the number of gratings that appeared on each trial in a single location. Gratings could appear at two possible time points, T1 and T2, separated by 250 ms, in four possible sequences: both present, both absent, or present at either T1 or T2 only. When present, gratings were tilted clockwise or counterclockwise from either the vertical or horizontal axis. On each trial, a precue tone directed attention to either one or both time points. A response cue tone matched the precue on 75% of trials and indicated which time point to report. Participants reported whether or not they saw a grating at the response-cued time point and, if they saw the grating, its tilt. Results: If temporal attention only resolved competition between stimulus representations, it should have improved performance only on two-target trials. However, we found that temporal attention enhanced perceptual sensitivity to a similar degree even without temporal competition, on one-target trials. Therefore, the results suggest that voluntary temporal attention selects stimuli by enhancing perception at attended time points prior to a competitive stage.

Startup funds from Boston University to R.D. and National Defense Science and Engineering Graduate Fellowship to K.T.

Attention: Features, objects

Monday, May 19, 8:30 am – 12:30 pm, Pavilion

43.417 ATTENTION IN FLUX: OBJECT-BASED ATTENTION IS FLEXIBLE TO BOTH LOW- AND HIGH-LEVEL CHANGES IN REAL-WORLD OBJECTS

Kelly McEvoy[1], Dick Dubbelde[2], Sarah Shomstein[1]; [1]The George Washington University, [2]Georgetown University

Attentional selection operates on an object-based representation using both simple geometric shapes (e.g., rectangles) 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 meaning) to object-based attentional selection in real-world objects. We adapted a two-rectangle task in which following a brief exogenous cue (150 ms) and a delay (70 ms), the attended object abruptly changed according to one of four conditions: an object with the same border but different semantic category, an object with a different border but same semantic category, an 'outlier' object with neither the border nor category in common, or no change at all. 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 (OBE) compared to border and category. Objects that undergo no change will mimic classic object-based attentional selection (e.g., same object advantage). As predicted, OBEs were observed for all object change conditions, excluding outlier objects, with a significantly larger OBE for same category objects. These results were replicated in several follow-up experiments. Overall, object-based attentional guidance was modulated by the degree of low- and high-level changes in object properties. Our results suggest that object-based attention persists despite altering object properties in real-world objects, informing current models of attentional mechanisms and extending them to more naturalistic environments.

NSF BCS 2022572 to SS

43.418 EVIDENCE FOR A SHAPE-SIMILARITY GAIN MODEL FOR OBJECT-BASED ATTENTION

Brendan Valentine[1] (valen207@msu.edu), Xiaoli Zhang[1], Melisa Menceloglu[1], Taosheng Liu[1]; [1]Michigan State University

Previous work on feature-based attention has established two prominent models of the selection profile: feature-similarity gain and surround suppression. The former predicts a monotonic decrease in task performance as the target feature becomes more different from the attended feature, whereas the latter predicts a non-monotonic performance pattern where the lowest performance occurs for targets close to the attended feature with a rebound in performance for more distant features. While support for both models have been found using simple features, it is unclear whether the selection profile for object-based attention aligns with either model. The current study assessed the selection profile for simple shapes, as a first step toward more parametric investigations of object-based attention. The study used a newly developed standardized circular shape space that allowed object difference to be quantitatively measured. In two experiments, participants were directed to attend to two target shapes that systematically varied along the shape circle. Two distractor shapes then appeared, overlapping with the target shapes, and one shape in each pair underwent a brief luminance change. Participants reported the status of each target shape (no change, dimmer, brighter). Experiment 1 used finer sampling of the shape space with a maximum target difference of 90°, and Experiment 2 used a coarser sampling with maximum target difference of 180°. For both experiments, performance accuracy peaked when the two target shapes matched and then decreased in a monotonic manner as the two shapes became more different. These results align more with the feature-similarity gain model and suggest that an analogous shape-similarity gain effect operates at a higher level of complexity. Such a gain effect may support object-based selection to differentiate target objects along higher-order, holistic dimensions like shape.

This work was supported by NSF grant 2019995.

43.419 THE EFFECT OF FEATURE CHANGES ON MULTIPLE OBJECT TRACKING

Rachel Eng[1], Lana Trick[1]; [1]University of Guelph

Multiple object tracking (MOT: Pylyshyn & Storm, 1988) is the ability to monitor the positions of a subset of identical items (targets) among identical non-target items (distractors). This ability is believed to be required for everyday activities such as keeping track of children in a crowd or driving a vehicle. However, most real-world situations involve items that differ in surface features (e.g., colour and shape). Our previous research investigated the effect of target similarity and item uniqueness by using displays of 16 items that varied on two feature dimensions: colour (red, blue, green, yellow) and shape (circle, triangle, square, cross), such that each item was a unique combination. Every trial had four targets and 12 distractors. Targets could have the same colour or shape (Colour-share and Shape-share conditions, respectively), or no common features (e.g., the No Share condition: e.g., red circle, blue triangle, green square, yellow cross). We found that performance was significantly better when targets shared a colour or shape than when they did not (target similarity effect), though even performance in the No-share condition was superior to that when items were identical (the uniqueness benefit). To determine whether these two effects were stable across featural change, we compared performance in the four conditions when the items retained their colours and shapes to when they adopted novel colours and shapes during item motion, manipulating whether the items preserved feature grouping (e.g., all red items become purple) or disrupted it (e.g., red items become purple, pink, aqua, brown). In the Colour and Shape Share conditions, performance was significantly impaired when the change disrupted the grouping but not when grouping was preserved. In contrast, colour and shape changes did not affect performance in the No-share or Identical item conditions. This suggests the uniqueness benefit and target similarity effects reflect different mechanisms.

We would like to thank NSERC for funding this research.

43.420 ELECTROENCEPHALOGRAM DECODING OF THE ATTENTIONAL SELECTION AND TRACKING OF FEATURELESS OBJECTS

Henry Jones[1] (henryjones@uchicago.edu), Dawei Bai[2], Brian Scholl[2], Edward Awh[1]; [1]University of Chicago, [2]Yale University

Recent work leveraging multivariate decoding of EEG data has identified a signal that scales with the number of items in working memory (WM). This signal appears to be content-independent, generalizing across distinct visual features, and between single-feature and multi-feature items. One explanation for this content independence is that this signal reflects an abstract indexing process that binds items to their context in space and time for maintenance and access. To explore this possibility, we examined whether a similar load signal exists for “featureless objects”, which have no enduring properties from moment to moment. On each trial, participants viewed a dense grid of crosses of random orientations. A moving object was implemented by having a cross change from one random orientation to another, with these changes propagating through space and time. These transients yield a persisting trackable object, even though (a) there is no surface feature that is constant from one frame to the next, and (b) it is impossible even in principle to identify an object in any static frame. In the actual experiment, participants viewed a set of such moving featureless objects at once, and were cued to track 1 or 2 of them. After tracking the cued item(s), there was a brief delay, and then participants had to discriminate between the true final location of an object, and a nearby alternative location. EEG decoding found a signal that scaled with the number of featureless objects and generalized between all 3 phases of the trial: cueing, tracking, and the pre-test delay. In next steps, this neural signature will be compared directly to the previously identified WM load signal. Generalization of these load signals across task contexts would provide support for the theory that spatiotemporal indexing plays a role in WM maintenance.

43.421 FEATURE- VERSUS OBJECT-BASED ATTENTIONAL TEMPLATES DURING FEATURE, CONJUNCTION, AND OBJECT SEARCH

Rai Samar Ghulam Bari[1], Ziyi Wang[1], Anna Grubert[1]; [1]Durham University

Visual search is guided by attentional templates, i.e., target representations held in visual working memory. The content of such templates is still under debate, especially when targets are defined by combinations of task-relevant features. Previous research showed that target templates during conjunction search are feature- rather than object-based. However, these studies often compared guidance during feature versus conjunction search but failed to include (true) object-based control conditions. In this study, we systematically compare attentional guidance (as indexed by the N2pc component of the event-related potential) during feature, conjunction, and object search. We developed a stimulus set with 81 naturalistic object images (9 different object categories, each with 9 members defined by different combinations of 3x3 colours and shapes). Target identities were kept constant in each block in Experiment 1 but were cued on a trial-by-trial basis in Experiment 2 to increase the perceptual strength of the target representations and minimise potential long-term memory effects. In different tasks, participants searched for targets with a specific colour or shape (feature search), colour/shape combination (conjunction search), or object category (object search). Results were consistent in both experiments. Reaction times were fastest in the feature search and slowest in the object search with an intermediate conjunction search. N2pc amplitudes and latencies were identical in feature and conjunction search and partially matching distractors captured attention in the conjunction task. This mirrors previous findings that have demonstrated initial feature-based guidance during conjunction search. In contrast, and even though stimuli were physically identical, N2pc amplitudes were attenuated and N2pc latencies were delayed during object search. Furthermore, partially matching distractors were completely ignored. This suggests that attentional templates can contain holistic object-representations if required by task demands but that attentional guidance by such object-based templates is less efficient than feature-based guidance.

This work was funded by research grants of the Leverhulme Trust (RPG-2020-319) awarded to AG.

43.422 FEATURE INTEGRATION THEORY REVISITED: ATTENTION IS NOT NEEDED TO BIND STIMULUS FEATURES, BUT PREVENTS THEM FROM FALLING APART.

Surya Gayet[1] (surya.gayet@gmail.com), Süheyla Aydemir[1]; [1]Utrecht University, Helmholtz Institute

Any object that we look at comprises many different properties (e.g., a color, shape, orientation, and spatial location), all of which are processed in separate brain regions. According to Anne Treisman’s influential Feature Integration Theory of attention, these individual properties are processed pre-attentively, whereas attention is needed to 'bind' them into integrated objects. Here, we ask during which processing stage attention contributes to binding: during stimulus encoding (to encode stimuli as integrated objects) or during maintenance (to prevent objects from falling apart into their constituents)? In eight lab-based behavioral experiments (N = 8*34), participants memorized four briefly presented visual objects. To manipulate the amount of attention, two of these objects were (pre or post) cued to be likely targets (with 75% validity). After a delay, one object was probed, and participants had to report its features from memory (e.g., color, location, orientation, category). Recall performance was far better for cued than for uncued items, indicating the success of the attention manipulation. Also, stimulus properties were generally bound, so that recall performance on one feature strongly predicted recall performance on another feature of that same object. However, attention only increased binding strength between features (e.g., color and orientation), but not between features and locations, suggesting that feature-location binding is a pre-attentive process. Finally –and most surprisingly– the effect of attention on (feature-feature) binding did not depend on whether the cue preceded or succeeded the memory array, showing that attention contributes to binding exclusively during memory maintenance. These findings expand the Feature Integration Theory of attention, by showing that (1) not only the encoding of features occurs pre-attentively, but also the binding of features to locations, and by showing that (2) the role of attention is not to integrate objects during stimulus encoding, but to keep features bound together during stimulus maintenance.

43.423 IMPLICIT FEATURE-BASED SUPPRESSION IS EFFECTIVE AND ROBUST, EVEN IN THE FACE OF FEATURE-BASED GAIN, WHILE EXPLICIT FEATURE-BASED SUPPRESSION IS INEFFECTIVE AND WEAK

Andrew Rodriguez[1] (rodri818@msu.edu), Morgan R. Dodd[1], Mark W. Becker[1]; [1]Michigan State University

In visual search literature, evidence for feature-based suppression (FBS) has been controversial and mixed. Much of the mixed findings are due to differences between implicit FBS, the unconscious ignoring of a feature that appears as a distractor with high regularity, and explicit FBS, the active biasing of attention away from a feature. Given this discrepancy, we examined the benefits of implicit and explicit FBS in a Landolt C search task. Across four experiments, we had participants search for a C with a horizontal gap among several distractor Cs with vertical gaps. This search task had two phases. The first phase had one color (ignored color) that appeared on a higher percentage of trials (80%) and always appear as a distractor. The second phase removed the contingency - the previously ignored color now could appear as the target. For the experiments that measured the implicit suppression effect, participants were not made aware of the ignored color. For the experiment that measured the explicit suppression effect, participants were cued that the target would not appear in the ignored color. Our results showed a robust implicit suppression effect, with faster RTs when the ignored color was present as a distractor than when ignored color was absent. Furthermore, the implicit suppression benefit required minimal training of the to-be-ignored feature and persisted even when participants were given positive cues about the target's color. However, we found a weak effect of explicit suppression that required extensive training before it became beneficial. Our findings suggest that implicit FBS can effectively guide attention while explicit FBS is ineffective unless there is ample training of the to-be-ignored feature, limiting its utility as a real-world mechanism for allocating attention.

43.424 INCIDENTAL LEARNING ABOUT RELEVANT AND IRRELEVANT FEATURE VALUES ENHANCES EARLY STAGES OF ATTENTIONAL SELECTION

Kevin Ortego[1], Douglas Addleman[2], Viola Stoermer[1]; [1]Dartmouth College, [2]Gonzaga University

Statistical regularities in the environment influence how our brains process and prioritize information, even when we are not explicitly aware of these regularities. For example, a target item occurring more frequently in a particular location or color will lead to faster response times in visual search tasks. Similar behavioral benefits are observed from learning that certain features are commonly irrelevant (e.g. that items in a particular color are frequently distractors). It is currently debated whether these learning benefits are realized via enhancements at early processing stages or through later modulations of response and decision thresholds. We tested this by recording EEG in two experiments in which participants (N=32, N=25, ongoing) viewed a circular visual search array and identified the gap direction of a Landolt C with a left or right gap among seven distractors with top/bottom gaps. Critically, and unbeknownst to participants, we manipulated the color probabilities such that the target occurred more frequently in one particular color (target feature learning), or such that distractors occurred more frequently in a particular color (with equiprobable target colors; distractor learning). This resulted in "valid" trials where the targets/distractors matched the most likely color and "invalid" trials where this association was reversed. As predicted, responses were faster for valid relative to invalid trials in both experiments, and correspondingly the N2pc, an ERP component associated with early attentional selection, occurred approximately ~30ms earlier on valid trials across both experiments, demonstrating that both relevant and irrelevant feature learning can modulate attentional selection within ~200ms of stimulus onset. Later response and memory processes, indexed by the LPC component, were also modulated by the probability structure, but the pattern differed for relevant and irrelevant feature learning. Together our results suggest similar attentional processing benefits, but differential memory-related effects across these two modes of learning.

43.425 **PROACTIVE ENHANCEMENT OF 'TO**-BE-**ATTENDED' AND 'TO**-BE-**IGNORED' FEATURES DURING** CUED VISUAL SEARCH IN OLDER AND YOUNGER ADULTS

Punyawish Patumhirunruksa[1] (pun@docchula.com), Viola Stoermer[3], Chaipat Chunharas[1,2]; [1]Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, [2]King Chulalongkorn Memorial Hospital, Thai Red Cross Society, Bangkok, Thailand, [3]Dartmouth College

Feature-based attention is an ability used to search for target features among distractors in an environment. Previous research has shown that older adults show deficits in feature-based attention compared to younger adults, in particular when it comes to ignoring distractors. However, these age-related declines appear to depend on how feature-based attention is engaged. Here we tested how proactive attentional control is affected by age. Two groups of participants (10 younger and 10 older adults, >60 years) performed a visual search task in which a 'to-be-attended' or 'to-be-ignored' cue indicated the likely target or distractor color with 80% probability at the beginning of each block. Following the cue, participants were instructed to search for a ring with a horizontal gap among distractors with vertical gaps. To assess whether the cues would modulate the saliency of the cued color proactively, on a subset of trials, participants performed a probe-color detection task in which they were presented with a probe that could either match the cued color or not. Our results showed that both groups' responses during search were reliably faster for valid relative to invalid target cueing trials (younger adults: 795 ms vs. 926 ms, older adults: 1782 ms vs. 2244 ms) but no reliable differences in distractor cuing trials (younger adults: 807 ms vs 828 ms, older adults: 1874 ms vs 1930 ms). For the probe task, both groups' accuracy was higher when the probe appeared in the cued color relative to a neutral gray or the uncued color in both target and distractor cueing conditions. This indicates that older and younger adults initially enhanced the cued color regardless of whether it predicted the target or distractor, suggesting that they failed to proactively ignore the anticipated distractor color.

43.426 THE EFFECT OF INTERNAL ATTENTIONAL SHIFTS ON VISUAL FEATURE ERRORS

Caitlin V. Cunningham[1], Julie D. Golomb[1]; [1]Department of Psychology, The Ohio State University

In visual contexts, spatial attention facilitates processing for items appearing at an attended location and is crucial for binding multiple features together within an object. Previous studies reported that manipulating external spatial attention (e.g., shifting or splitting top-down attention) can cause errors (swapping and/or distortion) when reporting a target item's features. However, less work has investigated errors resulting from manipulations of internal spatial attention. Here, we asked whether spatial shifts of internal attention within working memory (WM) have systematic effects on feature reports, and if so, whether these errors differ from those caused by shifts of external attention. In the current study, participants viewed an array of four colored squares. After a 1s retention period, a spatial postcue oriented internal attention toward one of the items' locations, followed by a 500ms delay. On some trials (two-cue trials), a second postcue appeared, indicating that participants should shift their spatial attention within WM. After a short (50ms) or long (500ms) delay, participants reported the color of the most recently cued item (target) on a continuous color wheel. Using probabilistic mixture modeling, we found that the short delay resulted in performance decrements compared to the long delay in two-cue trials, suggesting that the short delay provided insufficient time for participants to update their spatial attention within WM. Interestingly, while incomplete shifts of external attention have been found to cause swapping errors (higher likelihood of misreporting the color of the first-cued item), we instead found that after the short delay, participants were more likely to report the color opposite, characteristic of feature avoidance errors. These errors were not observed after the long delay, suggesting that while participants were able to shift their internal spatial attention, the new item's features may not have been accessible by the time of report after the short delay.

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

43.427 DIFFERENTIAL NEURAL ACTIVATION FOR SHAPE- AND LOCATION-BASED ATTENTION

Ishita Agarwal[1] (agarw467@purdue.edu), Sung-Mu Lee[1], Jinho Lee[1], Anne B Sereno[1]; [1]Purdue University

Attention selectively modulates neural activity, enhancing the processing of task-relevant visual features. However, it remains unclear how neural activity differs between shape-selective and spatial-selective attentional focus in different brain regions. Many existing attention-related tasks are complex, making them difficult for individuals with attentional deficits to perform, or require a significant amount of practice (e.g., dual-task or task-switching paradigms). Additionally, many imaging studies use visual stimuli such as shapes, objects, or faces without adequately controlling for factors like luminance, color, and size, which introduces variability in brain responses, reducing reliability and sensitivity. Finally, previous paradigms often focus attention on a single dimension of an object (e.g., color or location) rather than examining multiple aspects of attention simultaneously, limiting their generalizability. To address these issues, this study used identical stimuli in a simple, well-controlled experimental paradigm designed to reliably assess the effect of varying attentional focus. Participants performed a 1-back repetition task where, in separate fMRI blocks but using the same stimuli, participants need to pay attention to either the shape or the location of images. Whole-brain analyses across MNI-defined brain regions revealed that attending to shape, compared to passive viewing condition, elicited increased activity in the ventral stream such as temporal parietal junction. In contrast, attending to location, compared to passive viewing condition, predominantly elicited increased activity in dorsal stream such as supramarginal gyrus. These findings highlight region-specific neural activation, demonstrating that shape-selective and spatial-selective attentional focus selectively engages distinct brain regions. We will discuss how these findings advance our understanding of how attention to shape and location of objects dynamically modulates neural activity within the visual system.

NIH CTSI and Purdue University

43.428 PREPARATORY ATTENTION TO VISUAL FEATURES SPREADS GLOBALLY

Melisa Menceloglu[1] (mencelo1@msu.edu), Nora Indenberga[1], Sydney T. George[1], Seeun Oh[1], Søren K. Andersen[2], Taosheng Liu[1]; [1]Michigan State University, [2]University of Southern Denmark

Attention to visual features not only facilitates neural responses to task-relevant stimuli but also to task-irrelevant stimuli at other locations that share the attended feature. This global attentional facilitation of features across the visual field has been reported during active selection (e.g., attending to blue dots among other colored dots). Here, we investigated whether this global spread also occurs during preparatory attention (e.g., preparing to select blue among other colored dots, before the presentation of such dots). We presented two overlapping fields of blue and orange dots to the left and right of a central fixation cross. We recorded EEG while observers (N=20) were cued to attend to a specific dot field (e.g., blue dots on the left) to detect coherent motion at threshold. In the immediate-onset condition, dot fields on both sides were presented simultaneously following the cue; whereas in the delayed-onset condition, attended-side dot fields were presented after a ~1.5-2.7s delay, extending the preparatory period. We used the frequency tagging method to measure the effects of attention on the visual cortical responses by flickering the dot fields at four different frequencies and comparing their inter-trial phase coherence (ITPC). We observed a local feature-based attention effect on the attended side: ITPC elicited by the attended-color dots was stronger than unattended-color dots. We also observed a global feature-based attention effect on the unattended side, in the immediate-onset condition: ITPC elicited by the dots matching the attended color was stronger than unattended color. Notably, we observed a robust, albeit reduced, global effect in the delayed-onset condition during the preparatory period. These results show that preparatory attention to a feature can bias the analysis of that feature in the absence of a stimulus in the spatially attended region, suggesting that the global spread of feature-based attention can be a purely endogenous process.

This work was supported in part by grants from NIH (R01EY032071) and NSF (2019995).

43.429 SPATIAL ATTENTION TO MULTIPLE STIMULI DOES NOT REDUCE EVOKED SSVEP POWER RELATIVE TO FOCAL ATTENTION

Amelia H. Harrison[1], Thomas C. Sprague[1]; [1]University of California, Santa Barbara

Typically, when spatial attention is directed to multiple locations simultaneously, perceptual sensitivity to stimuli at those locations is reduced relative to attending only a single location (e.g., Ling & Carrasco, 2006; Popovkina et al., 2021). Additionally, neural responses to visual stimuli are generally attenuated under distributed attention compared to focal attention conditions (e.g., McMains & Somers, 2005; Toffanin et al., 2009). However, fMRI data from a previous study in our lab show a different pattern: focal spatial attention enhances the neural representation of a stimulus compared to when the stimulus is ignored, but the degree of enhancement does not decrease as an additional stimulus is attended (Harrison et al., VSS 2024). To resolve the discrepancies between this finding and prior literature, we conducted an EEG experiment in which participants performed a selective attention task modeled after our previous fMRI experiment. fMRI and EEG have been suggested to assay complementary attentional modulations arising from distinct physiological processes (e.g., Itthipuripat et al., 2019), motivating its use for better understanding the neural mechanisms surrounding distributed spatial attention. Participants were cued on each trial to attend to the fixation point, to one cued peripheral location, or to two cued peripheral locations where flickering stimuli (20 and 24 Hz) appeared on every trial. In our pilot dataset, we compared steady-state evoked potential (SSVEP) amplitudes for each stimulus across attentional conditions and found that while attending to a single stimulus increases evoked power, the effects of distributed attention, compared to focal attention, are variable. In line with our prior fMRI results, there is no clear evidence that distributed attention reduces attentional enhancement compared to focal attention. This is consistent with a model where attentional feedback signals enhance sensory responses to a similar degree when attention is focused as when it is distributed.

Research supported by a University of California, Santa Barbara Academic Senate Research Grant and an Alfred P Sloan Research Fellowship.

43.430 THE ROLE OF MID-LEVEL VISUAL PROCESSES IN THE WORD IDENTIFICATION BOTTLENECK

Mary Catington[1], Michael Pratte[1]; [1]Mississippi State University

In any particular moment only a subset of incoming visual information can be perceived due to limitations in attentional processing. Recent research has shown that low-level visual information about two objects, such as their color and luminance, can be perceived in parallel. However, high-level information such as word meaning is limited by a serial processing bottleneck, such that only one of two simultaneously presented words can be recognized. Although this limitation could be due to constraints on high-level aspects of word recognition, here we explore the possibility that it is driven by attentional limitations on the mid-level visual processes that make word identification possible. One such mid-level visual process is figure-ground segmentation, by which common local elements are grouped together to form a coherent surface. This is one of several mid-level processes known to rely on attention, and is crucial for boundary ownership and object identification. In Experiment 1 participants attempted to identify coherent textures in noise for one item, or for two items presented simultaneously. The results reveal a cost when attempting to process two items simultaneously such that, unlike color or luminance, texture segmentation could not be performed across multiple objects in parallel. However, unlike word identification, segmentation was not purely serial such that some information about two objects could be processed at the same time. Therefore, the inability to simultaneously read two words observed in previous studies may partially stem from attentional constraints in the mid-level visual processes that are necessary prerequisites for word recognition. There are several such mid-level processes, and future work will examine how limitations across them may give rise to the severe serial bottleneck in word recognition.

This work was supported by National Institutes of Health NIMH Grant #R15MH113075-02A1.

43.431 SELECTIVE ATTENTION WARPS THE REPRESENTATION OF SPACE THROUGHOUT CORTICAL VISUAL NETWORKS

Jen D. Holmberg[1], Emily X. Meschke[1], Yuerou Tang[1], Amanda LeBel[1], Jack L. Gallant[1]; [1]UC Berkeley

Visual attention systems in the brain selectively enhance visual information that is relevant to the target or goal (Cukur et al., 2013). Most object-based visual attention studies have focused on how attention affects the representations of high-level semantic categories or objects. However, high-level semantic category representations rely on the integration of low-level visual representations from earlier stages of visual processing. Given that the visual system is densely interconnected, it is important to understand how object-based attention modulates low-level visual representations to facilitate downstream object-based perception. Here we examine whether object-based attention affects spatial representations across cortical regions of interest (ROIs) during a naturalistic movie-watching task.

Six participants watched compilations of movie clips either while fixating passively or while covertly searching for “humans” or “vehicles” in the movies (Cukur et al., 2013). High-level semantic features and low-level visual features were recovered from the movies, and these were used to fit voxelwise encoding models to predict brain responses in each participant, under each attention condition (Dupré la Tour et al., 2022). Task-specific functional networks were recovered using model connectivity (Meschke et al., 2022), and were used as ROIs in the analyses. To assess semantic representation changes across attention conditions, the semantic model weights were correlated with ‘human’ and ‘vehicle’ semantic templates (as in Cukur et al., 2013). To assess spatial representation changes across attention conditions, the visual model weights were correlated with a spatial template that partitioned the visual field into ‘center’ and ‘periphery’. These analyses revealed that selective attention to either semantic target shifts spatial representations towards the periphery in the majority of networks. Furthermore, in some networks, the magnitude of the spatial representation shift was correlated with the magnitude of the semantic representation shift. These results suggest that object-based selective attention is supported by changes in low-level spatial representations.

Object Recognition: Neural mechanisms

Monday, May 19, 8:30 am – 12:30 pm, Pavilion

43.432 CHEMOGENETIC INVESTIGATION OF IMAGE RECOGNITION IN RHESUS MONKEYS

Phelix Rodriguez[1], Bing Li[1], Mark AG Eldridge[1,2], Walter Lerchner[1], Barry J Richmond[1]; [1]National Institute of Mental Health, [2]Newcastle University

Image recognition depends on the ventral visual stream. Inferior temporal cortex (ITC) is a brain region thought to be the last element of the ventral visual stream. Previous studies have shown that neurons in higher level visual areas, including area TE, a subregion of ITC, exhibit less activity when presented with stimuli that subjects have seen previously, an effect known as ‘repetition suppression’. We have demonstrated that the bilateral lesion of area TE causes deficits in object recognition. We examined the role of area TE in image recognition using an inhibitory chemogenetic agent. Two rhesus monkeys received bilateral injections of virus expressing inhibitory DREADDs (designer receptors exclusively activated by designer drugs) to area TE. Subjects were trained to complete a ‘serial recognition’ task by incorporating progressively longer intervals (more intervening trials) between first and second presentations of a novel stimulus. Subjects were tested on two versions of the serial recognition task with all interval sizes (0-128 trials between first and second presentation), one with continuously presented stimuli and the other with briefly presented stimuli, after receiving a systemic injection of either saline or DREADD-activating ligand, deschloroclozapine (DCZ). Early results show a reduction in overall task performance and a higher rate of objects reported as familiar using briefly-presented stimuli but not on the task using continuously-presented stimuli. Thus, area TE is needed to identify a briefly-presented stimulus as unfamiliar. The inhibitory effects of DREADDs on TE activity may mimic the repetition suppression, which leads to increased reporting of familiarity.

Intramural Research Program; National Institute of Mental Health; National Institutes of Health; Department of Health and Human Services (annual report number ZIAMH002032)

43.433 VISUAL STIMULI AMPLIFIES THE EFFECT OF OPTOGENETIC STIMULATION IN THE INFEROTEMPORAL CORTEX IN MONKEYS

Alvin Dinh[1], Behnam Behnam Karami1[1,2], Reza Azadi[1], Arash Afraz[1]; [1]1National Institute of Mental Health, Bethesda, MD, [2]Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

We have previously demonstrated that the detectability of stimulation in the inferior temporal (IT) cortex of monkeys varies with the visual stimuli presented. Particularly, detectability is lowest when no image is shown to the animals. This suggests that the perturbability of neurons in the IT cortex depends on the cortical activity state, which is influenced by visual inputs. Consequently, identical cortical stimulation can evoke different levels of neuronal activation depending on the visual stimulus. To investigate this interaction, we employed optogenetic stimulation using the excitatory vector pAAV-CaMKIIa-C1V1(t/t)-TS-EYFP, targeting excitatory neurons. Monkeys were trained to maintain central fixation, while a series of images were presented for 1 second each, with 200 ms intervals between images. During half of the image presentations, randomly selected, a square light pulse was delivered to the transduced tissue 400 ms after image onset, lasting for 200 ms. Our results indicate that in conditions where no image was displayed, neural activity during cortical stimulation was significantly lower. This finding aligns with the animals' reduced performance in detecting cortical stimulation in the absence of visual stimuli, suggesting that the perturbability of the stimulated site is modulated by visual input.

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

43.434 DYNAMIC OBJECT PROCESSING IN MACAQUE IT CORTEX: TEMPORAL DYNAMICS AND MODEL LIMITATIONS

Matteo Dunnhofer[1,2] (matteo.dunnhofer@uniud.it), Christian Micheloni[2], Kohitij Kar[1]; [1]York University, [2]University of Udine

The macaque inferior temporal (IT) cortex, the apex of the ventral visual stream, plays a crucial role in object recognition. While most studies have relied on static images to predict neural responses and build encoding models, how dynamic visual inputs are transformed into IT responses remains incompletely understood. Here, we presented 200 videos (500 ms each, 18 frames at 60 Hz) to two macaques while recording neural activity from 132 reliable sites in the IT cortex. Each video contained objects moving within a fixed background, enabling us to investigate the temporal dynamics of IT population responses. We addressed three key questions. First, we asked whether IT neurons exhibit reliable activity beyond the initial transient response (70–150 ms) in response to videos. Neural responses demonstrated significant reliability throughout the video duration (~0.62 Spearman-R split-half correlation), suggesting that dynamic stimuli engage sustained processing in the IT cortex. Second, we assessed whether these later responses could predict object identity using linear decoders. Decoding performance was significantly above chance (~58% classification accuracy ~500ms post video onset, chance-level=10%), indicating that IT activity carries discriminative information for object recognition beyond the initial response phase. Third, we evaluated how well feature activations from feedforward models (e.g., convolutional neural networks) could explain IT responses along its entire reliable dynamics. The early (70-170ms) responses were significantly better predicted (~51% explained variance) by any frame-based model activation compared to later response phases (470-570 ms, %EV ~19%), highlighting a critical limitation of feedforward architectures in accounting for dynamic neural processing. Our findings reveal that dynamic stimuli elicit sustained and informative responses in the IT cortex. The inability of standard feedforward models to explain later neural responses suggests the need for models incorporating recurrent or temporal mechanisms to explain IT representations better.

MD was funded by the European Union (MSCA Project 101151834 - PRINNEVOT). KK was supported by Canada Research Chair Program, Simons Foundation Autism Research Initiative (SFARI, 967073), and a NSERC DG.

43.435 A DATA-DRIVEN ANALYSIS OF THE PERCEPTUAL AND NEURAL RESPONSES TO NATURAL OBJECTS REVEALS ORGANISING PRINCIPLES OF HUMAN VISUAL COGNITION

Tim Andrews[1] (ta505@york.ac.uk), David Watson[1]; [1]University of York

A key challenge in understanding the functional organisation of visual cortex stems from the fact that only a small proportion of the objects experienced during natural viewing can be presented in a typical experiment. This constraint often leads to experimental designs that compare responses to objects from experimenter-defined stimulus conditions, potentially limiting the interpretation of the data. To overcome this issue, we used images from the THINGS initiative, which provides a systematic sampling of natural objects. A data-driven analysis was then applied to reveal the functional organisation of the visual brain, incorporating both perceptual and neural responses to these objects. Perceptual properties of the objects were taken from an analysis of similarity judgements, whereas neural properties were taken from through whole brain fMRI responses to the same objects. Partial least squares regression (PLSR) was then used to predict neural responses across the brain from the perceptual properties while simultaneously applying dimensionality reduction. The PLSR model accurately predicted neural responses across visual cortex using only a small number of components. These components revealed smooth, graded topographies, which were similar in both hemispheres, and captured a variety of object properties including animacy, real-world size, and object category. However, they did not accord in any simple way with previous theoretical perspectives on object perception. Instead, our findings suggest that the visual cortex encodes information in a statistically efficient manner, reflecting natural variability among objects.

43.436 BOUND BY SIGHT, SHARED IN THE MIND: THE TEMPORAL EMERGENCE OF OBJECT CO-OCCURRENCE REPRESENTATIONS

Lu-Chun Yeh[1] (luchun.yeh@gmail.com), Marius V. Peelen[2], Daniel Kaiser[1,3]; [1]Justus Liebig University Gießen, Germany, [2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, The Netherland, [3]Center for Mind, Brain and Behavior (CMBB), Philipps University Marburg, Justus Liebig University Gießen, and Technical University Darmstadt, Germany

In real-world scenes, objects reliably appear alongside other objects (e.g., pots, forks, mugs, and food can all be found in the kitchen). Prior fMRI work suggests that object representations in the parahippocampal place area are organized by such co-occurrence statistics among objects. However, it remains unclear when these representations emerge, whether they differ from other representational dimensions (e.g., category), and whether they reflect a mutual priming between co-occurring objects or the activation of shared scene representations. To investigate these questions, we conducted an EEG experiment where participants viewed individual objects while performing a one-back task. We orthogonally manipulated the objects' co-occurrence context (kitchen vs. garden) and category (tool vs. non-tool), and matched their overall shape across the contexts and categories. To characterize the time course of representations related to co-occurrence and category, linear classifiers were trained on EEG response patterns from a subset of objects across the co-occurrence contexts or object categories and tested on an independent subset of objects. Results revealed orthogonal representations of co-occurrence and category, which both emerged during an early (co-occurrence: 116–263 ms, category: 170–255 ms) and a late (co-occurrence: 368–467 ms, category: 274–499 ms) time window. To examine whether representations of co-occurrence are driven by the activation of common scene representations, we trained classifiers on EEG response patterns evoked by kitchens and gardens (obtained in separate runs) and tested them on response patterns evoked by objects associated with the two contexts. Accurate cross-classification from 360–499 ms suggested that late representations of object co-occurrence are indeed driven by the activation of the associated scene context, while earlier representations likely reflect a co-activation of objects. Together, our findings delineate how objects are rapidly associated with one another and, in turn, their common context to facilitate behavior in typically structured environments.

LCY is supported by the MSCA programme (101149060). DK is supported by the DFG (SFB/TRR135,222641018; KA4683/5-1, 518483074, KA4683/6-1, 536053998), "The Adaptive Mind", and an ERC Starting Grant (PEP, ERC-2022-STG 101076057).

43.437 CORTICAL REPRESENTATIONS SUPPORTING COARSE AND FINE OBJECT CATEGORIZATION

Margaret M Henderson[1] (mmhender@cmu.edu), Sungjoon Park[1], Leila Wehbe[1], Michael J Tarr[1]; [1]Carnegie Mellon University

Intermediate visual features may be sufficient to support certain types of object categorization, even in the absence of recognizable high-level properties. We hypothesize that the role of these features depends on the granularity of a task: fine-grained distinctions (types of birds) may require high-level, complex features, while coarser distinctions (animals vs. vehicles) may be accessible from low-level features alone. Differences in the diagnostic features relevant to these different tasks may also lead to task-dependent differences in how object images are cortically encoded during coarse and fine categorization. Across both behavioral and fMRI studies, we leveraged a computational texture synthesis procedure (Gatys et al.; 2015, NeurIPS) to generate "texturized" versions of target object images by matching their summary statistics at different layers of a deep neural network. This results in images that vary continuously in their feature complexity. Participants viewed these images and performed a 2-alternative forced choice task discriminating the image category at either a coarser (superordinate) or a finer (basic) level. We found that observers could behaviorally discriminate a subset of object categories at an above-chance level based on the simplest texturized images tested. Categorization of simple texture images was highest for coarse categories (vs. fine), natural objects (vs. artificial), and color images (vs. grayscale). In the brain, we used multivariate classification within higher visual cortex to demonstrate evidence for distributed neural representations of both coarse and fine object categories. As in the behavioral data, discriminability of these representations was lower for texturized as compared to original images, but coarse category information could be decoded with above-chance accuracy even from cortical responses to simple texturized images. Taken together, these results indicate that intermediate visual features contribute to object categorization in a manner that depends on task precision.

This study was funded by a grant from Apple Inc to MJT, and a postdoctoral fellowship from the Carnegie Mellon Neuroscience Institute to MMH.

43.438 DISTINCT CONNECTIVITY FINGERPRINTS REVEAL FUNCTIONAL SPECIALIZATION IN THE VENTRAL VISUAL STREAM

Isaac Liao[1] (liao.798@osu.edu), David E. Osher; [1]The Ohio State University

The ventral visual stream consists of a mosaic of regions specialized for processing distinct high-level visual categories like faces, bodies, objects, and scenes. It has been argued that the functional activations of brain regions rely on their specialized connectivity patterns; the function of each region is determined in large part by its unique set of connections to the rest of the brain. However, the connectivity patterns that characterize these ventral visual regions are still poorly understood. Here, we characterized the connectivity fingerprints for 18 functional regions of interest (fROI) by building linear models to predict each region based on its connectivity patterns to the whole brain, at the level of individual subjects. We then compared the connectivity patterns that predict each fROI with their task-based response. We found that, in general, the activity of a fROI tends to be predicted by distant regions that share selectivity of the same stimulus domain. The connectivity fingerprints for a fROI also tend to be most similar to regions that perform the same function. Notably, some fROIs have characteristic connectivity patterns that may distinguish them within and between domains. Together, our results demonstrate common motifs in the connectivity patterns of high-level visual regions. Future work will focus on using TMS to non-invasively disrupt the regions that predict function in the ventral visual stream.

43.439 ENHANCED PROTOCOL FOR ISOLATING HIGH-LEVEL VISUAL RESPONSES USING SSVEP

Ana Rozman[1] (a.rozman@sussex.ac.uk), Abigail Flowers[1], Jenny Bosten[1]; [1]University of Sussex

SSVEP signals are strongly influenced by low-level visual features, and it is challenging to use them to extract responses to higher level visual features. Aiming to improve on existing approaches to isolating high-level (e.g., object-specific) responses using SSVEP, we developed an enhanced implementation of image scrambling methods. We presented food texture stimuli modulating at 2 Hz in 4 conditions, and recorded 64-channel EEG while participants passively viewed trials of 30 seconds duration. Two of the stimulus conditions were based on existing methods for targeting object responses using SSVEP: Semantic wavelet-induced frequency tagging (SWIFT: Koenig-Robert & VanRullen, 2013, NeuroImage), and a progressive (sinusoidal) phase scramble. Two additional conditions controlled for modulation of low-level image elements within each approach, where two different phase-scrambled images or two different SWIFT-scrambled images were modulated. In current implementations of SWIFT or phase-scrambling, SSVEP signals are assumed to isolate high level responses if low-level image features are constant over the stimulus modulation. Our additional conditions account for remaining modulations of low-level features when we compare SSVEP signals for (i) original versus scrambled stimulus modulations and (ii) scrambled versus scrambled stimulus modulations. We make these comparisons by regressing (i) against (ii) across the 64 channels, and analysing the residuals, using baseline-corrected and summed SSVEP amplitudes in the frequency domain for the harmonic series of 2 Hz. Results show that differences between conditions in signals captured by residuals are more anterior than signals based on raw modulations (e.g., SSVEPs to SWIFT modulations alone). These findings suggest that our protocol can better isolate a signature of higher-level responses to passively viewed object textures than scrambling methods without the additional control. Our next aim is to adapt these methods to study higher-level color representations using SSVEP.

The work was funded by ERC grant 949242 COLOURCODE to JMB.

43.440 NATURALISTIC STIMULATION ELICITS CATEGORY-SELECTIVE NEURAL RESPONSES

Xueying Ren[1] (xueying.ren@vanderbilt.edu), Kendrick Kay[2], James Booth[1], Franco Pestilli[3], Sophia Vinci-Booher[1]; [1]Dept. of Psychology and Human Development, Vanderbilt University, Nashville, Tennessee, USA, [2]Dept. of Radiology, University of Minnesota, Minneapolis, Minnesota, USA, [3]Dept. of Psychology, University of Texas at Austin, Austin, Texas, USA

Cortical regions within the ventral visual stream demonstrate category selectivity, with areas such as the fusiform face area showing increased activity when viewing faces compared to other categories. Traditional methods for localizing category-selective responses typically use isolated stimuli on a gray background with a central fixation, potentially limiting our understanding of how categories are processed during real-world visual experiences. This study explored how the human brain processes four different visual categories embedded in complex scenes under free-viewing conditions to offer insight into how the brain processes and integrates visual information in real-world settings. To this end, we collected fMRI data from 10 adult participants as they viewed images selected from the Sesame Street Archive, which includes a prevalence of categories uncommon in other image datasets. Each image featured at least one of four categories: faces, places, words, and numbers, which are prevalent in educational and everyday contexts. Univariate analyses at the subject level revealed individual variability in brain activations for different categories. Notably, faces and places elicited relatively more consistent activation patterns across participants, whereas representations for words and numbers exhibited larger individual differences. These findings suggest that certain category-selective brain responses, particularly for faces and places, can be effectively captured using complex scenes under free-viewing conditions. However, we observed individual variability in responses to learned object categories, such as words or numbers, which may reflect differences in reading and math abilities. By using naturalistic stimulation with rich visual contexts, this research provides new insights into the nuances of visual processing at the individual level.

5R01HD114489

43.441 SINGLE PULSE ELECTRICAL STIMULATION REVEALS STRONGER INPUTS THAN OUTPUTS IN DORSAL AREAS

Maria Guadalupe Yanez-Ramos[1] (yanez-ramos.mariaguadalupe@mayo.edu), Gabriela Ojeda Valencia[1], Harvey Huang[2], Nick Gregg[3], Zeeshan Qadir[1], Morgan Montoya[1], Kendrick Kay[5], Greg Worrell[3], Kai Miller[4], Dora Hermes[1]; [1]Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, [2]Mayo Clinic Medical Scientist Training Program, Rochester, MN, [3]Department of Neurology, Mayo Clinic, Rochester, MN, [4]Department of Neurological Surgery, Mayo Clinic, Rochester, MN, [5]Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, Minneapolis, MN

Neural processing in areas along the visual pathways is influenced by inputs from connected visual areas. Although the connectivity between these areas has been studied using post-mortem dissection and diffusion MRI, little is known about the directionality of these connections in the living human brain. To understand the strength and directionality of human brain connectivity, capturing the inputs to one region when stimulating another is necessary. Cortico-cortical evoked potentials (CCEPs) recorded via stereo-electroencephalography electrodes are particularly suited for this purpose. A previous study using cortico-cortical spectral responses showed that visual feedforward connectivity was larger compared to feedback connectivity. Here, we leverage CCEPs to quantify the strength of the connectivity between visual pathways. We recorded CCEPs in 11 patients and used ‘Canonical Response Parametrization’ to capture CCEPs independent of their waveform. Connectivity strength was quantified in each direction by average response reliability, measured as the coefficient of determination (CoD), and connectivity sparsity as the percentage of significant connections from all possible connections. First, feedforward connections from V1-V3 to higher-order visual areas were more reliable and less sparse compared to feedback connections, similar to previous studies. Across 702 possible feedforward connections, the average CoD was 0.164 (SEM = 0.010), with 19-49% significant connections, whereas across 708 possible feedback connections, the CoD was lower (0.081 ± 0.007), with only 5-17% significant connections. Second, between pathways, ventral to dorsal connectivity was more reliable and less sparse compared to dorsal to ventral connectivity (98 possible connections each way, reliability: CoD = 0.102 ± 0.020 vs. 0.0543 ± 0.012, sparsity: 23% vs. 2%). These findings show that connectivity is sparse and not always reciprocal. Moreover, they suggest that visual feedforward influences are stronger compared to feedback influences. Finally, the ventral pathway may more strongly modulate the dorsal pathway compared to the reverse.

This work was supported by NIH R01MH122258 and R01EY035533.

43.442 NETWORK ARCHITECTURE OF OBJECT RECOGNITION: INVESTIGATING INTEGRATION BETWEEN DORSAL AND VENTRAL VISUAL PATHWAYS

Claire Simmons[1] (csimmon2@andrew.cmu.edu), Vlad Ayzenberg[2], Marlene Behrmann[3]; [1]Carnegie Mellon University, [2]University of Pennsylvania, [3]University of Pittsburgh

While mounting evidence indicates that both the dorsal and ventral visual pathways participate in object perception, their relationship remains controversial. Previous work has identified specific instances of pathway interaction, but whether these reflect a single integrated network or distinct systems with context-dependent coupling remains unclear. We investigated this question using multiple network analyses of fMRI data acquired in response to viewing objects. Eighteen participants viewed 96 object images while undergoing 3T fMRI scanning (3 runs, TR=2s). Images were presented in blocks of 15 (800ms display) with interleaved fixations. Functional connectivity (FC), psychophysiological interactions (PPI), and Granger causality analysis (GCA) were computed between regions of interest in both pathways, defined using probabilistic parcels registered to native space. Results revealed both distinct processing and integrated function. FC analysis showed widespread connectivity between pathways, with stronger patterns in the left hemisphere. PPI analysis demonstrated task-specific connectivity changes during object processing, with lateral occipital (LO) regions showing stronger modulation than posterior intraparietal sulcus (pIPS). This directional influence contrasts with previous tool-specific findings. Direct pathway comparisons through subtraction analysis revealed complementary functional connectivity patterns but distinct task-modulated interactions. GCA uncovered asymmetric information flow, with stronger dorsal-to-ventral directed connectivity in response to objects processing (mean GCA value ~7.5 for left pIPS to LO) compared to scrambled displays (mean ~1.0). Our results suggest that object recognition relies on an integrated network with hierarchical organization, rather than purely independent pathways. While regions maintain distinct processing roles, they show coordinated activity through specific spatial and temporal patterns. This work further develops a network-level understanding of how dorsal and ventral pathways interact during object recognition, reconciling previous findings of both pathway independence and interaction within a broader framework of coordinated but specialized processing.

This research was supported by grants from the National Eye Institute, NIH (RO1 EY027018) to M.B and from NIGMS T32GM142630 to CS. MB acknowledges support from P30 CORE award EY08098 from the National Eye Institute, NIH, and unrestricted supporting funds from The Research to Prevent Blindness Inc, NY, and the Eye & Ear Foundation of Pittsburgh.

43.443 PERCEPTUAL DEFICITS IN PSYCHOSIS: THE ROLE OF HIGHER LEVEL VISUAL AREAS AND PRIOR KNOWLEDGE

Anh Pham[1] (pham0598@umn.edu), Scott R. Sponheim[2,1], Cheryl Olman[1]; [1]University of Minnesota, [2]Minneapolis VA Health Care System

Psychosis leads to the disruption of visual perception, which can be explained by the hierarchical predictive coding model as an impairment of the feedforward and feedback interactions between sensory input and prior knowledge. However, it is essential to have direct evidence from behavior and brain activity to support the model. Our study used an object recognition task with stimuli designed to control low-level features while varying recognizability. Using 7T fMRI, data were collected with 37 controls (CON), 15 bipolar disorder individuals (BP) and 5 relatives (BREL), and 19 people with schizophrenia (SCZ) and 10 relatives (SREL). Stimuli consisted of rough sketches of the outlines and textures of real-world objects formed by controlling the orientation of line segments on a grid. Three image conditions were high recognizability (meaningful - MF), low recognizability (meaningless - ML), and a baseline with randomized line segments. Regions of interest (ROIs) were V1, defined as the intersection of visually responsive voxels clustered in the calcarine sulcus, and post-hoc ROIs in fusiform, parietal, ventral temporal, lateral occipital (LOC), and prefrontal cortex, defined as clusters of voxels showing significant differences in responses to MF and ML stimuli. We found no significant group differences in BOLD responses to MF vs. ML objects in LOC; contrasts between ML objects and random were also generally consistent across groups, suggesting no significant differences in basic grouping processes. In contrast, the intraparietal sulcus and fusiform cortex showed greater BOLD enhancement to MF objects in SREL compared to ML than either SCZ or CON, and BP also showed larger response enhancements by object recognizability. This finding points toward higher visual areas, and differential use of prior knowledge, rather than low- and intermediate-level visual processing (edge detection and grouping) as a source of perceptual differences in psychosis patients and individuals with genetic liability for psychosis.

NIH R01 MH112583

43.445 THE CODING OF SPIKY OBJECTS IN HUMAN OCCIPITOTEMPORAL AND POSTERIOR PARIETAL CORTICES

Yaoda Xu[1], Marvin Chun[1]; [1]Yale University

An important aspect of understanding primate vision is to determine the key features in visual object coding. In the macaque inferotemporal (IT) cortex, spikiness and animacy are two principal features mediating the coding of a diverse array of objects, with distinctive IT subregions preferring unique combinations of these features (Bao et al., 2020). Similar results, however, were not observed in the human brain. Here, we constructed well-matched spiky and stubby stimuli and reexamined their coding in the human brain with fMRI. Using the output of a convolutional neural network shown to mirror the macaque IT cortex in spiky and stubby object representations, we selected pairs of inanimate spiky and stubby objects matched for semantic category (e.g., tripod vs. camera). The object images were further equated in low-level visual features, including luminance, contrast, and spatial frequency. Across 12 participants, contrasting spiky with stubby objects consistently revealed three areas preferring spiky objects: a small bilateral ventral activation between face- and scene-selective areas (proximal to the location of the macaque's inanimate-spiky area), a large bilateral lateral activation within a separately localized body-selective area, and a bilateral dorsal activation along superior/anterior intra-parietal sulcus. Meanwhile, no areas were found to prefer stubby objects. The same three spiky areas were also activated when participants viewed matching animate spiky-stubby stimulus pairs (e.g., crane vs. penguin), indicating their spiky-object preference is animacy-independent. Furthermore, within the lateral body-selective area, only the posterior part truly preferred animacy/bodies, as it was activated by both the animate spiky and stubby objects compared to the inanimate ones; the anterior part of the lateral body-selective area prefers spikiness but not animacy/bodies. Together, these results document for the first time a network of human brain areas preferring spiky over stubby objects and highlight the potential importance of spikiness in human visual object perception.

This research was supported by NIH grant R01EY030854 to YX.

Object Recognition: Categories

MONDAY, MAY 19, 8:30 AM – 12:30 PM, PAVILION

43.446 DISSECTING SPARSE CIRCUITS TO HIGH-LEVEL VISUAL CATEGORIES IN DEEP NEURAL NETWORKS

Jeffery Andrade[1] (jandrade01@g.harvard.edu), George Alvarez[1], Talia Konkle[1]; [1]Harvard University

While humans easily recognize innumerable object categories, the underlying computational paths from retina to category-level representations are still being unraveled. Convolutional neural networks (CNNs) like AlexNet have remarkable competence in visual categorization, and thus offer a unique case study for understanding the hierarchical routing of visual information. Extending work from Hamblin et al., 2023, here we develop a method to extract the relevant connections involved in the computation of each output category, and assess the effectiveness of this sparser sub-network. The key idea is that not all connections are necessarily involved in the computation of any one category; thus, for each of the 1000 category-level output units in the Alexnet, our algorithm assigns scores to connections based on their contribution to the category unit's outputs and prunes the lowest-scored connections to a specified sparsity. Our goal is to identify the sparsest circuit through the network that still maintains the original function. To evaluate how well the extracted circuits reflect the output unit's original functionality, we introduce a new metric–circuit substitution accuracy (CSA). We find that circuits need only 5.0% (median) of connections to achieve 85% of the unpruned CSA. Surprisingly, we observed that CSA initially increases with pruning and

often actually exceeds the unpruned baseline at its peak (median peak CSA = 188.0% median unpruned CSA) with just 13.3% (median) of connections. We hypothesize that the full network must employ inhibition to negotiate between competing, interfering pathways. Finally, the "anatomical overlap" amongst these category circuits ranged from <1% to >99% shared circuitry, revealing a range of implicit modularization in the network's categorical processing routes. Broadly, this work presents a novel method for gaining insight into the functional neuroanatomy of neural networks, and offers a foundation for understanding the hierarchical computations involved in the emergence of category-level information in visual systems.

Kempner Institute for the Study of Natural and Artificial Intelligence

43.447 MACAQUE SPATIOTEMPORAL NEURAL DYNAMICS DURING PERCEPTION OF OBJECT-OBJECT OCCLUSION IMAGES

Wenxuan Guo[1] (wg2361@columbia.edu), Matthew Ainsworth[2], Tim Kietzmann[3], Marieke Mur[4], Nikolaus Kriegeskorte[1]; [1]Columbia University, [2]University of Oxford, [3]University of Osnabrück, [4]Western University

The dynamic neural mechanisms for recognizing partially occluded objects are not fully understood. Previous studies often used partial fragments (Tang et al., 2014) or geometric shapes (Bushnell et al., 2011; Namima & Pasupathy, 2021). We investigated neural dynamics under ecologically valid conditions where objects occlude each other as in natural scenes. We recorded neural activity using Utah arrays implanted in foveal V4 and posterior TE of two macaque monkeys. Fixating monkeys viewed eight single objects and 56 object-object occlusion stimuli (250 ms duration) — resulting from all pairings of the objects. Instantaneous firing rates were estimated using a 50 ms sliding window spike count. Using cross-temporal decoding (Meyers et al., 2008), we investigated the temporal dynamics of neural population coding for occluder and occluded objects using linear SVMs. Decodability of a front object versus other front objects was computed by averaging decoding accuracy across pairs sharing the same back object. Similarly, we assessed discriminability for each occluded object versus others. Results showed that representations of front objects emerge earlier, are more decodable and stable over time than the occluded objects. Additionally, representations in posterior TE lag behind those in V4 but exhibit a more stably decodable temporal code. To assess whether spatial and temporal representations are separable, we applied tensor component analysis (TCA; Williams et al., 2018). We modeled the neural data tensor by decomposing it into a spatial mode (representing object-specific neural patterns) and two temporal modes (capturing dynamics for front and back objects separately). We computed cross-validated variance explained (R^2) compared to a baseline model. TCA models explained significantly more variance in V4 than in posterior TE, indicating that spatial and temporal representations are more separable in V4. TEp, intriguingly, exhibited less space-time separable dynamics, but more sustained decodability of front and back objects.

43.448 COMPETITION OR COOPERATION IN WORD AND FACE PROCESSING? A CROSS-SECTIONAL STUDY

Randi Starrfelt[1] (randi.starrfelt@psy.ku.dk), Anders Petersen[1], Peter Laursen[2], Erling Nørkær[1]; [1]University of Copenhagen, Denmark, [2]Bispebjerg Hospital, Copenhagen, Denmark

Visual processing of faces and words rely on a number of shared low-level processes, but the degree to which higher order visual and memory processes are shared or independent between these stimulus categories is debated. There are suggestions that the two interact and even compete during development, so that learning to read impacts the (cerebral) organization of the face recognition system. Whether there is a systematic relationship between face and word processing on the behavioral level in literate adults, and if so, in what direction, remains unclear. Here, we take advantage of unique data from a representative sample of Danish adults aged 41-71 (n=694) tested with the Cognitive Function Scanner test battery. Data on immediate and delayed (1h) face and word recognition, as well as a test of general visual ability were collected as part of a larger study of cognitive abilities. In pre-registered analyses, we test the relationship between face and word processing and its direction, and whether this potential relationship may be explained by general visual ability. In addition, we explore whether the relationship between word and face processing changes with recognition delay. The relationship between d' for words and faces was significant at both immediate and delayed recognition, in a positive direction. Adding scores from the general visual ability test to the regression did not abolish this relationship. Performance for both faces and words decrease with recognition delay, but more so for words than faces. The results suggest that word and face processing do rely on some shared higher order processes, and contradict the hypothesis of a negative relationship (competition) between performance with faces and words. Interestingly, the positive relationship between faces and words does not seem to be explained by a common underlying visual factor

43.449 EXTENSIVE EXPERIENCE REMODELS NEURAL TASK CIRCUITRY TO INCREASE AUTOMATICITY OF CATEGORIZATION

Patrick H. Cox[1], Clara A. Scholl[2], Marissa Laws[3], Xiong Jiang[4], Maximilian Riesenhuber[4]; [1]Lehigh University, [2]Applied Physics Laboratory, Johns Hopkins, [3]Carnegie Mellon University, [4]Georgetown University

Object category learning is an extensively studied foundational cognitive process. Typical human category learning paradigms are brief, lasting only a few hours, and generally show a sharpening of shape tuning in visual areas and the formation of task-dependent category responses in prefrontal areas. Other studies have identified a "frontal bottleneck" in the prefrontal cortex that limits processing to one task at a time, potentially limiting our ability to categorize objects in multitasking scenarios. However, many categorizations made in daily life are practiced over months and even years. This extensive practice may be associated with qualitative changes in behavior and underlying neural processing that make categorization more automatic. Here we tested the hypothesis that extensive training leads to a spatiotemporal shift in the neural circuitry underlying categorization. We trained participants on 30,000+ trials over 5-10 weeks to categorize novel morphed car stimuli using a mobile app. We used fMRI and EEG rapid adaptation (RA) techniques to identify the brain regions and temporal dynamics that underlie the categorization process after initial learning (~4 hours over 1 week) and after extensive practice (~16 more hours over 4 more weeks, ~30,000 trials in total). Converging evidence from these EEG and fMRI analyses showed that extensive experience fundamentally remodeled neural task circuitry: ventral occipitotemporal regions in the visual cortex that initially showed shape selectivity changed their profile to become category selective. These brain areas showed decreased functional connectivity with prefrontal cortex and increased connectivity with motor areas after extensive training, supporting the hypothesis that extensive experience leads to a bypassing of the “frontal bottleneck”. Crucially, the decreased connectivity with the prefrontal cortex correlated with increased dual-tasking ability, supporting the hypothesis that extensive practice leads to increased automaticity.

NSF (BCS-1232530)

43.450 EXTREME VALUE THEORY FOR MODELING CATEGORY DECISION BOUNDARIES IN VISUAL RECOGNITION

Jin Huang[1] (jhuang24@nd.edu), Deeksha Arun[1], Terrance Boult[2], Walter Scheirer[1]; [1]Department of Computer Science and Engineering, University of Notre Dame, [2]Department of Computer Science, University of Colorado, Colorado Springs

Several possibilities exist for modeling decision boundaries in category learning, with varying degrees of human fidelity. This paper finds evidence for preferentially focusing representational resources on the extremes of the distribution of visual inputs in a generative model as an alternative to the central tendency models that are commonly used for prototypes and exemplars. The notion of treating extrema near a decision boundary as features in visual recognition is not new, but a comprehensive statistical framework of recognition based on extrema has yet to emerge for category learning. Here we suggest that the statistical Extreme Value Theory provides such a framework. In Experiment 1, vertical line stimuli that vary in a single dimension of length are used to assess how human subjects and statistical models assign category membership to a gap region between two categories previously shown as training data. A Weibull fit better predicts an observed human shift when moving from uniform to enriched or long tails during training. In Experiment 2, more complex 2D rendered face sequences drawn from morph spaces are used as stimuli. Again, the Weibull fit better predicts an observed human shift when training data are sampled differently. An extrema-based model lends new insight into how discriminative information is encoded in the brain with implications for decision making in machine learning.

43.451 FOOD-SPECIFIC AREAS IN PRIMATE INFEROTEMPORAL CORTEX

Pinglei Bao[1], Baoqi Gong[1]; [1]Peking University

Visual information about food is crucial for survival in primates, yet the neural mechanisms for identifying and processing food stimuli remain unclear. Previous studies in humans have localized regions in the ventral temporal cortex that show a strong preference for food stimuli using fMRI datasets (Allen et al., 2021). However, the response of individual neurons within these regions has not been explored due to the lack of suitable animal models. To address this, we conducted a food localizer experiment with four macaques, comparing photorealistic and cutoff food images against corresponding non-food images. This analysis revealed a consistent food network across subjects, identifying three distinct food-selective regions spanning from posterior to anterior in the inferotemporal cortex. We then targeted the middle food region using Neuropixels probes for recordings. A high concentration of the food-preferred neurons was identified with two subtypes of food-preferred neurons: one responding more strongly to photorealistic food images and the other showing selectivity for cutoff food images. These neuron types also differed in cortical depth, with cutoff food-preferred neurons in superficial layers and photorealistic food-preferred neurons in deeper layers. To further investigate whether macaque food regions are homologous to human food areas, we calculated the correlation between mean neuronal responses and human fMRI voxel responses to the same set of images. Regions with high correlation significantly overlapped with human food areas. Overall, our results indicate that the macaque food area closely mirrors the human food network, suggesting it as a valuable model for investigating the neural mechanisms underlying food-related visual processing.

43.452 INVERSION EFFECT IN RECOGNIZING OBJECTS FROM DYNAMIC CUES

Shifa Maqsood[1] (smaqsood@udel.edu), Niloufar Faridani[1], Maryam Vaziri-Pashkam[1]; [1]Movement and Visual Perception Lab, Department of Psychological & Brain Sciences, University of Delaware

Humans have the ability to identify and categorise objects in a dynamic visual scene with remarkable accuracy. To achieve this accuracy, our visual system does not only rely on static information of the objects but it can also use dynamic cues. . In static images of objects it has been shown that turning the image upside down hinders recognition. Similar effects have been found in dynamic videos of human action. However, this inversion effect has not been investigated for a broader set of categories when only dynamic cues are available. Here we used object kinematograms to investigate the inversion effect on object recognition using dynamic cues. We used videos of 3 animate and 3 inanimate categories of objects, extracted motion cues from the videos and projected them on randomly moving dots to create object kinematograms that only contain dynamic information. We showed the videos either upright or inverted to participants in an online experiment. For each video, participants were asked to answer three questions: 1) identify if the stimulus was an animal or an object, 2) identify the category to which the object belonged and 3) name the object. Our results showed that the accuracies were higher for animals than objects ($p < 0.0001$). We observed a significant inversion effect for both animate and inanimate objects ($p < 0.001$) with no difference between the two ($p = 0.1336$). In addition, the classification accuracy for identifying the 6 objects categories was higher in the upright than the inverted condition ($p <0.001$) and the ability to name the object was significantly hindered when the videos were inverted ($p <0.001$). These results document for the first time strong inversion effects and suggest the role of configural processing in object categorization based on dynamic cues.

43.453 PERCEPTUAL LEARNING IN DERMATOLOGY: THE IMPACT OF ADAPTIVE COMPARISONS IN ACCELERATING SKIN LESION DIFFERENTIATION

Victoria Jacoby[1], Christine Massey[1], Philip Kellman[1]; [1]University of California, Los Angeles

Processes of perceptual learning can be accelerated through direct comparison of items. Previous work showed that adaptive learning methods can enhance the benefits of comparison by creating targeted contrasts between confused categories (Jacoby, Massey, & Kellman, 2024). Here, we tested whether adaptive comparisons could improve learning in the difficult task of skin lesion differentiation. Undergraduate participants were assigned to learn 10 categories of skin lesions in one of three conditions. In the Single-Item Condition, participants classified individual exemplars sequentially using a standard category learning format. The ARTS system (Mettler, Massey & Kellman, 2016), an adaptive interleaving system, determined category spacing based on learner performance. In the Adaptively Triggered Comparisons (ATC) Condition, participants completed the single-item trials with their responses monitored for confusions. Repeated misclassifications between categories triggered a comparison trial, in which two exemplars from the confused categories were presented simultaneously for discrimination. Finally, the Non-Adaptive Comparison (NAC) Condition included a similar number of comparison trials as the ATC Condition, but these occurred at fixed intervals with randomly selected categories. Participants completed trials until they met mastery criteria based on accuracy and response times, followed by immediate and one-week delayed posttests. Results: Average time to mastery was lowest in the ATC condition, followed by the Single-Item and NAC conditions. Posttest accuracy and learning efficiency (calculated as posttest accuracy divided by time invested) was numerically highest in the ATC condition, followed by the Single-Item and NAC conditions. An ANOVA revealed significant differences for time to mastery and posttest accuracy, with marginal effects for efficiency. Follow-up contrasts indicated that the ATC Condition significantly outperformed the NAC Condition across all measures and reached mastery reliably faster than the Single-Item Condition. These results provide initial evidence that adaptive comparisons can enhance the perceptual learning of skin lesion classifications, demonstrated primarily through quicker learning rates.

We gratefully acknowledge support for this research from National Institutes of Health grant R01CA236791 to PK and CM.

43.454 VISION AND SEMANTICS: INSIGHTS INTO ROCK CATEGORY LEARNING AMONG GEOLOGY UNDERGRADUATES

Anna K. Lawrance[1] (annalawrance@uvic.ca), Mateusz Janiszewski[2], Hilda Deborah[3], Andy J. Fraass[1], Duncan Johannessen[1], Lucinda J. Leonard[1], Dipendra J. Mandal[3], Brett D. Roads[4], James W. Tanaka[1]; [1]University of Victoria, [2]Aalto University, [3]Norwegian University of Science and Technology, [4]University College London

To the untrained eye, rocks offer limited perceptual information to aid in accurate categorizations—making geology an ideal domain to study the development of semantic and visuoperceptual knowledge. This study examined the formation and restructuring of perceived rock-type categories in undergraduates (N=48) enrolled in an introductory-level geology course. Through this work, we addressed three questions: 1) How do categories for rock knowledge develop? 2) How does the acquisition of real-world expertise reshape semantic and visuoperceptual categories? 3) Is the nature of one's category restructuring indicative of academic performance? We quantified category restructuring trajectories using PsiZ, a machine learning package that generates a multi-dimensional category representation (i.e., psychological embedding) based on the participant's similarity judgments (Roads & Love, 2020). On each trial, participants were presented with a visual array of nine rocks—depicted by images on visuoperceptual trials and labels (e.g., "basalt") on semantic trials—and were asked to select the two most similar peripherally presented rocks to the reference rock. Visuoperceptual and semantic category structures were assessed at the start and end of the course. How does category structure relate to academic success? A comparison of the top and bottom 25% of students, based on lab-test performance, revealed diverging trajectories. Post-instruction, image and label judgments of high-performers were highly correlated demonstrating strong integration of visuoperceptual and semantic knowledge. Contrastingly, low-performers displayed faulty visuoperceptual and semantic knowledge as demonstrated by poorly differentiated and conceptually inaccurate clusters of rock images and labels. Strikingly, despite no differences in prior exposure to the field, the two groups showed distinct pre-instruction visuoperceptual structures. High-performers exhibited significantly greater exemplar differentiation; this differentiation (among high-performers) was then maintained, but reconfigured (to assemble conceptually-accurate rock-type clusters) post-instruction. These findings suggest that the groups approached the task, at both timepoints, with markedly different levels of perceptual sensitivity.

This research was funded by an NSERC Discovery Grant held by Dr. James W. Tanaka.

43.455 A BEDSIDE ASSESSMENT OF THE BRAIN VISUAL FUNCTION OF PATIENTS WITH DISORDERS OF CONSCIOUSNESS—BASED ON FAST PERIODIC VISUAL STIMULATION ODDBALL PARADIGM AND EEG

Zhiqing Deng[1] (zhiqingdeng@m.scnu.edu.cn), Ruixue Wang[1], Gexiu Wang[1], Yichong Zhang[1], Jie Gao[1], Can Wang[1], Chao Zheng[1], Fuying Zhu[1], Zina Li[1], Pengmin Qin[1,2], Qiuyou Xie[3,4], Juan Chen[1,2]; [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]Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education, Guangzhou, Guangdong Province, 510631, China, [3]Joint Center for Disorders of Consciousness, Department of Rehabilitation Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, 510220, China, [4]School of Rehabilitation Sciences, Southern Medical University

Disorders of consciousness (DoC), mainly caused by traumatic brain injury, are typically with impairment in sensory processing. Among sensory modalities, vision is one of the most critical senses in humans. Accurate evaluation of the remaining vision function along the visual hierarchy from low-level visual features to high-level object recognition in DoC patients is crucial for their waking and rehabilitation. It is also challenging due to their limited ability to understand instructions and respond. Here, we proposed a method to rapidly evaluate DoC patients' visual function along the visual hierarchy at the bedside without participants' responses. A series of images were presented at 6 Hz (base frequency), with oddball stimuli presented following every four base stimuli (oddball frequency, 1.2 Hz) according to the fast periodic visual stimulation (FPVS) oddball paradigm while recording electroencephalographic signals. We observed electrophysiological response at 6 Hz in healthy controls, revealing their normal abilities to receive visual information. Oddball images generated an electrophysiological response at 1.2 Hz in healthy controls, revealing normal abilities to discriminate oddball stimuli from base stimuli. Unresponsive wakefulness syndrome/vegetative state (UWS/VS), minimally conscious state (MCS), and emergence from MCS (EMCS) retained the ability to receive visual information across almost all visual levels. However, UWS/VS primarily retained visual discrimination in low-level visual conditions, while MCS and EMCS demonstrated preserved discrimination in low-level, middle-level, and high-level visual conditions. Classification and prognosis results suggested that high-level visual functions were the strongest predictors of both current consciousness state and changes in consciousness state. These findings indicate that the FPVS oddball paradigm across various visual levels is a highly-sensitive, efficient, no-report tool for assessing visual processing in DoC patients, with high-level visual function indices predicting consciousness recovery. It advances our understanding of visual processing in DoC, offers insights for improving diagnostics, and guides patient awakening and rehabilitation strategies.

This work was supported by grants from the National Natural Science Foundation of China (No. 31970981 to JC, No. 82171174 and No. 82371184 to QX), the National Science and Technology Innovation 2030 Major Program (STI2030-Major Projects 2022ZD0204802 to JC), and the China Scholarship Council to JC.

Perceptual Organization: Parts, wholes, shapes and objects

MONDAY, MAY 19, 8:30 AM – 12:30 PM, PAVILION

43.456 EFFECTS OF ELEMENT FILL ON PERCEIVED SIMILARITY OF ENLARGED TEXTURES

Zoe Howard[1] (zoeshoward@gmail.com), Anna Chinni[1], Ruth Kimchi[2], Karen B. Schloss[1]; [1]University of Wisconsin-Madison, [2]University of Haifa

In information visualizations (e.g., charts, maps), visual features such as texture and color are used to represent concepts. Visualizations may be designed with particular sizes in mind (e.g., a printed page at arm's length), but they can be scaled down for smaller displays (phone screens) or scaled up for larger displays (projected slides). Thus, it is important to understand how interpretations of visualizations change with scale. For color, differences must be magnified to maintain discriminability for smaller visualization sizes (Stone et al., 2014; Szafir, 2017). Here, we studied effects of scaling on perceived texture similarity. Our approach was motivated by work comparing similarity between original patterns of local elements vs. enlarged versions that were either (a) proportional to the original pattern (i.e. dilation; "zoom-scaled") vs. unproportional (i.e., tiled; "repeat-scaled") (Goldmeier, 1936/1972; Kimchi & Palmer, 1982). Original patterns of few local elements were more similar to zoom-scaled stimuli, whereas original patterns of many local elements were more similar to repeat-scaled stimuli. Similar reversals occurred for changes in element size, reflecting a shift from part-based to texture-based representations (Kimchi & Palmer, 1982). Using Kimchi and Palmer's (1982) paradigm, we tested for analogous effects of scaling type for 36 textures commonly used in data visualizations, which varied in shape/fill/size/number (n=34; within-subject). Effects of number and size were similar to Kimchi and Palmer (1982), but we observed a new effect of element fill that interacted with size (p<.001). For filled elements, participants responded near chance for large elements and chose repeat-scaled patterns for small elements. But, for unfilled elements, they chose repeat-scaled patterns for patterns of large and small elements. These results suggest that texture fill modulates a shift of pattern perception from part-based to texture-based representations. These findings could have important implications for interpretations of textures with various scales in data visualizations.

NSF award BCS-2419493 to K.B.S; Hilldale Undergraduate/Faculty Research Fellowship University of Wisconsin, Madison

43.457 LEARNING SAME-DIFFERENT RELATIONS OF VISUAL PROPERTIES BY HUMANS AND DEEP CONVOLUTIONAL NEURAL NETWORKS

Philip Kellman[1], Nicholas Baker[2], Austin Phillips[1], Patrick Garrigan[3], Hongjing Lu[1]; [1]University of California, Los Angeles, [2]Loyola University of Chicago, [3]St. Joseph's University

Considerable research suggests that deep convolutional neural networks (DCNNs) trained for image classification do not access abstract relational properties needed for perception of global shape (Baker, Lu, Erlikhman, and Kellman, 2018) or detection of relations, including same-different relations (Puebla & Bowers, 2022; Baker, Garrigan, Phillips & Kellman, 2023). Due to its inherently comparative architecture, we hypothesized that a twin network, built on AlexNet sub-networks and trained on same-different comparison tasks, might show better capability for learning and generalization of same-different relations. We tested the network (pre-trained on ImageNet classification) and human participants on same-different learning for pairs of objects that could match or differ in color, texture, or shape. Each condition contained same-different trials, where same trials could match on either of two dimensions, producing three conditions: color-texture (shape irrelevant), color-shape (texture irrelevant), and shape-texture (color irrelevant). Human participants were trained on 256 trials. Human asymptotic learning performance in the shape-color and shape-texture conditions was reliably greater than in the color-texture condition. In contrast, the model yielded opposite learning results, showing better learning performance in the color-texture (shape irrelevant) condition than in the shape-color and shape-texture conditions. This difference highlights the network's limited capability to compare shape information in same-different judgments. Generalization testing using new stimulus values on the trained dimensions in each condition showed poor performance by the model, with a maximum accuracy of .62 on color-texture test stimuli after color-texture training, and no other accuracies higher than .55. Conversely, human participants showed excellent generalization to new stimulus values (.8), slightly exceeding performance levels shown at the end of training. These results suggest two salient limitations of same-different learning in a twin network: poor performance where shape is the basis of comparison, and difficulty in generalizing the same-different relation, even to new stimulus values within a trained stimulus dimension.

We gratefully acknowledge support for this research from National Institutes of Health grant R01CA236791 to PK.

43.458 SPATIOTEMPORAL CORRELATION STRUCTURE IN ORIENTATION REPULSION

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

Visual appearance is modulated by spatial context. Orientation repulsion is one such contextual phenomenon wherein a central target appears tilted against the orientation of a surrounding inducer. We employed a psychophysical reverse correlation technique to investigate the temporal window contributing to repulsion at several target–inducer distances. If repulsion results from passive filtering, the window would shift backward with increasing distance due to longer latencies of contextual signals. At 8° below the fixation point, an oriented grating, serving as the inducer, was presented within an annular window and randomly switched the orientation every 33 ms. In the middle of this time sequence, a horizontal target Gabor patch was flashed for 33 ms at the center of the inducer. After they disappeared, two probe Gabor patches, tilted 5° and −5° from the horizontal, were presented side by side, and participants judged which probe appeared closer in orientation to the target. The proportion of judging the target as tilted against the inducer was plotted against the relative time between the target and inducer and then fitted with a Gaussian function. This procedure was separately applied to three blocked conditions for target–inducer distance, defined as the radius of the inducer annulus (2°, 3.5°, and 5°). The temporal window was significantly positive for an inducer orientation of ±20°, consistent with the optimal orientation relationship for repulsion. Contrary to the above prediction, the window significantly shifted forward as distance increased. This shift is considered to reflect the time lag of contextual signals sent from the target to inducer, rather than from the inducer to target. Therefore, contextual modulation in orientation repulsion may be viewed as an active process: after a behaviorally relevant target is detected, query signals radiate from the target representation toward its surrounds to acquire information about the spatial context.

Supported by JSPS KAKENHI 21J20400, 22KJ0555, 24KJ0233, 18H05523, 23H01052, and 23K25749.

43.459 ROLE OF PERCEIVED OBJECT STRUCTURE ON A SPATIAL PERCEPTION ILLUSION

Fernando Castillo Rodriguez[1], Cathleen M. Moore[1]; [1]University of Iowa

Perceiving distance between things is essential for successful action. However, our perceptions are sometimes inaccurate. An example is an illusion originally labeled “object-based warping” because it suggested that our representation of space is distorted by perceived object structure (Lebed et al., 2023; Vickery & Chun, 2010). This illusion occurs when the separation between two stimuli appears to be larger or smaller than it actually is, depending on whether they appear inside or outside of an object, respectively. Because these misperceptions do not occur when there is no object present, they were interpreted as perceived object structure distorting perceived space. Alternatively, these misperceptions may be driven by interactions among image features including proximity of stimuli to contrast edges (Baker et al., 2024). We tested whether perceived object structure contributes to these spatial misperceptions or if instead they can be accounted for entirely by image-level interactions. We used binocular disparity to manipulate perceived object structure while maintaining nearly identical image-level information across four conditions. Two rectangles were presented at different depths defined only by binocular disparity. Two vertically separated diamonds were presented at one of the two depths thus appearing in the same depth plane as one of the two rectangles. The perceived separation between the diamonds differed depending on the rectangle with which the dots appeared. Because image level information was nearly identical across conditions, the perceived differences in separation can be attributed to perceived object structure. How, specifically, perceived object structure impacts perceived distance remains an open question.

43.460 TOPOLOGICAL STRUCTURE AND THE CREATION OF VISUAL COMPLEXITY

Ashna Shah[1], Sami Yousif[1]; [1]University of North Carolina at Chapel Hill

Imagine different Tetris pieces. They are all made up of four blocks, yet some appear distinctly more complex than others. Why is that? One compelling idea is that the different pieces vary in the topology of their underlying skeletal structure; we know that differences in skeletal structure predict ratings of complexity and visual engagement (Sun & Firestone, 2021). But what if participants are not asked to judge the complexity of objects, but generate objects of varying complexity? Here, we had participants complete a block-building (or block-destroying!) task in which they were asked to either modify or maintain the complexity of Tetris-esque objects by altering the number of blocks in the object. We measured changes in the number of specific topological features in the skeletal structure of the objects (e.g., T-junctions, crosses, and holes). There were strikingly different patterns of block placement depending on the instructions participants were given. When asked to increase complexity, participants added more T-junctions and crosses in the objects but added fewer L-junctions and holes (compared to the maintain-complexity condition). In two additional experiments, we did the opposite: We asked participants to remove blocks to either decrease or maintain the complexity. We found the same pattern: When asked to decrease complexity, participants retained more L-junctions and holes in the objects but retained fewer T-junctions and crosses (compared to the maintain-complexity condition). The instructions in all of these tasks were minimal: Participants were given virtually no guidelines about how to alter the blocks except to increase/decrease/maintain the complexity of the objects. Yet the objects they created as a result differed dramatically in their overall structure (a fact which is immediately apparent when looking at the resulting objects). Thus, these results demonstrate that topological structure influences not only how people perceive complexity, but also how they create complexity.

43.461 SYMMETRY DETECTION AND FIGURE-GROUND SEGMENTATION IN PERCEPTUAL ORGANIZATION OF RICH IMAGES

Lisa Koßmann[1] (lisa.kossmann@kuleuven.be), Gonzalo Muradás Odriozola[1,2], Christophe Bossens[1], Johan Wagemans[1]; [1]KU Leuven, Department of Brain & Cognition, [2]KU Leuven, Department of Electrical Engineering (ESAT)

Symmetry plays a critical role in perceptual organization, contributing to tasks like perceptual grouping, figure-ground segmentation, shape perception and aesthetics. However, most studies rely on simple geometric patterns, leaving its role in complex, real-world stimuli underexplored. Here, we present findings from a preregistered online study involving 901 participants and a set of 200 real-world photographs and 200 artworks to investigate symmetry detection and its interplay with figure-ground segmentation. Participants identified symmetric regions in images by superimposing bounding boxes and axes of symmetry on self-selected image regions, allowing multiple regions per image. For each region, they rated salience (how much it stood out to them) and strength of symmetry (ranging from imperfect to nearly perfect). These human annotations were compared to symmetry detection by a deep learning model (U-Net). Additionally, object detection (YOLO11) and segmentation (SAM2) algorithms were used to explore how figure-ground segmentation influenced symmetry perception. Our results confirm the dominance of vertical symmetry axes in real-world scenes (72.91%) and artworks (73.87%). Bayesian regression analysis revealed that as the number of marked regions increased, their salience and size decreased while symmetry strength slightly increased. The U-Net model showed alignment with human annotations in high-salience, high-symmetry regions, with overlaps of 55% and 53% for natural scenes and artworks, respectively. Overall performance was better for natural scenes (Weighted F1-Score = 0.43) than for artworks (WF1 = 0.40), but it is clear that machine vision models still have a way to go before they can mimic human symmetry perception. Critically, symmetry detection was heavily influenced by object presence and recognizability within the image, suggesting a strong interaction between symmetry perception, figure-ground segmentation, and object detection. These findings bridge the gap between traditional symmetry research and real-world visual contexts, offering insights into the computational and perceptual mechanisms underlying symmetry detection and in complex images.

This work is funded by an ERC Advanced Grant (No. 101053925, GRAPPA) awarded to JW.

43.463 LINEAR EXTRAPOLATION DOES NOT EXPLAIN THE MISLOCALIZATION OF A TRIANGLE'S OCCLUDED VERTEX

Tess White[1], Jacob Tack[1], Gideon Caplovitz[1]; [1]University of Nevada, Reno

In previous year's VSS conferences, we presented a variety of stimulus configurations in which the perceived intersection point of two converging lines is systematically biased away from the true intersection point. We have found a consistent pattern of result for configurations that include partially occluded triangles, partial triangles without explicit occluders and even pairs of converging lines. The bias persists under both binocular and monocular viewing and is modulated by both the length of the visible lines and the convergence angle, with greater biases observed for more acute angles. In triangle configurations, observers underestimate the inferred vertex (closer to the base than it really is). While this contraction is somewhat consistent with interpolation models of contour continuation in amodal completion, such models are based on smooth-curvilinear interpolation and do not readily account for completing a singular vertex or intersection point. Rather than interpolation, our current research seeks to determine whether the contraction is a result of an extrapolation mechanism. Across two experiments our observers were presented with linear line segments matched in length and orientation to conditions used in our previous experiments and performed a linear extrapolation task in which they place a dot on the other end of the screen where they thought the line segment would extrapolate to. In both Experiments, we found systematic biases in how observers extrapolate with a pattern of results that conflict with our previous findings. For example, acute line angles relative to vertical tend to be pulled towards the vertical, a finding that is consistent with expansion not contraction. Unlike the results of our previous studies, two-line extrapolation results in the strongest effects in obtuse angle conditions (120°). Based on these data, we conclude that linear extrapolation does not explain the systematic bias of a triangle's vertex or two converging lines' intersection point.

43.464 LATERALIZATION OF CROWDING WITH STEREO-DEFINED 3D LETTERS: TWO PATHWAYS FOR 3D SHAPE PERCEPTION

Anthony Cate[1]; [1]Roanoke College

INTRODUCTION: Perception of the 2D outlines shapes of concave and convex 3D forms relies on qualitatively different visual features, which suggests that there may be two distinct pathways for perceiving 3D shape (Cate & Behrmann, 2010). This study investigated whether the basic shape features of concavities and convexities have different spatial extents, based on the idea that visual crowding is related to isolation field size (e.g. Pelli et al., 2004). METHODS: Participants with normal or corrected vision identified capital letters (stroke width 1.5˚) presented within random dot stereograms on a CRT monitor synchronized with LCD shutter glasses. The background portion of each stereogram filled the display, and had a uniform disparity of either 6' crossed, 6' uncrossed, or 0. 3D letters with either convex or concave depth structure were produced by creating regions with smoothly changing disparities that appeared either nearer or farther than the background. These relative disparity differences extended to 6'. On each trial, an array of three vertically arranged letters appeared briefly (150 ms) at 11˚ eccentricity, randomly to the left or right of fixation. Participants identified the middle letter. The two flanker letters were separated from this target either by a tight spacing that produced crowding (5.5˚), or by a loose spacing that did not (16.5˚), as established by a previous study (Cate & Hartman, VSS 2016). Visual crowding was as measured the difference in correct identifications for the loose minus tight spacing conditions. RESULTS: Crowding was significantly higher in the right hemifield for concave letters, and in the left hemifield for convex letters. CONCLUSION: This may suggest that concave and convex 3D forms recruit shape perception processes that differ according to the parts-based and holistic processing distinction attributed to the left and right cerebral hemispheres (e.g. Robertson & Lamb, 1991).

Spatial Vision: Crowding and eccentricity

Monday, May 19, 8:30 am – 12:30 pm, Pavilion

43.465 SPATIOTEMPORAL CONTEXT SHAPES CENTER-SURROUND DYNAMICS

Lisa Schwetlick[1,2], Peter Neri[2]; [1]EPFL, [2]ENS

The ability to discriminate visual features is profoundly influenced by the surrounding context. Understanding center-surround dynamics is essential for uncovering the mechanisms of contextual interactions that play a role in natural scene processing. Previous research has highlighted complex dependencies of orientation tuning on the spatial and temporal relationships between a central probe and its surround. Here, we extend this work by investigating how orientation noise and contextual configurations influence orientation discrimination in a 2AFC task. Participants identified which of two patches contained a target orientation while the surround elements varied in shape (full ring, four diagonal petals, or two opposing petals) and grating content (aligned with target orientation, orthogonal to target orientation, or plaid combining both). Both patches were independently corrupted by orientation noise to support reverse correlation analysis. In the absence of a surround, noise aligned with target orientation increased the probability of classifying the corresponding patch as containing the target, while anti-target noise orthogonal to the target reduced it. In the presence of a plaid surround, these noise effects remained similar, but were attenuated. When the surrounds contained a target-aligned grating, the effect of target-aligned noise was selectively reduced, while surrounds with orthogonal-to-target gratings reduced the impact of orthogonal-to-target noise. While surround effects were mainly driven by surround content (grating orientation), we also observed subtle effects of surround shape (petal alignment with the target). Moreover, these effects become evident only when the surround is presented at the same time as- or after the target, but not when the surround is shown before the target. These results reveal nuanced spatiotemporal contextual effects that dynamically modulate local orientation tuning, and provide critical constraints for developing models of contextual interactions in vision.

43.466 GESTALT MODULATION OF SURROUND SUPPRESSION IN MACAQUE V1

Cai-Xia Chen[1], Dan-Qing Jiang[1], Xing-Nan Zhao[1], Xin Wang[1], Shi-Ming Tang[1], Cong Yu[1]; [1]Peking University

Orientation discrimination of a line target is impaired when flanking lines are added. This crowding effect is at least partially attributable to neuronal surround suppression. The Michael Herzog lab also reported that when the target and flanking lines differ in length, the resultant separate Gestalts reduce crowding, suggesting that surround suppression is modulated by Gestalt organization. We investigated the neuronal responses associated with this de-crowding effect. Using two-photon calcium imaging, we simultaneously recorded responses of large populations of V1 and V4 neurons at various target-flanker distances, with the flanker set to the same or longer length, in awake, fixating macaques. The target was a 0.8° long line presented in the parafovea (1.2-3.4° eccentricity), while the flankers were a pair of parallel lines either the same length as the target or longer (4°). The results revealed: (1) Compared to V1 neuronal responses to the target line alone, surround suppression was significant with the addition of flanking lines. However, with longer flanking lines, less suppression was evident at certain target-flanker distances. (2) Using a SVM to classify the same and longer flanker lengths based on PCA-transformed population neuronal responses to the target line, a decoding accuracy of 70% could be reached across all target-flanker distances with up to 25 PCs. (3) Upon reducing the population neuronal responses to a low-dimensional space using multidimensional scaling (MDS), the distances between different orientations were greater under the longer flanker condition compared to the same-length flanker condition. (4) In contrast, V4 neuronal responses were further suppressed by longer flanking lines. We conclude that Gestalt modulation not only reduces surround suppression but also enhances the separability of target orientation at the population representation level. The opposing effects seen in V4 suggest that the Gestalt modulation effects in V1 may not be driven by downstream feedback mechanisms.

STI2030-Major Projects grant (2022ZD0204600)

43.467 **CHILDREN'S READING ABILITY IS BETTER** PREDICTED BY FOVEAL CROWDING THAN ACUITY

Sarah J Waugh[1] (s.j.waugh@hud.ac.uk), Emma Martindale, Monika A Formankiewicz, Leticia Álvaro, Denis G Pelli; [1]University of Huddersfield, [2]Anglia Ruskin University, [3]Universidad Complutense de Madrid, [4]New York University

Visual acuity and crowding distance develop differently but which measure relates better to reading ability and/or reading speed? Visual acuity with isolated Sloan letters and crowding distance with Pelli trigrams and repeated optotypes (Pelli, Waugh, Martelli et al., 2016) were measured. In Group 1 (N=200 children aged 3-11 years), teacher-assessed literacy/reading indicators were also obtained. In Group 2 (N=72 children aged 4-10 years), reading speed was measured with a Rapid Serial Visual Presentation (RSVP) and an "Ordinary" reading task. Teacher-assessed PiRA (Progress in Reading Assessment) reading ages were also obtained. In Group 1, log visual acuity (r=-0.35) and log crowding distance (r=-0.66 and -0.65) decreased and teacher-assessed reading performance (r=0.94) increased with age (Pearson r; p<0.0001). Removing age through partial correlation revealed a significant impact of crowding distance (trigram and repeated optotypes; r=-0.24 and -0.25, both p<0.001) on teacher-assessed literacy/reading performance, not true for visual acuity (r=-0.13, p>0.05). In Group 2, relationships between log visual acuity (r=-0.36), log crowding distance (r=-0.72 and -0.71) and teacher-assessed reading age (r=0.90) with actual age were similar to those in Group 1 (Pearson r; p<0.01). In readers, crowding distances were significantly correlated with PiRA reading age (r=-0.59, -0.51, p<0.0001), whereas visual acuity was not (r=-0.04, p>0.1). Reading speed in readers improved with age (RSVP r=0.62; Ordinary r=0.63; p<0.0001)) and teacher-assessed reading age (RSVP r=0.67; Ordinary r=0.72; p<0.0001). Crowding distance was significantly correlated with RSVP reading speed (r=-0.50, -0.45; p<0.001), whereas visual acuity was not (r=0.057; p>0.10). Similarly, crowding distance had a stronger relationship with ordinary reading speed (r=-

0.63, -0.52; p<0.0001) than did visual acuity (r=-0.29, p<0.05). Foveal crowding distance is much better than acuity at predicting children's reading ability. Thus clinically, measurement of crowding distance could prove valuable, especially in children whose reading might be impaired by vision.

URN020-01 from University of Huddersfield to Waugh. R01 EY027964-01A1 to Pelli.

43.468 CAPTURING APPEARANCE REVEALS ILLUSORY LETTERS IN VISUAL CROWDING

Bilge Sayim[1], Olivia Koechli[2], Natalia Melnik[3]; [1]CNRS, University of Lille, [2]ipw Winterthur, [3]Otto-von-Guericke-University Magdeburg

In typical experiments on visual crowding (the deteriorating influence of clutter on target perception), observers are usually informed about the stimulus category they have to report. For example, observers are asked to report a target letter. This prior information strongly limits the response space to a few categorical (letter) responses, and may influence how targets are perceived. Here, we investigated to what extent prior experience with letter stimuli increased the likelihood to subsequently report letters when letter-like stimuli were presented. Targets consisted of letters and letter-like stimuli (modified letters with shortened and added lines), created with lines positioned on a virtual 3-by-3 dot grid. Targets were presented either in isolation or flanked by Xs (crowded) at 10° eccentricity to the left or right of fixation. Observers' task was to capture target appearance by placing lines on a freely viewed response grid. There were two groups of observers: In the Letters First (LF) condition, observers were first presented with letters; in the Letters Second (LS) condition, observers were first presented with letter-like stimuli. We hypothesized that compared to the LS condition, the prior experience of letters in the LF condition would bias observers to report letters instead of the presented letter-like targets. We analyzed to what degree the appearance captures resembled the presented targets. The results showed strong deviations of the captured from the presented targets, especially when the targets were crowded. Quantifying how often observers reported the corresponding letter targets when presented with letter-like stimuli revealed that LF observers erroneously 'corrected' the letter-like stimuli to letters more frequently than LS observers. Overall, our results show that prior experience of letter stimuli strongly influenced appearance reports. We suggest that systematic biases in typical crowding paradigms can be revealed --and avoided-- by appearance-based methods.

43.469 CROWDING PREDICTS READING SPEED AND COMFORT ACROSS FONTS AND PARTICIPANTS

Maria Pombo[1], Minjung Kim[2], Denis G. Pelli[1]; [1]New York University, [2]Meta Platforms, Inc.

Crowding and reading speed vary more than two-fold across participants and also vary across fonts. Here, we try many fonts and examine crowding and reading speed. Seventy-four online participants performed three objective tasks — visual crowding, ordinary reading, RSVP reading — and two subjective rating tasks — comfort and beauty. Each participant did these for 3 of 12 fonts. Fonts ranged from common text fonts (e.g., Times New Roman) to display fonts (e.g., Zapfino, an intricate script font, and Omfug, a bubbly, graffiti-like font). In the crowding task, participants identified the middle of three letters presented at ±5º horizontal eccentricity. QUEST measured the threshold spacing. In the ordinary reading task, participants read from a short story and answered reading retention questions. For RSVP, on each trial, 3 words were presented, one at a time. Participants then identified the three words among foils. QUEST measured the threshold word duration. Participants also rated (on a 7-point scale) their comfort after a fixed-rate RSVP reading task. Lastly, to assess the beauty of the font, participants rated how much beauty they felt from looking at a page of Latin Lorem Ipsum text. This emphasizes looking without language processing. Results reveal strong correlations between crowding and both reading comfort (r = –0.93) and speed (r = –0.89 for RSVP, r = –0.58 for ordinary reading). A mixed-effects linear model with font as a fixed effect and participant as a random effect explains the variance in log crowding distance (89%), log reading speed (84%), and log comfort (43%). Text fonts generally produce low crowding and faster, more comfortable reading, while display fonts with high crowding and slower reading are better suited for titles and ads. In sum, measuring a person's crowding online predicts their reading performance and comfort across fonts.

43.470 THE NATURE OF TEMPORAL CROWDING - THE ROLE OF FORWARD AND BACKWARD INTERFERENCE.

Ilanit Hochmitz[1], Yaffa Yeshurun[1]; [1]University of Haifa

Temporal crowding refers to impaired target identification when it is preceded and/or succeeded by other irrelevant items. Critically, this temporal interference occurs with stimulus onset asynchronies (SOAs) that exceed the typical limits of visual masking (i.e., SOA>150ms), and was even found with an SOA of almost half a second. Recently, we have directly compared temporal crowding and visual masking, demonstrating that temporal crowding is not merely 'particularly long' masking, but rather the two phenomena rely on different perceptual processes. Furthermore, unlike suggested by some models for its spatial counterpart —spatial crowding—the interference brought about by temporal crowding cannot be accounted for by simple pooling (i.e., averaging) of information across time. In the current study, we examined whether temporal crowding primarily arises from forward interference or backward interference. Across three experiments, participants performed an orientation estimation task. A sequence of three randomly oriented stimuli was presented to the periphery. SOAs ranged from 175 to 475ms. Depending on the experiment, the target was either the first, second, or third item in the sequence. Participants had to reproduce the target orientation by rotating a probe line, and the measure of performance was the angular difference between the target's true orientation and the reported orientation. Mixture modeling analysis revealed that temporal crowding was strongest for the second item, showing significant effects of SOA on target encoding precision and the rate of reporting the orientation of a non-target item, but not on the guessing rate. A similar, yet weaker, pattern of effects was found for the first item, while minimal interference was observed when the target was the third item. These findings suggest that crowding reflects a combination of strong backward interference and weaker forward interference.

43.471 SOLVING A MAZE WITH TUNNEL VISION

Yelda Semizer[1] (yelda.semizer@njit.edu), Kian Motahari[1], Benjamin Balas[2], Ruth Rosenholtz[3]; [1]New Jersey Institute of Technology, [2]North Dakota State University, [3]NVIDIA Research

Solving a maze often requires a series of eye movements. Peripheral information, limited by visual crowding, determines ability to perceive the path ahead, and informs the next fixation location. We examine to what extent peripheral vision is used for maze solving using a gaze-contingent paradigm by only rendering a portion of the maze within a circular window centered at the current gaze point, outside of which the maze is masked or degraded. If the visual system uses peripheral vision to solve a maze, performance should diminish when there is no peripheral information available. For a given maze design, performance should increase as a function of window size, up to the point at which the window size reaches that of the functional viewing field. If peripheral information is available but degraded outside the window, performance should be better but constrained by the usefulness of the remaining information in signaling the structure of the path ahead. If the information available to peripheral vision is not informative about the path ahead, performance should be poor. Observers solved mazes mentally within an eye-tracking setup. Only the central gaze region was presented in clear view while the peripheral regions were masked to varying degrees of visual angle (1, 2, 4, or 7 in radius). Three different masking conditions were applied, where the peripheral region was either blurred, fully masked, or masked by a grid-like structure matching the appearance of the maze paths and walls. Results showed that performance in the grid condition was worse than in the masked condition, which was worse than in the blur condition. Additionally, performance increased with window size, with no difference between 4 and 7 degrees. These findings suggest that the amount of peripheral information available, its relevance, and usefulness for the task are critical in mental maze solving.

43.472 CROWDED DYNAMIC FIXATION FOR ONLINE PSYCHOPHYSICS

Fengping Hu[1] (fh986@nyu.edu), Joyce Yi Xin Chen[2], Denis G Pelli[1], Jonathan Winawer[1]; [1]New York University, [2]University of Michigan - Ann Arbor

Online vision testing allows efficient collection of data from a diverse participant pool. However, many vision science studies depend on participants looking at a fixation mark. Traditionally, vision scientists check fixation with a camera-based gaze tracking device. However, current online implementations of gaze tracking using webcams are not sufficiently accurate. A recent alternative strategy took advantage of the natural tendency to coordinate hand and eye movements, encouraging accurate fixation by asking participants to track a moving fixation mark with a mouse-controlled cursor (Kurzawski, Pombo, et al., 2023). This dynamic fixation task greatly reduces peeking at a peripheral target, but does not eliminate it. Here, to further improve fixation, we exploit the “crowding” phenomenon by adding clutter around the fixation mark. We call this “crowded dynamic fixation”. We assessed fixation accuracy while measuring a peripheral threshold. Compared to stationary fixation, dynamic fixation without clutter reduced RMS fixation error by 39%. Crowded dynamic fixation reduced RMS error even more, by 53%. With a 1.5º tolerance, this corresponds to “peeking” on 9% of trials with a stationary fixation, 4% with a dynamic fixation, and 0% with crowded dynamic fixation. This improvement in fixation eliminated implausibly low thresholds on the peripheral task, presumably by eliminating peeking. We conclude that, by preventing peeking, crowded dynamic fixation brings accurate gaze control to online testing.

Funding: NIH grants R01-EY027401 (JW), R01-EY027964 (DP) and R01-EY033628(JW).

43.473 LEFT/RIGHT ASYMMETRIES IN VISUAL CROWDING

Valorie Wiseman[1,3], Summer Sheremata[2,3]; [1]Neuroscience Graduate Program, [2]Department of Psychology, [3]Florida Atlantic University

When spotting signs in a busy area, searching for a friend in a crowd, or even reading this text, visual crowding plays a role in many of our daily tasks. Crowding is the phenomenon that occurs when the presence of nearby distractors makes it hard to identify an object. The role of attention in crowding is well-supported, but the specifics remain the subject of active debate. Asymmetries in the visual field have been identified, with greater crowding effects seen in the upper, as compared to the lower, visual field. Greater crowding has also been identified in the left, as compared to the right, visual field. However, these asymmetries have been found using letter stimuli. It is possible, however, that left/right visual field asymmetries reflect language processing or another process underlying visual crowding. Therefore, in order to rule out asymmetries in language processing, we asked participants to perform a crowding experiment in which contrast-defined gabors were presented to 5° to left and right of fixation. In the crowding block, flanking gabors were present at distances of 2°, 3°, or 4° center-to-center. In the isolation block, no distractors were present. Accuracy increased with target-distractor distance across the 3 conditions. Importantly, accuracy was lower (greater crowding) in the left, as compared to right, visual field. These asymmetries were not seen in the isolation block. We, therefore, found that observers are more accurate at identifying crowded targets and, therefore, have significantly decreased crowding effects in the right visual field when compared to the left visual field. Furthermore, these results cannot be explained by asymmetric language processing. This rightward bias in crowding establishes a greater understanding of the role of visual field asymmetries in visual crowding.

This work was supported by the June Jones fellowship to V.W.

43.474 CROWDING DOES NOT IMPROVE PRECISION

Ayberk Ozkirli[1], David Pascucci[1,2], Michael H. Herzog[1]; [1]Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, [2]Psychophysics and Neural Dynamics Lab, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

In crowding, target perception deteriorates in the presence of flankers, i.e., performance with flankers is worse than in a target only condition. However, a recent study by Cicchini et al. (2022) suggests that “crowding improves overall performance” compared to the target alone condition through optimal integration of target and flanker features, i.e., a ‘superiority’ effect. Here, we attempted to replicate their findings using the original paradigm. Participants reproduced the orientation of an oval target, which was presented either in isolation or flanked by ovals of varying orientations. In Cicchini et al., performance in the flanked conditions was superior compared to the target alone condition for almost all orientation differences between the target and flankers. With the very same paradigm, we could not reproduce their results:

the target alone condition resulted in best performance. When the flankers and target shared the same orientation or differed by 90°, performance was comparable to that in the target alone condition, with deterioration observed for intermediate orientation differences. Therefore, our data falsify the optimal integration model proposed by Cicchini et al. to explain crowding and its alleged benefits.

43.475 RETURN TO ISOTROPY: CROWDING ZONES BECOME LESS ANISOTROPIC IN FAR PERIPHERY

Evalie C. Rehor[1] (evalie_rehor@berkeley.edu), Daniel R. Coates[2], Susana T. L. Chung[1]; [1]University of California, Berkeley, [2]University of Houston

Crowding, the impaired ability to recognize an object in clutter, has been suggested as the bottleneck of object recognition, especially in peripheral vision. A signature of crowding is its radial-tangential anisotropy in peripheral vision —the extent of crowding is larger along the radial than the tangential direction. Nandy and Tjan (2012) proposed that this anisotropy is a result of the saccade-confounded image statistics in peripheral vision, which affects the direction along saccades (radial with respect to fixation) more than other directions. Because saccades made by humans are usually <15°, these authors predicted that the radial-tangential anisotropy would be smaller at eccentricities >15°. We tested this prediction by comparing the crowding extent (critical spacing) along the radial and tangential directions at 5°, 10°, 20° and 30° eccentricities in the right visual field. Participants (n=4) identified the orientation of a Tumbling-E presented alone, or flanked by two other Tumbling-Es positioned either radially or tangentially relative to the target-E, at a range of target-flanker spacings and letter sizes. The critical target-flanker spacing yielding 79%-correct was determined psychophysically, for each combination of eccentricity × radial/tangential conditions. Fixation was monitored using an eyetracker. As expected, the critical spacing increased with eccentricity. However, the increase was not isotropic: 5.5× radially and 7.8× tangentially (5° vs. 30° eccentricity). Moreover, the radial-tangential ratio of critical spacing was not a constant across eccentricity ($p=0.005$); instead, it increased slightly from 5° to 10° eccentricity, then diminished at larger eccentricities. Post-hoc analyses showed that the ratio at 30° eccentricity (1.50±0.15) was significantly different from the ratios at 5° (2.13±0.21) and 10° (2.31±0.31) eccentricities. Our finding of a reduction in the radial-tangential anisotropy of the crowding extent at large eccentricities is consistent with the prediction of Nandy and Tjan (2012) on the role of saccadic eye movements in modulating the crowding extent.

Grant Support: NIH Grant T32-EY007043

Color, Light and Materials: Optics, models

Monday, May 19, 8:30 am – 12:30 pm, Pavilion

43.476 THE EFFECT OF WAVELENGTH ON BEHAVIORAL RESPONSES TO SCATTERED LIGHT

Yaw Buabeng[1] (yaw.buabeng@uga.edu), Billy Hammond[1], Lisa Renzi-Hammond[1]; [1]University of Georgia

Intraocular light scatter significantly impairs visual performance and is a critical factor in limiting the ability to perform vision-dependent tasks, such as safe driving. Visual stimuli—ranging from lighting and road signs to lenses and ocular implants—can be optimized to minimize the detrimental effects of scatter. The success of such optimizations relies on accurate data concerning the behavioral effect of light spread across different wavelengths within the eye. This study aims to provide foundational data to inform and enhance the design of visual aids and stimuli for improved visual performance. Light spread was measured using two small light points (2 mm apertures in a light shield with a collapsible baffle) back lit with homogeneous monochromatic light (a 1000-W Xenon source used in conjunction with narrow-band interference filters). These points were perceived as one point of light when adjacent. Two-point thresholds were defined as the minimum separation necessary to perceive two non-overlapping points. Subjects sat approximately 67 mm away from the source. Eye position was stabilized using an adjustable chin-and-forehead rest assembly (alignment with the optic axis was done using a small aperture near the plane of the eye physically aligned with the stimulus). 60 young healthy subjects with good acuity were tested. The wavelengths were equated for energy in order to isolate the effects of wavelength (i.e., an action spectrum). The results indicated that the shortest wavelengths required the largest separation ($p<0.001$) with a general monotonic decrease as the wavelength increased (plateauing around 580 nm). These data indicate a strong effect of wavelength on the perception of intraocular scattered light.

43.477 HOW SHARP IS THE (AVERAGE) RETINAL IMAGE?

Charlie S. Burlingham[1] (cburlingham@meta.com), Ian M. Erkelens[2], Oliver S. Cossairt[1], Phillip Guan[1]; [1]Reality Labs Research, Meta, [2]Reality Labs, Meta

Most models of visual perception assume a fully-in-focus retinal image, but nearly the entire retinal image is blurred by optical aberrations. The magnitude of this optical blur in everyday life and its variation across the retina are unclear. Some off-axis aberrations (e.g., astigmatism, coma, spherical) are well-characterized and thought to be relatively invariant to scene depth. On the other hand, defocus aberration depends strongly on scene depth, focal distance, and pupil size — all of which can vary substantially across environments and tasks. Sprague et al. (2016) previously used mobile eye tracking and a geometric model to estimate the natural statistics of defocus blur in human observers performing everyday tasks. However, their analysis only examined the central 20º due to camera field-of-view and eye box limits. In the farther periphery, where other off-axis aberrations may be larger, the natural statistics of blur are unknown. In this work we aim to estimate these statistics using a recently developed set of wide-field (80º x 50º) eye models (Hastings et al., 2024) in realistically modeled indoor and outdoor 3D environments. We use Blender and Zemax to simulate a fixating, accommodating observer, whose fixation depths statistically match those of human observers measured in nine everyday tasks (Burlingham et al., 2024). Using this platform, we model the average blur field across the retina, compare this with previous estimates in the central 20º, and examine variability in blur magnitudes arising from differences in environment, focal distance, pupil size, and individual refractive errors. Our simulations of optical aberrations in everyday life can help identify optical bottlenecks on visual perception, and guide the development of novel displays and rendering methods.

43.478 VARIATIONS IN HUMAN OPTICS EXPLAIN IDIOSYNCRATIC PATTERNS IN THE RED-GREEN SPATIAL CONTRAST SENSITIVITY FUNCTION

Semin Oh[1] (ohsemin1990@gmail.com), Fangfang Hong[2], Derek Nankivil[3], Edda B. Haggerty[4], Robert Stauble[5], Nicolas P. Cottaris[2], John Buch[3], Geoffrey K. Aguirre[2], David H. Brainard[2]; [1]Justus Liebig University Giessen, [2]University of Pennsylvania, [3]Johnson & Johnson, [4]Growth Minded Co., Wayne, US, [5]SeeSharp, Philadelphia, US

The chromatic spatial contrast sensitivity function (csCSF) provides a fundamental characterization of pattern vision. Models of the csCSF often focus on group-averaged data, potentially obscuring meaningful individual variation. Here, we explored individual differences in the csCSF by presenting participants (N = 32) with L-M cone-contrast Gabor images (0.75° SD; stimulus size = 7°) at five different spatial frequencies (3, 6, 9, 12, 18 cpd) using a custom Maxwellian-view hyperspectral display. To support our scientific goals, we did not adjust the stimuli for individual isoluminance. On each trial participants reported the stimulus orientation (tilted +/- 45° relative to vertical). Contrast was varied to determine threshold. While the group-averaged csCSF was low-pass, individual csCSFs had diverse shapes; many exhibited a notch at intermediate spatial frequencies. We modeled the data using a computational simulation. Retinal images were obtained by convolving each stimulus with 79 different polychromatic (i.e. including chromatic aberration) optical point spread functions (PSFs), derived from published wavefront-aberration measurements in a different set of participants. For each PSF, we found a computational-observer csCSF by i) computing the Poisson-distributed cone excitations of a simulated cone mosaic for each retinal image; ii) training a linear classifier to discriminate the stimuli; iii) evaluating classifier performance across contrast. For each of our participants' csCSFs, we found the best-fitting of the 79 computational csCSFs, with one free gain parameter for overall sensitivity. This model captured individual csCSF shapes (R2 = 0.831), outperforming an individually-scaled group-average csCSF model (R2 = 0.685), and approaching the measurement-noise-limited performance of a model that fits CS at each spatial frequency for each participant separately (R2 = 0.890). These findings indicate that optical factors (including chromatic aberration) coupled with sampling by the cone mosaic account for individual variation of the L-M csCSF, and should be incorporated into models of human pattern vision.

43.479 S-CONE INCREMENT AND DECREMENT PERCEPTUAL SCALES ARE NEARLY LINEAR, BUT S-CONE NOISE GROWS WITH CONTRAST: RESULTS FROM MLDS AND PEDESTAL DISCRIMINATION

Yangyi Shi[1] (shi.yang@northeastern.edu), Rhea T. Eskew, Jr.[1]; [1]Northeastern University

The short-wavelength (S) cone photoreceptors exhibit unique characteristics compared to the long-wavelength (L) and middle-wavelength (M) cones. To investigate the perceptual representation of S-cone contrasts at suprathreshold levels, we conducted two psychophysical experiments, building on prior studies with achromatic stimuli (Shi & Eskew, 2024; Shooner & Mullen, 2022). In the first experiment, we used maximum likelihood difference scaling (MLDS) (Knoblauch & Maloney, 2008; Maloney & Yang, 2003), with stimuli ranging from twice the detection threshold to the maximum monitor contrast. Observers compared two pairs of contrasts in each trial and judged which pair appeared more similar. In the second experiment, the same observers performed a two alternative forced choice pedestal discrimination task with the same stimuli used in perceptual scaling. Both experiments separately examined S-cone increment (S+) and decrement (S−) stimuli. Our results reveal that, unlike our previous achromatic findings (Shi & Eskew, 2024), the estimated perceptual scales for both S+ and S- are approximately linear and approximately symmetric. However, consistent with Gabree et al. (2018), pedestal discrimination thresholds rise more rapidly for S+ pedestals than for S- pedestals. At higher pedestal contrasts, S+ pedestals produce stronger masking effects than S− pedestals. The original MLDS model, which assumes a constant magnitude of perceptual noise, cannot account for the increasing discrimination thresholds when the scales are linear, as in both S+ and S− conditions. We propose and fit a variable noise MLDS model, where noise grows with the square root of contrast. As Kingdom (2016) showed, whether noise is constant with contrast or not has little effect on perceptual scales generated by a difference-scaling method, but contrast-dependent noise strongly affects discrimination threshold functions. Our results show that with S cone stimuli, the increasing noise allows us to predict pedestal discrimination thresholds from MLDS results.

NSF BCS-2239356

43.480 A WISHART PROCESS MODEL COMBINED WITH ADAPTIVE SAMPLING FOR EFFICIENTLY CAPTURING DISCRIMINATION THRESHOLDS IN HIGH-DIMENSIONAL STIMULUS SPACES

Alex Williams[1,2], Fangfang Hong[3], Craig Sanders[4], Michael Shvartsman[5], Phillip Guan[4], David Brainard[3]; [1]Center for Neural Science, New York University, [2]Center for Computational Neuroscience, Flatiron Institute, [3]Department of Psychology, University of Pennsylvania, [4]Reality Lab Research, Meta, [5]FAIR, Meta

Traditional methods for trial selection do not support exhaustive characterization of perceptual thresholds in high-dimensional settings; they require exponentially more data with increasing stimulus dimensionality. Here, we study the efficacy of adaptive sampling for trial selection combined with a Wishart Process model (WPM) for dense characterization of thresholds. Specifically, we simulated color-discrimination performance using the CIELAB color space, which is intended to be perceptually uniform and which coarsely approximates human color sensitivity. We restricted attention to a plane in color space. We generated CIELAB-based 'ground-truth' responses for a forced-choice task in which an observer identifies the odd-colored stimulus (comparison) from two other stimuli that are identical to each other (reference). The reference (R & B calibrated monitor channel intensities) and comparison stimuli (R+ΔR & B+ΔB) were chosen using AEPsych, a non-parametric method that adaptively selects informative trials (Owens et al., 2021). The monitor G intensity was

held fixed. We fit the generated responses with a semi-parametric WPM which expresses smoothness of performance over the stimulus space, on the assumption of multivariate-Gaussian performance-limiting internal noise. This approach accurately recovered the 'ground-truth' thresholds with 2,800 trials, ~10x fewer trials than required by conventional methods (~30,000 assuming 25 reference stimuli x 8 comparison directions/reference x 150 trials/direction). Evaluation using the Bures-Wasserstein distance showed close agreement between 'ground-truth' threshold ellipses and model predictions (mean = 0.009, SD = 0.003; compare to a baseline set by distance between the 'ground-truth' ellipses and their inscribed circles: mean = 0.040, SD = 0.010). Crucially, our approach densely characterizes discrimination, interpolating to predict thresholds for any comparison direction around any reference. The approach generalizes to higher stimulus dimensions (full color space) and to any modality where smoothly varying multivariate-Gaussian noise limits performance, for example auditory localization and motor reaching.

43.481 EFFICIENT CHARACTERIZATION OF HUMAN COLOR DISCRIMINATION THRESHOLDS USING ADAPTIVE SAMPLING AND A WISHART PROCESS MODEL

Fangfang Hong[1] (fh862@sas.upenn.edu), Ruby Bouhassira[1], Craig Sanders[2], Michael Shvartsman[3], Alex Williams[4], Phillip Guan[2], David Brainard[1]; [1]Department of Psychology, University of Pennsylvania, [2]Reality Lab Research, Meta, [3]FAIR, Meta, [4]Center for Neural Science, New York University

Discrimination thresholds are foundational for understanding the limits and mechanisms of color vision. Fully characterizing these thresholds has been intractable because, with standard methods, trial number grows exponentially with stimulus dimensionality. We overcome this challenge using adaptive trial placement (AEPsych, Owens et al. 2021) and a semi-parametric Wishart Process model (WPM). The WPM leverages smoothness in performance variation and enables interpolation of thresholds across the full color gamut. As proof of concept, we measured color discrimination across the DKL isoluminant plane using a three-alternative forced-choice task. On each trial, participants (N = 5) identified the odd-colored stimulus. Nine reference stimuli were sampled, with comparison stimuli chosen around each reference by AEPsych (3,240 trials/participant). We also ran blocks of validation trials using the method of constant stimuli (MOCS) for both tested and untested reference stimuli (2,880 trials/participant). We fit the WPM to infer the covariance matrices of a multivariate Gaussian internal noise field. From the field, we predict percent correct discrimination for any reference and comparison. Thresholds from validation trials were highly correlated with WPM predictions, but were systematically lower (mean correlation = 0.878, mean ratio = 1.374). We think the bias is due to reduction in uncertainty for the MOCS trials. The data reveal consistent patterns across participants: (1) thresholds are smallest at the achromatic reference; (2) thresholds increase with reference distance from achromatic; (3) threshold ellipses are oriented toward the achromatic reference; (4) thresholds deviate substantially from predictions based on CIELAB. In sum, our approach enables a complete characterization of color thresholds across the isoluminant plane. We expect it will generalize to ellipsoids in full 3D color space, and the validation bias will diminish when MOCS trials are run fully interleaved. Moreover, our method induces a Riemannian geometry on color space that provides a candidate perceptual distance metric.

Meta

43.482 COLOR CODING FOR MULTI-CHANNEL COLOR PERCEPTION FROM THREE PHOTODETECTOR TYPES WITH WIDE OVERLAPPING SPECTRAL SENSITIVITY BANDS

Vasile Diaconu[1], Jocelyn Faubert[2]; [1]Université of Montréal, [2]Université of Montréal

Introduction. Human color perception generate typical chromatic sensations from various wavelengths of the visible spectrum by exciting three broadband sensitivity photodetector types. The retinal neural network compares, differentiates and redirects signals from the three photodetectors to the cortex through several chromatic pathways. The current concept that explains the creation of the different color pathways through the retina's neural network is not compatible with its implementation in a physical instrument to characterize colors in a way similar to the visual system so the mechanism of neural color-coding of the retina is partially known. Purpose. The present study presents a physical process of de-multiplexing signals emitted by two or more types of photodetectors with a wide band of overlapping spectral sensitivities to differentiate signals corresponding to the spectral zones that the photodetectors share with each other, and signals corresponding to the distinctive spectral sensitivity zone of each photodetector. Conclusion. The model adopts two fundamental principles from retinal neuron signal processing, such as the contrast of photoreceptor signals, as well the ON and OFF properties of retina's neural network to redirect the contrasts of photoreceptor signals to different chromatic channels. The concept of this model represents a good alternative to explain the process of color coding signals from three types of retinal photodetectors to activate the different channels involved in human color perception.

Color, Light and Materials: Lightness and brightness

Monday, May 19, 8:30 am – 12:30 pm, Pavilion

43.483 LAMINAR ORGANIZATION OF SHADOW-DISCOUNTED LIGHTNESS SIGNALS IN AREA V4

Fatemeh Didehvar[1] (dfatemeh@wustl.edu), Patrick Cavanagh[2], Tom P. Franken[1]; [1]Washington University in St. Louis School of Medicine, [2]Glendon College, York University

Although the light reflected by an object varies directly with the amount of light falling on it, we perceive the object's reflectance, its lightness (how black, gray, or white it is), as nearly constant. It is still poorly understood how the visual system does this: discounting the illumination to recover reflectance. Low-level theories emphasize spatial filtering operations at early stages of the visual system, whereas mid- and high-level theories propose that discounting occurs at higher stages. Here we leveraged high-density laminar recordings to study these computations in the primate brain. We used Neuropixels to record single- and multiunit neural activity from visual area V4 while a macaque viewed various scenes. On different trials we presented either shadow boards (checkerboards scenes partially in shadow), or paint boards, where the shadow edge was destroyed by averaging the luminance in each square. We presented the scenes such that the receptive field was centered on one square of the boards. We used current-source-density analysis to locate units to superficial, input or deep cortical layers. We then analyzed the shadow-discounted lightness signal (SDLS): we trained random forest decoders to predict the luminance of the square in the receptive field from neural responses to paint boards, and tested the decoders on responses not used during training, either to shadow boards (test-shadow) or to paint boards (test-paint). The SDLS, defined as the difference in predicted luminance between test-shadow and test-paint, was significantly positive (n=10 penetrations). We also found that the SDLS was not significant for neural populations in the input (granular) layer, but only in extragranular layers. Our experiments reveal shadow-discounted lightness signals in area V4. The laminar pattern suggests that these signals are computed at the level of V4 or higher areas, consistent with mid- and high-level theories of lightness.

This research was supported by the National Eye Institute of the National Institutes of Health under Award Number R00EY031795 and the Small Grants Program from the McDonnell Center for Systems Neuroscience.

43.484 EFFECT OF OPTICAL ABERRATION ON HUE AND LUMINANCE DEPENDENCY OF COLOR ASSIMILATION

Natsumi Tsuji[1], Tama Kanematsu[2], Kowa Koida[1]; [1]Toyohashi university of technology, [2]Kyushu University

Color assimilation is a perceptual phenomenon where the color appearance of the center region assimilates to the surrounding area. This phenomenon is driven by both optical and neural factors. The optical factors involve the smoothing of the retinal image due to chromatic aberration and light scattering, while the neural factors refer to the spatial summation during post-receptoral processing. While these factors are summed, the relative contributions of each to color assimilation remain unclear. To investigate this, we conducted a psychophysical experiment and optical simulation of the eye. In the psychophysical experiment, we measured the effects of inducer hue and background luminance on color assimilation using stimuli imitating a watercolor effect. Observers viewed illusions with varying background luminance and reported whether color assimilation was perceived. Color assimilation typically occurred when the luminance contrast between the inducer and background was low, although significant hue dependency was observed. Specifically, a relatively high luminance background was preferred for red inducers, while a low luminance background was preferred for blue inducers. Additionally, red inducers often elicited no color assimilation responses. The optical simulation, based on the spectral optical transfer function, was performed on the same images used in the psychophysical experiment. The resulting retinal images were analyzed for luminance contrast between the inducer and the center region. We found that retinal images with the lowest luminance contrast varied depending on the hue of the inducer. This hue dependency was consistent with the psychophysical results, suggesting that the hue and luminance dependencies of color assimilation could be attributed to optical aberrations.

43.485 THE EQUILUMINANT REMOTE CONTROLS ILLUSION

Alexander M Gokan[1] (ag6832a@american.edu), Arthur G Shapiro[1]; [1]American University

The remote control illusion (RCI) (Hedjar, Cowardin and Shapiro, 2018) is a method for measuring the spatial aspects of contrast perception. The luminance levels of two identical rectangular bars modulate at 2 Hz; when one bar is placed on a bright field and the other on a dark field, observers perceive the bars modulating in antiphase with each other (like contrast between the bars and their surround) and becomes light and dark at the same time (like the luminance of the bars). In RCI, bright rectangular flankers are added to both sides of the dark-field bar and dark flankers on both sides of the bright-field bar. An antiphase appearance occurs when flankers directly adjoin the bars or are separated by > 1°, but appear inphase when flankers are separated from the bars 20'. Here, we examine RCI with lights modulating along the L-M axis. In experiment 1, we compared the effect of luminance and equiluminance modulation. The configuration was similar to Hedjar et. al., and equiluminance was set using motion photometry. We varied the gap between bars and flankers and measured the proportion of antiphase trials each gap size. The luminance configuration replicates the original paper (that is, perception followed a "v" pattern: antiphase–inphase–antiphase as the gap between bar and flanker increased), but this pattern was not shown for the equiluminant configuration (typically, inphase-inphase-antiphase). In Experiment 2, we remeasured the patterns with modulation along color lines ranging from equiluminance to a 6% luminance and show that small shifts in color angles produce a transition between the equiluminance and luminance patterns. In Experiment 3, we show that inphase perception that arose when equiluminant flankers adjoin the bars, disappears when the width of the flankers increases, indicating that the LM contrast integrates over a larger area than luminance contrast.

43.486 RELATIVE LUMINANCE AND THE WATERCOLOR ILLUSION IMPACT FIGURE-GROUND IN A PROBED REGION PARADIGM

Tanner L. Lumpkin[1], Hannah V. Hyman[2], Ralph G. Hale[3]; [1]University of North Georgia

The watercolor illusion (WCI) is a color spreading illusion induced by contrasting outer and inner boarders, resulting in the dissemination of a hue similar in appearance to the lighter border. The illusion has been shown to be a strong Gestalt cue for figure-ground (F/G) organization. The color spreading area tends to be perceived as figure. In our present study, we examined the effects of the WCI on F/G assessment for regions with non-matched luminance contrast. Participants were exposed to two matched luminance conditions with gray and white ambiguous F/G regions. Each image was divided into two parts by a vertical wavy contour. Each had no WCI, WCI left, and WCI right versions. A probe was present in either the left or right region, and participants were instructed to report whether the probed region appeared as figure. Results showed a significant interaction between the probe and WCI for both matched conditions, indicating an effect of WCI on F/G assignment. The non-matched condition consisted of contrasting luminance regions. The conditions and reporting method was the same as the other two experiments. There was a significant interaction between the probe and grey region, indicating the gray region was seen as figure more often. However, the WCI failed to bias F/G in this experiment. This study furthers our understanding of the WCI interaction with contrasting luminance regions. These results enhance our understanding of color spreading mechanisms and how they interact with luminance, contrast, and perceptual organization.

43.487 EFFECT OF MATCHED RELATIVE CONTRAST ON AMBIGUOUS FIGURE-GROUND AND THE WATERCOLOR ILLUSION

Hannah Hyman[1], Tanner Lumpkin, Ralph Hale; [1]University of North Georgia

The watercolor illusion (WCI) is a visual phenomenon where a thin colored contour adjacent to a second darker contour induces color spreading, causing the enclosed region to appear tinted and more figure-like. Previous research has shown that the WCI can bias figure-ground (F/G) organization, with the WCI region typically seen as figure. This effect has also been observed to vary with luminance. In the present study, we further examined the interaction between WCI and luminance effects on ambiguous F/G organization. Unlike prior studies, we controlled for contrast by equalizing it between the background and possible figure regions, aiming to isolate the effects of luminance and illusion strength on F/G assignment. The WCI is known to be less visible on dark (black) backgrounds than on light (white) backgrounds, prompting us to investigate whether this difference in illusion magnitude affects F/G organization. We conducted three experiments using ambiguous F/G stimuli with central shapes composed of two regions in different luminance combinations: (1) white and white, (2) black and black, and (3) white and black. A mid-gray screen surrounded each stimulus. Each experiment included four conditions: no WCI, WCI on the left, WCI on the right, and WCI on both. A probe appeared in either the left or right region, and participants indicated if the probed region appeared as figure. Our findings reveal interactions between WCI and luminance on F/G assignment. This research is the first of its kind to explore color spreading and luminance in this manner, contributing novel insights into the role of luminance and the WCI on perceptual organization.

43.488 TESTING A MODEL OF ACHROMATIC COLOR APPEARANCE MATCHING FUNCTIONS: VARYING THE BACKGROUND

Osman Kavcar[1], Michael E. Rudd[1,2], Christabel Arthur[1], Michael A. Crognale[1,2]; [1]University of Nevada, Reno, [2]Center for Integrative Neuroscience

We investigated brightness induction using disk/annulus stimuli. Observers adjusted the brightness of a match disk to match the brightness of a fixed-luminance target disk surrounded by an annulus whose luminance was varied to induce brightness changes in the target. Unlike the classical paradigm (Wallach, 1948), the match disk was not surrounded by an annulus. When the resulting brightness matches are plotted against annulus luminance on a log-log scale, the relationship is well described by 2nd-order polynomial functions, which we refer to as achromatic color appearance matching functions (ACAMFs; Kavcar, Rudd, & Crognale, 2025). Rudd, Kavcar, and Crognale (2023) proposed a brightness induction model for our stimuli based on edge integration combined with contrast gain control, and the auxiliary assumption that observers discount the outer annulus edge on the target side with this display. The model predicts that the 2nd-order coefficients (k2) of the polynomial ACAMFs should be proportional to the 1st-order coefficients (k1) with a proportionality constant that depends on the contrast polarity of the target disk to its annulus, the target disk luminance, and the strength of the contrast gain control. Our previous work established experimentally that this relationship does not depend on the annulus size but does depend on the target disk luminance and contrast polarity. Here, we further tested the model by varying the background luminance. Our results support an additional model prediction that the relationship between 1st- and 2nd-order coefficients should not depend on the background. The slope of a linear regression line fitted to the k1 vs k2 plots remained virtually unchanged as the background luminance was varied; and was ~-3.30 for conditions in which the target was a luminance decrement, and -~1.74 for conditions in which the target was a luminance increment. These values replicate the values obtained in our earlier work.

Research reported in this paper was supported by National Institute of General Medical Sciences of the National Institutes of Health under grant number P30 GM145646. The content is solely the responsibility of the authors.

43.489 COMPARING DEEP NEURAL NETWORK ARCHITECTURES AS MODELS OF HUMAN LIGHTNESS AND ILLUSION PERCEPTION.

Jaykishan Patel[1,2], Richard Murray[1,2], Javier Vazquez-Corral[3,4], Konstantinos Derpanis[5]; [1]York University, [2]Department of Psychology and Center for Vision Research, [3]Department of Computer Science, Universitat Autònoma de Barcelona, [4]Computer Vision Center, [5]Department of Electrical Engineering and Computer Science

Inferring surface reflectance from luminance images is challenging, as identical luminance patterns can result from different combinations of illumination, reflectance, and 3D shape. Classical models have struggled with this ambiguity, but convolutional neural networks and vision transformers show promise in estimating surface color under varying illumination. This study tests the effects of model size, architecture, and the transferability of features from pretraining on a different task for reflectance estimation, along with whether different models utilize similar features for illusion perception. Using ResNet34, VGG13, MobileNetV3, and dense prediction transformer as encoders in a UNet architecture, we trained decoders to estimate reflectance from luminance images in a custom Blender-generated dataset. Encoders with pre-trained ImageNet or depth prediction weights were trained for reflectance estimation. Frozen encoders were utilized to evaluate whether their features could transfer to reflectance estimation without fine-tuning. Both frozen and fine-tuned models performed well on reflectance estimation, with frozen models being slightly less accurate. Model responses were computed for several illusions, including the argyle, Koffka-Adelson, snake, simultaneous contrast, White's, and checkerboard assimilation. Model responses were consistent with illusions perceived by human observers. Furthermore, illusion-like responses were weaker in control conditions, except for the argyle and assimilation illusions. Low-parameter models performed as well as high-parameter models on illusion perception but were less accurate in reflectance estimation, challenging the hypothesis that illusions arise from efficient coding. Saliency analysis showed that for all models, similar regions were responsible for perceived illusions, often focusing on shadowed areas. Saliency maps showed high correlation for regions contributing to illusion perception in all models, with slightly lower agreement for frozen models. These results suggest lightness illusions arise from visual systems' using natural scene statistics to generate an accurate perceptual correlate to reflectance and highlight the potential of deep learning architectures for modeling human lightness and color perception.

TUESDAY MORNING POSTERS IN BANYAN BREEZEWAY

Eye Movements: Social, individual differences, visual preferences

Tuesday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

53.301 AGE-RELATED DIFFERENCES IN THE TIME COURSE OF VISUAL NOVELTY PREFERENCE FOR CATEGORICAL AND FUNCTIONAL ASSOCIATIONS OF REAL-WORLD OBJECTS

Elena Allegretti[1] (elena.allegretti@uniroma1.it), Jennifer D. Ryan[2], Moreno I. Coco[1,3]; [1]Sapienza, University of Rome, Rome, Italy, [2]Rotman Research Institute, Baycrest, Toronto, Ontario, Canada, [3]I.R.C.C.S. Fondazione Santa Lucia, Rome, Italy

Novel information is attentionally prioritised, as empirically shown using the Visual Paired Comparison task (VPC). This attentional orienting reflects associative memory processes and is influenced by the strength of encoding for familiar items. Memory associations rely on perceptual and affordance-related features to establish categorical (e.g., apricot/grape) and functional (e.g., nut/nutcracker) relationships between objects. Yet, it is still unclear how these associations can impact their encoding strength in memory, indirectly influence attentional guidance when objects are explored, and whether age-related changes may alter these processes. This eye-tracking study adapted the classic VPC task to investigate the attentional dynamics of the novelty preference (NP) while manipulating categorical and functional relationships between object pairs. Twenty-five younger and twenty-four older adults were familiarised with object pairs, varying in the above-mentioned relationships, and then tested with novel and familiar pairs. We measured changes in the odds ratio of fixation proportion between novel and familiar pairs over the test trials, which were transformed into empirical logits. Growth-curve analysis revealed that functionally related pairs determined a stronger NP than categorically related pairs in younger adults than older ones. This indicates a greater reliance of older adults on categorical information to organise and store functional relationships in memory. These findings provide theoretical insights into how conceptual knowledge's hierarchical organisation mediates overt attention allocation, with action-oriented functional associations enabling stronger representations. Furthermore, they highlight age-related shifts in encoding conceptual associations, suggesting that categorical associations serve as cognitive scaffolds to support the organization and integration of functional relationships in ageing memory processes.

Funded by NSERC and CIHR. EA is supported by Progetti di Avvio alla Ricerca, AR12419073F84B6A (Sapienza, University of Rome).

53.302 TO LOOK OR NOT TO LIKE: OCULOMOTOR-CONTROL MECHANISMS ALTER STIMULUS-VALUE REPRESENTATIONS.

Niyatee Narkar[1] (nnarkar@uoguelph.ca), Cassidy D. Darechuk[2], Mark J. Fenske[3]; [1]University of Guelph, [2]University of Guelph, [3]University of Guelph

Cognitive-control mechanisms that determine which information becomes the focus of our attention and actions have affective consequences for associated visual stimuli. Ignoring or withholding a response from a stimulus, for example, can negatively impact its perceived value. Such stimulus devaluation is thought to be due to negative affect elicited by attention- and response-related inhibition. Using head-stabilized screen-based eye-tracking, we searched for similar effects in the oculomotor domain by combining tasks involving inhibitory control over eye-movements with affective evaluations of stimuli. Art-like patterns were first centrally presented in an oculomotor Go/No-go task. A central cue then prompted participants to either Go (look at an abrupt-onset stimulus appearing to the left or right of the pattern) or No-go (avoid making any eye-movements and instead maintain fixation on the pattern). Liking ratings obtained after each Go/No-go trial revealed that No-go-trial patterns were evaluated more negatively than Go-trial patterns, despite any fluency-related enhancement from longer foveal processing. Previously-unseen novel patterns were also disliked if rated shortly after a No-go trial than after a Go trial, suggesting lingering impacts of oculomotor inhibition on the coding of stimulus value. Similar results from experiments that interspersed anti-saccade or selective-looking trials with affective-evaluation trials suggest the mechanisms underlying ‘distractor devaluation’ and ‘No-go devaluation’ effects in other selective-attention and motor-response control domains may be similar to those influencing stimulus value in the oculomotor domain. Ongoing analysis of eye-movement data from selective-looking tasks will further reveal whether the trial-by-trial fluctuations in distractor suppression reflected in deviations in saccade trajectory can predict the magnitude of oculomotor distractor devaluation. This research underscores the potential significance of the link between inhibition and aversive response as a manifestation of the interaction between emotion and oculomotor-control systems as they work together to ensure that distracting or otherwise-problematic stimuli can be effectively avoided in the future.

53.303 DON'T STOP BELIEVIN'—IN THE LINK BETWEEN EFFICIENT SCANNING AND WORKING MEMORY CAPACITY

McKenzie Sheets[1] (kmulli21@vols.utk.edu), Micalee Segers[2], Tori Jones[3], Chloe Mitchell[4], Shannon Ross-Sheehy[5]; [1]University of Tennessee

There is a well-established relationship between working memory (WM) and visual attention. Previous studies have shown that the contents of WM can influence looking behavior (Soto et al., 2008; Woodman & Luck, 2007). Additionally, there has been a substantial amount of work indicating that individual differences in WM capacity (WMC) are related to performance on a variety of cognitive tasks (Unsworth & Engle, 2007). However, there is little work investigating the relationship between visual WMC and scanning patterns. Individual differences in WMC may be related to scanning efficiency (SE), such that individuals with high WMCs may employ more efficient scanning strategies than those with lower WMCs. To investigate this relationship, we used data collected from adults (n=55) who completed both a scanning task and a change detection task. Gaze was sampled at a rate of 500Hz using EyeLink 1000+ in both tasks. The scanning task consisted of four blocks of five 10s trials, in which participants were presented with two static faces (one male and one female), counterbalanced across blocks. For the change detection task, visual arrays consisted of colored circles (ss3,ss5,ss7,ss9), and each trial included a 100ms sample array, 900ms retention interval, and 3000ms test array. SE scores were calculated for each participant using fixation durations and subsequent saccade amplitudes (see Ross-Sheehy et al., 2022) during the scanning task. Preliminary results indicate a positive relationship between current fixation durations and subsequent saccade amplitudes, $t(54)=11.99, p<0.001$. WMC (k) was calculated for each participant based on their change detection performance. Participants were categorized as having high or low SE, and preliminary findings suggest a relationship between SE and WMC, as participants with low SE have higher WMC(M=4.77) than those with high SE(M= 4.30), $t(51)=-1.98, p=0.05$. While these findings do not support our hypothesis, further analyses will explore the relationship between WMC and SE.

53.304 THROUGH THEIR EYES: INVESTIGATING THE BROAD AUTISM PHENOTYPE IN PARENTS

Elena Esposito[1], Vanessa Harwood[1], Alisa Baron[1]; [1]University of Rhode Island

The Broad Autism Phenotype (BAP) refers to traits that resemble those of autism spectrum disorder (ASD), commonly seen in first-degree relatives of individuals with ASD suggesting a potential genetic contribution. Understanding how these traits manifest in communication remains limited. Previous research has shown differences in eye gaze patterns between children with and without ASD. This study explores whether similar patterns exist in parents of children with autism (pASD) during a phonemic restoration task. The Social Responsiveness Scale (SRS) and Broad Autism Phenotype Questionnaire (BAPQ) were also used to assess ASD-like traits related to social communication.The study involved 38 adults (both pASD and controls) in the U.S. Participants viewed a speaking face producing a consistent visual /ba/ sound, while the audio alternated between /ba/ and /a/, creating a phonemic restoration effect as the participants “fill in” missing phonemes based on context and linguistic expectations. Using linear mixed-effects models, preliminary eye gaze patterns to the eyes and mouth/jaw areas of 7 pairs of participants indicated no significant differences between groups. Independent sample t-tests revealed a trend toward significance for group differences in SRS scores (p = .06), with the pASD group scoring higher and no significant group difference was observed in BAPQ scores (p = .66). Although no significant group differences were found in eye gaze patterns or social communication traits, the study contributes to a deeper understanding of the BAP. Further analysis with a larger sample size may reveal more nuanced, significant findings.

53.305 READING THE MIND THROUGH THE EYES: EYE MOVEMENTS CONVEY USEFUL INFORMATION ON THEIR PERFORMER’S COGNITIVE STATE

Amit Zehngut[1], Shlomit Yuval-Greenberg[1]; [1]Tel-Aviv University

People who are engaged in effortful cognitive processing often avert their gaze peripherally. This gaze aversion (GA) behavior has been interpreted as a distraction-avoidance mechanism, yet evidence also suggests that it is linked to social cognition. Although GAs occur even when individuals are alone, they become larger and more frequent in the presence of others. Therefore, we hypothesize that, in addition to their perceptual role, GAs function as social signals, conveying information about the performer’s cognitive state. In two large-scale studies, we tested this hypothesis by examining how observers interpret others’ oculomotor behavior. In Experiment 1, participants (N=120) watched short (5s) muted videos of individuals, who were either thinking, listening, or performing a tapping task. We employed three between-subject masking conditions: in the no-mask condition, the videos were fully visible; in the eyes-only condition, only the eyes were visible; and in the no-eyes condition, the eyes were occluded while the rest of the image remained visible. Participants were asked to determine whether the depicted individual was thinking, listening, or tapping. Results showed that participants correctly identified the depicted individuals' cognitive state above chance-level. The effect was larger in the no-mask and eyes-only conditions but was attenuated in the no-eyes condition. In Experiment 2, we examined the spatial characteristics of this effect. Individuals were filmed while performing goal-directed saccades toward nine distinct locations. Participants (N=60) who viewed these videos tended to interpret left and right upward saccades with higher eccentricity as indicators of thinking. Taken together, these studies suggest that GAs with certain characteristics function as social signals indicative of effortful cognitive processing. Such signaling enables interlocutors to "read the mind" of others, thereby enhancing nonverbal communication. By integrating cognitive paradigms with ecologically valid social stimuli, this research underscores the important, yet rarely explored, link between social attention and oculomotion.

This study was funded by ISF grant 1960/19 to S.Y-G and by Minducate scholarship to A.Z

Eye Movements: Natural or complex tasks

Tuesday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

53.306 A HEAD-MOUNTED HIGH-RESOLUTION DUAL PURKINJE IMAGING EYE-TRACKER

Paul Jolly[1,2], Anis Idrizović[1,2], Soma Mizobuchi[1,2], Ruei-Jr Wu[1,2], Michele Rucci[1,2]; [1]University of Rochester, [2]Center for Visual Science, Rochester, USA

Measuring eye movements with high precision is crucial in vision science. Building upon the classical dual Purkinje imaging (DPI) method devised by Cornsweet and Crane in 1973, we have recently developed a digital DPI eye-tracker (dDPI; Wu et al, Journal of Vision 2023). This system leverages modern digital imaging technology and computational power to yield robust measurements with sub-arcminute precision while circumventing several limitations of the analog device. To achieve high-resolution, the dDPI requires immobilization of the observer's head, which limits the range of possible applications. Here we present a head-mounted dDPI, a system that enables precise oculomotor tracking while leaving the head free to move normally. The system uses a high-speed camera to image the first and fourth Purkinje reflections (the reflections from the cornea and lens) of an infrared beam. A GPU-based algorithm estimates the relative positions of the Purkinje images. The apparatus is capable of resolving arcminute rotations, as demonstrated by means of both artificial eyes and measurements from real eye movements. When coupled with a motion capture system that tracks head movements, this device yields an error in localizing the line of sight during normal head movements smaller than a few arcminutes. In sum, a head-mounted dDPI appears to be a valuable tool for studying active vision under natural conditions.

This work was supported by National Institutes of Health grants EY018363 and P30 EY001319.

53.307 TRACKING MOVING OBJECTS IN THE REAL WORLD

Jie Z. Wang[1] (jwang255@ur.rochester.edu), Paul Jolly[1], Michele A. Cox[1], Y. Howard Li[1], Ruitao Lin[1], Alina Neverodska[1], T. Scott Murdison[2], Michele Rucci[1]; [1]University of Rochester, NY, USA, [2]Meta Reality Labs, Redmond, WA, USA

Real-world objects often move in depth following complex trajectories. However, laboratory studies on human oculomotor tracking commonly use 2-D targets with uniform motions and predictable trajectories. Moreover, to reliably measure eye movements, the observer's head is typically immobilized, thus preventing examination of head-eye coordination, a crucial aspect of natural tracking behavior. Here we studied natural head-eye tracking of a real object moving in 3-D space. Precise measurements of eye and head movements were obtained by means of a custom apparatus that combines a specifically-designed magnetic induction coil-based eye-tracker with a motion-capture system. Subjects (N=8) were seated within a cubic cage while wearing scleral coils on both eyes and a tightly-fitting helmet equipped with coils and markers. Observers participated in a Plinko-inspired task: they were instructed to track a small disc dropped from the top of a slanted Galton board and report when the disc reached the bottom. The target disc traveled down approaching the observer and bounced between the obstacles on the board, resulting in a complex path. Our results show that observers tracked the target via a combination of eye saccades and head pursuits. In between saccades, head movements, but not eye movements, were effective in attenuating retinal motion. A subgroup of observers was also tested in a control condition in which the disc fell at slower speed along a track on the board, a condition more similar to standard oculomotor studies with predictable, slow motions. Here subjects relied on binocular smooth pursuit with occasional saccades, so that, on the retina, both the eye and head contributed to reducing the target's speed, with eye rotations playing the dominant role. These results show that natural 3-D tracking relies on distinct head/eye strategies depending on the target speed and the predictability of motion.

Supported by Reality Labs and NIH EY018363 and P30 EY001319

53.308 PREDICTING GAZE BEHAVIOR IN NATURAL WALKING ENVIRONMENTS

Lloyd Acquaye Thompson[1], Zachary Petroff[1], Stephanie M. Shields[1], Kathryn Bonnen[1]; [1]Indiana University Bloomington

Understanding human gaze behavior during naturalistic tasks is critical for understanding how visual processing supports real-world activities. Previous work has developed spatial saliency models that predict gaze in static images. Here we examine how well these spatial saliency models predict gaze in dynamic environments during naturalistic tasks. Specifically, we evaluated the gaze prediction performance of DeepGazeIIE (Linardos, Kümmerer, Press, & Bethge, 2021), a prominent spatial saliency model, for data from a previous study of walking in natural outdoor environments that captured scene video and gaze data with a mobile eye tracker (Bonnen et al., 2021). We extracted video frames from the head-mounted scene camera video and included them only if gaze estimation confidence was at least 95%. For each video frame, the model produced a log-likelihood map that indicates the likelihood of fixation at each point in the frame. We evaluated the network's performance on our dataset (without additional training) using the area under the curve (AUC) at each recorded fixation location. Initial results reveal that DeepGazeIIE provides good predictive accuracy for fixation patterns during natural walking (AUC=0.81). Notably, this is lower than the AUC reported on the original benchmark dataset (AUC=0.88). This performance gap may stem from differences between the two datasets, e.g., differences in the task and the visual environments. For example, walkers rarely look at their feet, but because feet are visually salient objects, the model predicts that walkers are highly likely to fixate on their feet. Future work will investigate and test strategies that address these limitations and better account for task-specific patterns (e.g., training a network specifically on egocentric video frames). Broadly, this study demonstrates the feasibility of leveraging existing spatial saliency models to analyze gaze behavior during natural tasks, offering insights into visual processing in dynamic contexts.

53.309 GAZE FOLLOWING IN MARMOSET MONKEYS IS CONTEXT DEPENDENT

Oviya Mohan[1], Dora Biro[1], Jude Mitchell[1]; [1]University of Rochester

Gaze following, the ability to direct one's attention to the target of another's, plays a key role in social interactions and can help individuals attend to relevant locations and events in their environment. Gaze following has been considered reflexive, in which case it might operate invariant to whether or not it leads to obtaining relevant social information. Although gaze following has been demonstrated in the common marmoset (Callithrix jacchus), the stimuli used are often not representative of the natural social settings in which it typically occurs. We presented head-fixed marmosets with naturalistic stimuli (videos) of freely behaving marmosets exhibiting directed gazing as a means of eliciting reliable gaze following behavior. Each video featured a marmoset entering a center box and gazing toward a box on either the right or the left. Two trial types, one where a target marmoset appeared in the cued box and another where no target marmoset appeared in either box, were presented interleaved covering a range of +/- 10 visual degrees (previously verified to fall within the oculomotor range of marmosets). The first saccades that landed outside the center box were categorized into cued and distractor regions for valid trials (head-fixed marmoset looks at cueing marmoset). In both trial types, significantly larger number of first saccades landed in the cued region than in the distractor region ($p < 0.001$) confirming reliable gaze following by the head-fixed marmoset in this paradigm. We further tested if gaze following is impacted in a non-informative context. We included a novel trial type in which the cued and distractor boxes were occluded by cardboard, eliminating the possibility of viewing the entry of the target marmoset. Preliminary results show differences in gazing behavior in this context indicating that gaze following is context dependent.

Funding: NIH grant EY030998

53.310 DECREASED EYE MOVEMENT IS ASSOCIATED WITH PERCEIVED BLURRING OF OBJECT BOUNDARIES IN MEDITATION

Briana Alcindor[1], Amy Patel[2], Lekhana Ramachandra[3], Jovelis Toribio[4], Melisa B. Sharkly[5], Robert G. Alexander[6]; [1]New York Institute of Technology

Meditation has been linked to many changes in perception, including colors becoming more vivid, changes in brightness, and the blurring of boundaries between objects. The mechanisms behind these perceptual changes are unknown. Here, we describe a “meditation” task that we conducted in the lab, which we can use to test different possibilities. In the current experiment, we presented an image of a clay pot on a computer monitor for 15 minutes. Participants sat at a fixed distance in front of the monitor and were told to focus all their attention on the clay pot and exclude all other thoughts or feelings. This task strongly resembles many forms of meditation. We then asked participants a series of debriefing questions designed to elicit comments about perceptual effects that might have occurred but without the potential for participants to simply agree with our expectations. While many participants reported that nothing happened during the task, 27 participants (out of 58) reported perceptual changes. Consistent with our predictions, the eye tracking data revealed that participants who reported perceived changes to the boundary had both smaller and fewer saccades. This suggests a connection between reduced eye movements and perceptual fading of object boundaries during meditation. The reduction of eye movements was not found in participants who reported other perceptual changes such as brightness. These other changes are presumably driven by some other underlying mechanism. By knowing when the effects occur during the experiment, we can explore whether other physiological events happen around those times and thus begin to determine the mechanisms underlying these experiences.

This work was supported by a New York Tech Institutional Support for Research and Creativity grant.

53.311 EDIA: AN OPEN-SOURCE TOOLBOX FOR VIRTUAL REALITY-BASED EYE TRACKING RESEARCH USING UNITY

Felix Klotzsche[1,3] (klotzsche@cbs.mpg.de), Jeroen de Mooij[2], Sven Ohl[3], Michael Gaebler[1]; [1]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2]thefirstfloor.nl, Rotterdam, Netherlands, [3]Humboldt-Universität zu Berlin, Germany

Virtual reality (VR) is increasingly used in vision science experiments, with recent headsets featuring built-in eye tracking proving particularly valuable for researchers. However, these headsets are not research-grade, can require time-intensive software development, and raise questions about the quality of their eye tracking data for scientific purposes. The open-source EDIA toolbox addresses these challenges by supporting vision scientists to design and conduct experiments with eye tracking in VR, using the Unity game engine. A key feature of EDIA is its integration with multiple commercially available VR headsets. This enables user-friendly data extraction from built-in eye trackers at their native temporal resolution, all accessible via a standardized interface. EDIA provides tools for logging, visualizing, and streaming (eye tracking) data, and allows experimenters to easily build reusable components. This flexibility facilitates hardware switching and reusing code across experiments and setups. EDIA supports data streaming via the LabStreamingLayer network protocol, enabling synchronization with external data sources such as EEG. Additionally, it offers remote monitoring and control of experimental applications, a feature particularly relevant for mobile VR setups. For experiments implemented using EDIA, researchers can alter relevant study parameters (e.g., the design matrix) using platform-independent configuration files. This allows researchers and their assistants to customize experimental sessions without modifying Unity code. A validation module enables researchers to monitor and compare eye tracking data quality throughout an experiment, across studies, and between different devices. We demonstrate EDIA in a study evaluating eye tracking performance in five state-of-the-art VR headsets: three mobile models (Meta Quest Pro, PICO 4 Enterprise, HTC Vive Focus 3) and two tethered devices (HTC Vive Pro Eye, Varjo Aero). Using an empirical sample of 24 participants, we compare the headsets' spatial accuracy, precision, and latency (relative to concurrent electrooculography) under conditions with and without head movements.

This research was supported by the cooperation project between the Max Planck Society and the Fraunhofer Gesellschaft (grant: project NEUROHUM).

53.312 EYE-TRACKING DEVICES CAN BLOCK YOUR PERIPHERAL VISION AND IMPAIR PERFORMANCE: EVIDENCE FROM A FOOTBALL SPECIFIC PERIMETER TEST

Božo Vukojević[1] (bozo.vukojevic@gmail.com), Svitlana Pinchuk[1], Christian Vater[1]; [1]University of Bern, Switzerland

Peripheral vision is crucial in football, where players must monitor movements beyond their direct line of sight (Vater et al., 2019).

Traditional perimeter tests fail to be representative for the natural sports environment, leading to a gap in assessing peripheral vision in football. This study evaluates the performance of two eye-tracking systems — Pupil Labs Neon (Field Eye Tracker) and Pupil Labs Core (Lab Eye Tracker)—within a VR setup simulating a football-specific perimeter test. Twenty-two sport science students participated in a controlled, within-subject design experiment, where each completed the test using both eye-trackers in a counterbalanced order. As independent variable, the eccentricity at which peripheral players appear was systematically varied (10 – 110° eccentricity on each side; a full 220° span). VR environment was built using Unreal Engine 5. Key performance metrics included reaction time and accuracy in detecting peripheral stimuli. Additional variables of head orientation and body movements were controlled with a 14-camera Optitrack system as well as the Heart Rate Variability (HRV) indices with Movisens EcgMove4 device. It was predicted that the Pupil Labs Neon will have higher accuracy than Pupil Labs Core due to unobstructed design. No differences in reaction time or HRV were expected. Results showed that the Pupil Labs Neon significantly outperformed the Pupil Labs Core in terms of accuracy ($F(3,60)=71.68$, $p<.001$), with a large effect size $\eta^2 = 0.582$, indicating a notable decline in accuracy as eccentricity increased, particularly for stimuli beyond 80 degrees of eccentricity ($p<.01$), with no differences at lower degrees of eccentricity. No significant differences in reaction time or HRV metrics were observed between the two systems. These findings suggest that the design of the eye trackers, which minimizes peripheral obstructions, enhances accuracy in football-specific tasks, making it a superior choice for sports applications requiring peripheral vision assessments.

Project funded by Swiss National Science Foundation (Grant Number 211220)

53.313 THE ROLE OF TORSION DURING SELF-MOTION

Andrés H. Méndez[1] (andres.mendez@ub.edu), Cristina de la Malla[1], Joan López-Moliner[1]; [1]Vision and Control of Action Group, Institute of Neurosciences, Universitat de Barcelona, Catalonia, Spain

Rigorously tracking eye and head behavior in space is key for constructing realistic models of the stimuli that reaches our eyes and elucidating the neural mechanisms underlying visually guided behaviors. The 3D orientation and movement of the eyes in the head and the head in the world not only determine the line of sight but also generate the moment-to-moment patterns of visual flow that stimulate the retina. While horizontal and vertical eye movements have been characterized in real settings, eye torsion remains largely unexplored outside the lab, and its relevance for visual perception has often been overlooked. Here, we leveraged head-mounted technology to measure torsional eye movements during locomotion. We asked ten subjects to wear a head-mounted device in a static condition and while walking and fixating a distant target. We combined manual with automatic coding to extract eye torsion and compared eye azimuth, elevation and torsion with head yaw, pitch and roll, respectively. We then applied optic flow algorithms to the raw head-centered and corrected retina-centered videos to describe the consequences of eye counter-roll in the retinal flow. We show that eye torsion effectively compensates for head roll during locomotion, altering the incoming visual flow in ways that could be relevant for the extraction of self-motion parameters.

This work was supported by grant PID2023-150081NB-I00 to JLM and PID2023-150883NB-I00 and CNS2022-135808 to CM, funded by MICIU/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR. AM was supported by grant PRE2021-097688 funded by MICIU/AEI/10.13039/501100011033 and by the FSE+

Decision Making: Metacognition

Tuesday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

53.314 CONFIDENCE INCREASES WITH NUMBER OF ALTERNATIVES IN A DIGIT CATEGORIZATION TASK WHEN PERFORMANCE IS MATCHED

Yunxuan Zheng[1] (yzheng447@gatech.edu), Nadia Haddara[1], Dobromir Rahnev[1]; [1]Georgia Tech

Metacognition, the ability to evaluate one's own decision-making process, often manifests as confidence judgments. While numerous computational models have been developed to explain confidence in perceptual decision making, most focus on 2-choice tasks. However, real-world decisions frequently involve multiple alternatives, and how confidence is computed in such contexts remains less understood. To address this gap, we examined confidence judgments across varying numbers of choice options. Two hundred participants performed a perceptual discrimination task on noisy images of hand-written digits from the MNIST dataset. Participants identified the digit from a given set of options and rated their confidence on a 4-point scale. The task included three conditions: a 2-choice condition (discriminating between digits 5 and 6), a 4-choice condition (discriminating among digits 5–8), and an 8-choice condition (discriminating among digits 1–8). Accuracy was approximately matched across conditions by using a 2-up-1-down staircase procedure that continuously adjusted the noise level of the images for each of the three conditions independently (mean ± SD: 0.66 ± 0.02, 0.63 ± 0.02, 0.64 ± 0.02 for 2-, 4-, and 8-choice, respectively). Despite matched accuracy, we found that reaction time (RT) increased with number of choices (mean ± SD: 0.76 ± 0.25, 0.94 ± 0.24, 1.03 ± 0.23; all pairwise p's < 0.001), an effect consistent with the notion that motor preparation may take longer in the presence of more choice alternatives. Critically, we also found that confidence ratings also increased as choice set size increased (mean ± SD: 1.88 ± 0.51, 2.21 ± 0.54, 2.43 ± 0.52; all pairwise p's < 0.001). These results demonstrate that confidence can be strongly influenced by the number of available choices and support the notion that understanding confidence requires that we move beyond simple 2-choice tasks.

53.315 DISTINCT EFFECTS OF STIMULUS RELIABILITY AND BOUNDARY DISTANCE ON CONFIDENCE AND ACCURACY IN PERCEPTUAL DECISION MAKING

Kai Xue[1] (kxue33@gatech.edu), Herrick Fung, Dobromir Rahnev; [1]Georgia Institute of Technology

Stimulus manipulations in perceptual decision-making tasks can be categorized into two broad categories: manipulating the quality of the sensory input (stimulus reliability) and manipulating the distance between the sensory feature of interest and the decision boundary (boundary distance). However, despite the fundamental difference between these two manipulations, it remains controversial whether they lead to dissociable behavioral effects. Here, we demonstrate robust dissociations between these manipulations across two large experiments (N = 78). In Experiment 1, subjects indicated the orientation of a Gabor patch relative to a decision boundary set at 45°, with stimulus reliability and boundary distance manipulated through contrast and the offset from 45°, respectively. Experiment 2 replicated this design using dot motion stimuli, where stimulus reliability and boundary distance were manipulated through motion coherence and the offset from 45°, respectively. Across the two experiments, we found that, the boundary distance manipulation, as compared to stimulus reliability, had a larger effect on accuracy but a smaller effect on confidence. This led to some of the strongest confidence-accuracy dissociations observed in perceptual tasks. Specifically, high-reliability/low-distance conditions produced much higher confidence but lower accuracy compared to low-reliability/high-distance conditions. We also found a qualitative dissociation in confidence patterns for correct and error trials. Specifically, boundary distance manipulations produced the "folded-X pattern" where easier conditions increased confidence for correct trials but decreased it for error trials. In contrast, stimulus reliability manipulations violated the folded-X pattern, such that where easier conditions increased confidence for both correct and error trials. These results demonstrate that manipulations of stimulus reliability and boundary distance have fundamentally different effects on the relationship between accuracy and confidence, suggesting that the mapping between confidence and accuracy depends critically on how task difficulty is manipulated.

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

53.316 VISUAL CONFIDENCE IS AN ONLINE PROCESS

Marshall L. Green[1], Dobromir Rahnev[1]; [1]Georgia Institute of Technology

A central question in visual metacognition is whether confidence is computed strictly after or together with the perceptual decision. On one hand, several prominent models such as 2DSD (Pleskac & Busemeyer, 2010) postulate that confidence is based exclusively on information accrued after the perceptual decision (i.e., confidence is a post-decision process). On the other hand, many other models posit that confidence is computed simultaneously with the perceptual decision (i.e., confidence is an online process). Here we aim to adjudicate between these two theories. We developed a novel task that required continuous report of both choice and confidence. Specifically, participants (N = 25) estimated the global motion direction of dot motion stimuli by centering a confidence interval on the perceived motion direction and provided confidence by adjusting the width of the confidence interval. We recorded angular error and confidence every 16 ms. We found that for the first 500 ms, participants exhibited chance performance and extremely low confidence. This initial period was followed by a rapid increase in both performance and confidence, with both reaching a plateau around 1.5-2 seconds after stimulus onset. Critically, the confidence-accuracy correlation was near zero during the first 500 ms but then rapidly increased at the same point in time that performance and confidence changed, suggesting that confidence did not exhibit any post-decision delay. This pattern was the same regardless of motion coherence, except that stimuli with higher coherence led to steeper increases in performance, confidence, and the confidence-accuracy correlation. These time courses were fit well by a circular drift-diffusion model (cDDM) where decision and confidence are simultaneously made based on a noisy 2D accumulator. Altogether, our results strongly suggest that confidence is an online process that is computed simultaneously with the perceptual decision and challenge models where confidence is a strictly post-decision process.

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

53.317 DID THAT ANGRY FACE CAPTURE MY ATTENTION? METACOGNITIVE MONITORING OF ATTENTIONAL BIAS IN ANXIETY

Cheongil Kim[1], Juyoen Hur[1], Jeong Hyeon Park[1], Sang Chul Chong[1]; [1]Yonsei University

Are anxious individuals aware that their attention is excessively captured by threat-related stimuli, such as an angry face? If so, how accurate is this awareness? Accurate attentional monitoring is crucial for anxious individuals to effectively control their maladaptive attentional bias toward threat, as it enables them to recognize whether and to what extent attentional control is necessary. However, despite extensive research on attentional bias and control, little is known about attentional monitoring in anxiety. Adopting a novel approach that involves average facial expression and attentional allocation judgments (N = 70 young adults), the present study investigates how accurately individuals monitor their attentional bias toward an angry face and whether this ability is associated with anxiety levels (i.e., state and trait anxiety scores) measured using the State-Trait Anxiety Inventory (STAI). We presented two faces with different intensities of anger (i.e., near neutral and full-blown-angry faces) simultaneously and asked participants to judge the average facial expression of the two faces and their relative attentional allocation between them (e.g., 39% vs. 61%). We first quantified the strength of attentional bias by assessing the extent of bias toward anger in the average facial expression judgment. Then, we measured attentional monitoring abilities based on how accurately participants' attentional allocation judgments corresponded to their attentional bias. Our findings demonstrate that individuals can monitor their attentional bias toward an angry face. However, anxious individuals tend to underestimate their attentional bias compared to non-anxious individuals, despite their intact ability to monitor trial-by-trial variations in attentional bias. This underestimation may lead to a misjudgment of the necessity for attentional control and could explain why anxious individuals exhibit impaired attentional control. This study provides a novel theoretical framework that incorporates attentional monitoring to understand attentional mechanisms in anxiety more comprehensively.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2022-NR070542).

53.318 METACOGNITION DURING ILLUSORY EVIDENCE MANIPULATION

Nora Bradford[1], Brian Maniscalco[1], Megan A. K. Peters[1]; [1]University of California, Irvine

Research on visual metacognitive judgments has shown that stronger evidence supporting one's choice usually leads to higher confidence judgments even in the presence of stronger evidence opposing the choice, dubbed a 'positive evidence bias'. Thus far, this bias has typically been demonstrated in visual tasks which manipulate the objective (physical) amount of evidence available to the observer. To investigate whether and how the positive evidence bias translates to perceived (illusory) positive evidence, we presented a version of the convex hull numerosity illusion, amplified by also manipulating total surface area, in a 2AFC task accompanied by post-decisional confidence ratings to 50 healthy human participants. On each trial, observers viewed two patches of 11-32 dots, where the ratios of the two patches' convex hulls ranged from 1:1 to 1:1.4. The ratios of total surface areas occupied by the dots in both patches covaried with the convex hull ratios. This predictably affected the perceived relative numerosity of the patches. We categorized trials as 'congruent' if the larger-convex hull patch was indeed the more numerous patch, and 'incongruent' otherwise. Participants' visual discrimination performance on congruent trials was higher than on incongruent trials, confirming that larger convex hulls did lead to higher perceived numerosity. Confidence covaried with perceived numerosity for congruent trials, but systematically deviated from this pattern in incongruent trials. These findings provide preliminary evidence that illusory shifts in perceived numerosity can affect type 1 judgments and type 2 judgments, but may do so in different ways. Our results suggest a powerful way forward for separating out the computations underlying visual type 1 performance and subjective confidence.

53.319 PROBABILISTIC DECODING REVEALS THE DYNAMICS OF SENSORY UNCERTAINTY

Jeffrey Nestor[1] (jnestor@bu.edu), Karen Tian[1], Angus Chapman[1], Jenny Motzer[1], Rachel Denison[1]; [1]Boston University

Motivation: Humans are constantly faced with uncertainty about the state of the world due to neural and environmental noise. One can compensate for sensory uncertainty by integrating prior information about a stimulus feature, but it is unclear how uncertainty representations unfold over time and dynamically incorporate prior information. Here we first sought to establish whether trial-by-trial uncertainty representations are decodable from electroencephalography (EEG) data. Then we leveraged uncertainty decoding to ask whether and how prior information sharpens sensory representations over time. Methods: We recorded EEG data while, on each trial, participants estimated the location of a low-contrast grating target and then reported their positional uncertainty by adjusting an arc to bet on the precision of their estimate. Targets were either preceded by a spatial cue which provided a Gaussian prior over target location or an uninformative neutral cue. We trained probabilistic decoders on EEG data which estimated a posterior probability distribution over target locations for each trial and each time point, allowing us to extract time-resolved location predictions as well as estimates of sensory uncertainty. Results: Behaviorally, reported uncertainty correlated across trials with position error. In neutral trials, EEG decoders predicted target location above chance starting ~150 ms after stimulus onset and continuing through the response period. Decoded uncertainty predicted decoding error both across time points and across trials. Importantly, we also observed trial-by-trial correlations between reported uncertainty and decoded uncertainty, demonstrating that uncertainty decoded from EEG meaningfully reflected sensory uncertainty. In spatially cued trials, decoded uncertainty was significantly reduced relative to neutral trials even before target onset, showing that the brain represents predictive spatial information in preparation for the target. Conclusion: Dynamic, behaviorally-relevant representations of sensory uncertainty can be decoded from EEG data and are shaped by prior information in an anticipatory fashion.

Startup funding from Boston University to R.D., National Defense Science and Engineering Graduate Fellowship to K.T.

53.320 QUANTIFYING METACOGNITION RELATIVE TO PERFORMANCE CAPACITY ACROSS THE VISUAL FIELD

Angela Shen[1] (angela.shen@uci.edu), Megan A. K. Peters[1,2]; [1]University of California, Irvine, [2]Program in Brain, Mind, & Consciousness, Canadian Institute for Advanced Research

Does visual metacognition accurately track performance differences due to visual field location? It seems not always: some evidence suggests observers can overestimate their performance in the visual periphery (perhaps especially for near-threshold stimuli) when performance is matched between central and peripheral stimulus locations. However, this effect has only been demonstrated across few performance levels or visual field locations. Here, we investigated the relationship between performance and confidence across a matched range of performance and at multiple stimulus levels and locations. In a 2AFC discrimination task, observers simultaneously reported the tilt (+/- 45 degrees) of a grating and their confidence on a 4-point scale. The grating appeared at one of four visual eccentricities (2, 4.67, 7.33, or 10 degrees of visual angle) on each trial, and one of eight polar angles (cardinal: 0°, 90°, 180°, 270°; intercardinal: 45°, 135°, 225°, 315°) in each block of trials. Thus, we recorded performance and confidence at 32 visual field locations. We staircased stimulus contrast at each eccentricity to obtain discrimination thresholds, selected four contrast levels around each threshold value, and fit psychometric functions to performance and to confidence ratings for each location. We applied a novel analytic approach – relative psychometric function (RPF) analysis (Maniscalco et al., submitted) – to compute mean confidence rating for each location and across a matched range of performance levels (stemming from the 5 stimulus contrast levels). The RPF method provides an intuitive summary statistic, the area under the RPF curve (RPF AUC), that quantifies the relationship between two psychological processes – here, performance and confidence. Results comparing RPF AUC across visual field locations suggest that the quantitative relationship between confidence and performance varies as a function of eccentricity and polar angle. These findings represent a significant step towards understanding polar angle-based and eccentricity-based asymmetry in performance-controlled subjective confidence.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC; to AS) and the Canadian Institute for Advanced Research (CIFAR, Fellowship in Brain, Mind, & Consciousness; to MAKP).

53.321 SKILLED BUT UNAWARE OF IT: A DISSOCIATION BETWEEN PERCEPTUAL DECISION-MAKING AND CONFIDENCE JUDGEMENT

Yiran Ge[1], Siyuan Cheng[1], Nihong Chen[1]; [1]Tsinghua University, Beijing, China

People can introspect on their performance through confidence ratings. However, this process is not flawless. The degree to which individuals' confidence ratings align with the accuracy of their own perceptual decisions reflects metacognitive efficiency. In this study, we trained subjects on a basic visual motion task. Before and after training, we asked them to report their confidence on a three-point scale following each perceptual decision. To quantify metacognitive efficiency, we computed the ratio of the slopes of psychometric functions associated with high and low confidence. Surprisingly, while five days of training (3000 trials in total) improved sensory coding and refined their decision-making process, metacognition did not keep pace with these improvements. A model (Boundy-Singer et al., 2023) assuming confidence as a noisy estimate of decision reliability indicates that subjects' uncertainty about their decision at the matched sensory just-noticeable-difference actually increased. Error-related neural signals in the dorsolateral prefrontal cortex and anterior cingulate cortex also declined for the trained motion stimuli. Together, our findings reveal neural substrates underlying perceptual learning and highlight the dissociation between perceptual decision-making and confidence judgment in the human brain.

This work was supported by MOST 2021ZD0203600 and NSFC 31971031, 31930053.

53.322 TEMPORAL ERROR MONITORING IN TEMPORAL ORDER JUDGEMENTS: SUPPORT FOR A METACOGNITIVE STEARC EFFECT

Tutku Öztel[1,2] (toztel17@ku.edu.tr), Martin Wiener[1], Fuat Balci[2,3]; [1]George Mason University, [2]Koc University, [3]University of Manitoba

Previous studies have demonstrated that human participants can keep track of the magnitude and direction of their trial-to-trial errors in temporal, spatial, and numerical estimates, referred to as "metric error monitoring". Notably, these prior studies investigated metric error monitoring in an explicit timing/counting context, in which subjects were instructed to attend to the metric of interest. However, error monitoring also requires implicit monitoring of the magnitude and direction of temporal mismatches between the experience of different stimuli, in which temporal information is not explicitly estimated or reproduced. To address this, in three experiments we investigated whether participants (n=25-39; 3 experiments) could monitor errors in a visual temporal order judgment (TOJ) task, where they judged which of the two consecutive stimuli appeared first while fixating on a central fixation cross during head fixation (~35cm distance). In an initial and replication experiment, stimuli appeared on the horizontal axis (left/right), whereas in Experiment 3 they were on the vertical axis (top/bottom), in order to test for the potential influence of stimulus alignment. Participants also reported their confidence regarding the accuracy of their TOJ on a trial-by-trial basis on a 1-3 scale. The results of all three experiments showed that the confidence judgments for correct responses increased and for incorrect responses decreased with longer absolute, compared to relative, point of subjective simultaneity (PSS) centered- stimulus onset asynchronies (SOA; where PSS-centered SOA = SOA - PSS). A more granular analysis showed that participants could only correctly monitor their errors for left-first (Experiment 1 and Experiment 2) / bottom-first (Experiment 3) presentation orders, suggesting a metacognitive spatial–temporal association of response codes (STEARC) effect. Overall, the observed results provide evidence for metacognitive awareness of implicit metric errors, which can be explained by the linguistic asymmetries between spatial and temporal descriptions, as asserted by Conceptual Metaphor Theory (CMT).

53.323 THE RELATIVE PSYCHOMETRIC FUNCTION: A GENERAL ANALYSIS FRAMEWORK FOR RELATING PSYCHOLOGICAL PROCESSES

Megan A. K. Peters[1,2,3,4,5,6], Olenka Graham Castaneda[1,2], Brian Odegaard[7], Jorge Morales[8,9], Sivananda Rajananda[2], Rachel N. Denison[10], Brian Maniscalco[1,2]; [1]Department of Cognitive Sciences, University of California Irvine, [2]Department of Bioengineering, University of California Riverside, [3]Department of Logic & Philosophy of Science, University of California Irvine, [4]Center for the Theoretical Behavioral Sciences, University of California Irvine, [5]Center for the Neurobiology of Learning and Memory, University of California Irvine, [6]Program in Brain, Mind, & Consciousness, Canadian Institute for Advanced Research, [7]Department of Psychology, University of Florida, [8]Department of Psychology, Northeastern University, [9]Department of Philosophy and Religion, Northeastern University, [10]Department of Psychological and Brain Sciences, Boston University

Psychophysics seeks to quantitatively characterize relationships between objective properties of the world and subjective properties of perception. However, traditional approaches investigate psychophysical dependencies of perception on stimulus properties on a case by case basis rather than seeking to identify quantitative relationships among these psychological processes themselves. This latter goal is particularly important when the processes in question likely depend on each other in some way, such as is the case for subjective experience and task performance: typically, stronger physical stimuli lead to better performance and stronger subjective experiences of clarity, vividness, or confidence. But is the relationship between performance and subjective experience fixed, or can it vary, e.g. by task or attentional demands? Such questions are key for better understanding psychological processes in general, and subjective experience in particular. Here, we develop and showcase a new psychophysical method designed to answer such questions: relative psychometric function (RPF) analysis, which characterizes the nonlinear psychometric relationships between psychological processes and how these relationships change under different circumstances (e.g. experimental manipulations). We demonstrate the advantages of RPF analysis using a sample dataset in which human subjects discriminated random dot kinematogram stimuli which varied in dot motion coherence and overall dot density (dots per visual degree), and rated confidence. RPF analysis revealed systematic changes in the relationship between performance and two subjective measures (confidence and metacognitive sensitivity) due to dot density and task design choices. While these empirical results are intriguing in their own right, they also show how RPF analysis can reveal changes in quantitative relationships between any two psychological measures: performance, vividness, clarity, reaction time, confidence, and more. To encourage the scientific community to use RPF analysis on their data, we also present our open-source RPF toolbox.

Templeton World Charity Foundation (0567, to M.A.K.P. and R.N.D.); Canadian Institute for Advanced Research (Fellowship in the Brain, Mind, & Consciousness Program, to M.A.K.P.).

53.324 PERIPHERAL OVERCONFIDENCE IN A SCENE CATEGORIZATION TASK

Nino Sharvashidze[1], Matteo Toscani[2], Matteo Valsecchi[1];
[1]University of Bologna, [2]Bournemouth University

Peripheral appearance is heavily influenced by top-down assumptions about the statistics of our visual environment. But does metaperception, i.e. visual confidence take this into account? Evidence is mixed, studies (e.g., Odegaard et al., 2018; Solovey et al., 2015) reported overconfidence in peripheral detection tasks, but Toscani et al. (2021) reported underconfidence when observers discriminated the orientation of a peripheral patch. Here we tested whether underconfidence is always associated with peripheral discrimination or if it is contingent on whether the specific task conforms to the constraints of peripheral vision. We used a task where peripheral vision performs well – rapid scene categorization (Larson & Loschky, 2009). Our 12 participants categorized scenes as Desert, Beach, Mountain or Forest in a two-interval four-alternative forced choice task. In each interval, they viewed the scene either only in the periphery (scotoma) or only in the fovea (window). Subsequently, they indicated the interval in which they were more confident in their judgement. Task difficulty was manipulated by varying the scotoma size, with larger scotoma corresponding to less peripheral information, or the window size, with larger window corresponding to more central information. Accuracy decreased with the increasing size of the scotoma and increased with the increasing size of the window. We computed the probability of higher confidence in the periphery as a function of performance difference between the two conditions. Participants' points of equal confidence were systematically shifted towards higher foveal perceptual performance, demonstrating overconfidence in the periphery. This suggests that changing the task from local orientation discrimination to global scene categorization, i.e. a task where peripheral vision outperforms foveal vision, reversed the metaperceptual bias. Periphery is suited for detecting objects and categorizing scene gist, but not for discriminating fine details or local features. Metacognitive judgements seem to follow these inherent capabilities and constraints of peripheral vision.

The study was funded by the Italian Ministry of University and Research under the PRIN 2022 programme, project number 2022HEKCWH.

Object Recognition: Models

Tuesday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

53.325 ASSAYING THE EFFECT OF RECENT SENSORY HISTORY ON OBJECT CATEGORIZATION VIA HUMAN PSYCHOPHYSICS AND COMPUTATIONAL MODELING

Lynn K. A. Sörensen[1], Michael J. Lee[1,2], James J. DiCarlo[1,2];
[1]McGovern Institute for Brain Research at MIT, [2]MIT Quest for Intelligence

Sensory history is thought to strongly influence object perception. While the field has image-computable models of how images map to behavioral reports, we lack a comparable understanding of how dynamic sequences affect object perception. Here, we take initial steps to address this gap: we measure how sensory history affects human categorization performance (online psychophysics, N=500). Using 300 naturalistic videos and a binary object detection task, we compared pre-cued categorization reports based on video clips (200-1600ms) that end at a particular target frame with reports on the same frame shown statically for 200 ms. Surprisingly, single-frame-based reports explained substantial behavioral variance, even in the longest clips, challenging the notion that object recognition heavily depends on sensory history. Still, longer sensory history reports increasingly differed from frame-based recognition, yielding performance increases suggestive of evidence accumulation over time. Next, we focused on what mechanisms might explain these effects of sensory history. We hypothesized that frame-based encoding (e.g., via the ventral visual stream) combined with downstream temporal integration mechanisms may account for the emerging differences with longer sensory history. To test specific instantiations of this hypothesis, we augmented a pre-trained artificial neural network with diverse temporal decoders, including max-pooling, mean integration, leaky-integrators, and recurrent architectures (RNNs, GRUs, LSTMs), each optimized for categorization on a separate set of videos (40 repetitions). Interestingly, unlike simpler decoders, we found that non-linear temporal decoders increasingly captured the unique behavioral variance emerging with extended sensory history. Still, compared to human frame-based reports, frame-based ANN predictions (without temporal decoders) proved much less powerful at explaining human behavior overall, highlighting weaknesses of current image-based encoding models. Leveraging powerful, rapid, frame-based inferences as a foundation, our results demonstrate how sensory history could enrich object recognition through dynamic temporal integration of high-level visual representations.

This work was funded in part by the National Science Foundation-STC [CCF-1231216] and [2124136] (JJD) and a DFG Walter-Benjamin Fellowship 547591872 (LKAS).

53.326 CAPTURING THE REPRESENTATIONAL DYNAMICS OF FACE PERCEPTION IN DEEP RECURRENT NEURAL NETWORKS

Hossein Adeli[1] (hossein.adelijelodar@gmail.com), Chase W. King[1], Nikolaus Kriegeskorte[1]; [1]Columbia University

In order to understand the space of potential neural mechanisms of primate visual recognition and how they unfold over time, we investigate the representational dynamics of recurrent convolutional neural networks (RCNNs). We explore a family of models with bottom-up and lateral connections that were optimized for face-identification and object-recognition tasks. Using representational similarity analysis

(RSA), we observed that only models that were trained for face identification showed a late-emerging prominent distinction of identities as seen in the monkey face patch AM. Interestingly, early model responses strongly separated the objects from the faces. These findings suggest that the dynamics of face recognition that emerges in a hierarchical recurrent neural network prioritizes category-level recognition at early stages (face detection), triggering later category-specific computations that enable individual-level recognition as observed in neurophysiological findings (face discrimination). Our results also show that models that were trained simultaneously on both face identification and object recognition were more likely to show the signature of mirror symmetric viewpoint tuning in their intermediate representations as has been reported for monkey face patch AL. We then examined the tuning properties of individual units in the last layer of our network across timesteps. After embedding the face/non-face objects in a multi-dimensional representational space, for each unit the tuning axes were determined as the direction in which the unit responses increased (measured separately for face and non-face object clusters). The model exhibited a change of alignment between the face and object axes with increasing steps, resembling a late emerging identity discrimination tuning that was recently observed in primate face patches. Taken together, these results give us a candidate mechanistic account of primate face perception. The model is consistent with evidence on individual unit tuning and population geometry, revealing how the visual system dynamically separates first categories and later identities.

Research reported in this publication was supported in part by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under award number [RF1NS128897].

53.327 DO DEEP NEURAL NETWORKS PERCEIVE CONTEXTUAL VISUAL ILLUSIONS?

Hojin Jang[1,2], Pawan Sinha[2]; [1]Department of Brain and Cognitive Engineering, Korea University, [2]Department of Brain and Cognitive Sciences, MIT

Through a combination of innate circuit mechanisms and visual experience, human visual perception comes to incorporate sensitivity to contextual cues, which provide important information for interpreting the environment. While these contextual influences enable accurate perception in complex and dynamic settings, they can also give rise to systematic biases under certain conditions, as observed in phenomena like visual illusions. Classic illusions such as the Delboeuf, Ebbinghaus, Ponzo, and Müller-Lyer vividly demonstrate how context alters the perception of relative size, often leading to significant misjudgments. Modern deep neural networks (DNNs), which have demonstrated remarkable success in emulating human perceptual behaviors, raise the following question: Can artificial vision systems, that incorporate some of the architectural properties of their biological counterparts, and are trained on natural imagery, also develop susceptibility to such illusions? Addressing this question is not straightforward, given the challenges of querying the perceptual ‘experience’ of these systems. To address this question, we employed neuroscience-inspired methodologies, including univariate spatiotopic analysis to assess neural responses at target locations and multivariate decoding analysis to examine representational patterns across network layers. Using ImageNet-trained neural network models, our preliminary results have revealed that while neural responses in DNNs scale predictably with physical size and remain unaffected by contextual cues, multivariate decoding shows that illusion-like effects emerge in deeper layers, aligning with human perceptual biases. These findings suggest that hierarchical processing architectures and extensive visual training may drive susceptibility to contextual illusions.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (RS-2024-00451866).

53.328 FEATURE ACCENTUATION ALONG THE ENCODING AXES OF IT NEURONS UNCOVERS HIDDEN DIFFERENCES IN MODEL-BRAIN ALIGNMENT

Jacob S. Prince[1] (jacob.samuel.prince@gmail.com), Binxu Wang[1,2,3], Akshay V. Jagadeesh[3], Thomas Fel[1,2], Emily Lo[1], George A. Alvarez[1], Margaret S. Livingstone[3], Talia Konkle[1,2]; [1]Harvard University, [2]Kempner Institute for the Study of Natural and Artificial Intelligence, [3]Harvard Medical School

While deep neural network (DNN) encoding models increasingly achieve high predictivity of neural responses to natural images, it remains unclear whether these scores indicate algorithmic or mechanistic alignment between models and neural systems. Here we introduce a novel paradigm for rigorously testing DNN encoding models based on how well they can control neural responses. As a case study, we consider a resnet-50 and an adversarially robust variant, whose encoding models of IT neural responses to natural images achieve nearly identical R2 predictivity. However, using an explainable AI (xAI) technique called feature accentuation, we found dramatic differences in these models’ ability to control neural responses. Specifically, for each neural site, we synthesized image sets predicted to parametrically drive neural activity along the encoding axes in the target model’s feature space, which critically relies on the hierarchical computations and mechanisms of the target model. We presented these accentuated stimuli to the same monkey under identical recording conditions the day after synthesis. In this test of "parametric control," we found that stimuli from the robust model achieved precise modulation of neural firing: responses reliably and predictably aligned with each feature level. In contrast, baseline resnet-derived stimuli showed far weaker parametric control. Qualitatively, the robust model accentuations enhanced cohesive object contours, such as face-like curvatures, whereas baseline accentuations predominantly altered textural features, such as fur-like patterns. These results highlight that adversarially robust training may naturally pressure learning of more brain-relevant features, compared to standard objectives. More broadly, these results show that models with similar encoding predictivity for natural images can be distinguished through targeted tests of fine-grained parametric control along the encoding axes, revealing that some models offer better controllability than others. By bridging neuroAI and xAI, this approach emphasizes mechanistic alignment as a key goal for linking DNNs and brains.

This research was supported by NSF CAREER BCS-1942438 (TK), and an NDSEG grant (JSP)

53.329 FOVEATED SENSING WITH KNN-CONVOLUTIONAL NEURAL NETWORKS

Nicholas M. Blauch[1] (nblauch@fas.harvard.edu), George A. Alvarez[1], Talia Konkle[1]; [1]Harvard University

Human vision prioritizes the center of gaze through spatially-variant retinal sampling, leading to magnification of the fovea in cortical visual maps. In contrast, deep neural network models (DNNs) almost always operate on spatially uniform inputs, a severe mismatch that limits their use in understanding the active and foveated nature of human vision. Some work has explored foveated sampling in DNNs, however, these methods have been forced to wrangle retinal samples into grid-like representations, sacrificing faithful cortical retinotopy and creating undesirable warped receptive field shapes that depend on eccentricity. Here, we offer an alternative approach by adapting the model architecture to enable realistic foveated encoding of visual space. First, we use a spatially-variant input sensor derived from the log polar map model, which links retinal sampling to cortical magnification (Schwartz, 1980), but does not produce grid-like images. To handle the sensor's outputs, we convert spatial kernels for convolution and pooling into k-nearest neighborhoods (KNNs) defined in pixel space, and generalize convolution to KNNs. Filters are learned in a canonical reference frame, and are spatially mapped into each neighborhood for perception. This approach allows us to build hierarchical KNN convolutional neural networks (KNN-CNNs) closely matched to their CNN counterparts. Architecturally, these models naturally exhibit realistic cortical retinotopy and desirable receptive field properties, such as exponentially increasing size and constant shape as a function of eccentricity. Training these models end-to-end over natural images, we find that they perform competitively with resource-matched CNNs trained on grid-like foveated images, and exhibit increasing performance with multiple fixations. Broadly, this model class offers a more biologically-aligned sampling of the visual world, enabling future computational work to model the active and spatial nature of human vision, with applications in understanding visual recognition, crowding, and visual search. Last, this approach holds promise in building more neurally mappable models.

This work was supported by NSF CRCNS 2309041.

53.330 POPULATION RECEPTIVE FIELD MODELING: METHODS, CHALLENGES, AND INSIGHTS

Garikoitz Lerma-Usabiaga[1,2] (garikoitz@gmail.com), Chris I. Baker[3], Noah C. Benson[4], Tessa Dekker[5,6], Serge O. Dumoulin[7,8,9,10], Justin Gardner[11], Kalanit Grill-Spector[11], Kendrick Kay[12], Tomas Knapen[7,9], Eline R. Kupers[11,12], Fernanda Lenita Ribeiro[13,14], D. Samuel Schwarzkopf[15,16], Brian Wandell[11], Jonathan Winawer[17], Christian Windischberger[18], David Linhardt[18]; [1]BCBL. Basque Center on Cognition, Brain and Language. Spain, [2]Ikerbasque, Basque Foundation for Science. Spain, [3]Laboratory of Brain & Cognition, National Institute of Mental Health, Bethesda, MD, USA, [4]eScience Institute, University of Washington, Seattle, Washington, USA, [5]Institute of Ophthalmology, University College London, United Kingdom, [6]Experimental Psychology, University College London, United Kingdom, [7]Spinoza Centre for Neuroimaging, Netherlands, [8]Netherlands Institute for Neuroscience, Royal Netherlands Academy of Sciences, Netherlands, [9]Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Netherlands, [10]Experimental Psychology, Utrecht University, Netherlands, [11]Department of Psychology & Wu Tsai Neurosciences Institute, Stanford University, CA, USA, [12]Center for Magnetic Resonance Research (CMRR), Department of Radiology, University of Minnesota, Minneapolis, USA, [13]School of Electrical Engineering and Computer Science, The University of Queensland; Brisbane QLD, Australia, [14]Department of Medicine, Justus-Liebig University Giessen; Giessen, Hessen, Germany, [15]School of Optometry & Vision Science, University of Auckland, New Zealand, [16]Experimental Psychology, University College London, United Kingdom, [17]Department of Psychology & Center for Neural Science, New York University, NY, USA, [18]High Field MR Center, Center for Medical Physics and Biomedical Engineering, Medical University of Vienna. Austria

Population receptive field (pRF) modeling is a quantitative technique for characterizing neural responses to visual stimuli. It is most commonly used to measure the spatial sensitivity (position and spatial extent in the visual field) of neural populations. Recent advances have significantly expanded its applications and accessibility: pRF solutions across many voxels can now used by deep neural networks to automate segmentation of visual areas; some pRF models are now used to estimate temporal or spatiotemporal tuning; and pRFs are used in clinical applications to track patient disease progression and monitor treatment. These developments make pRF modeling more powerful and accessible than ever before, but also highlight the need for collaborative efforts to establish standards and quality checks, and to share expertise. In this work, we introduce an open collaborative initiative concerning methods for pRF modeling. The goals of the initiative are to examine key aspects of pRF analysis, from data acquisition to model fitting, and suggest how methodological choices influence pRF estimates. We identify challenges to common implementation and interpretation and propose solutions based on collective experience. For non-pRF users, we aim to provide accessible explanations of the method, along with suggested guidelines and potential applications. For current pRF users, we discuss practical challenges that may arise during implementation and suggest strategies to address these issues. For method developers, we highlight areas that need improvement and propose possible directions for future research and validation of models. We will present our initiative's current progress, discuss our planned publication, and invite interested practitioners to collaborate with the group.

Due to character limits, funding acknowledgments will appear on the poster

53.331 QUANTIFYING THE SIMILARITY OF NEURAL REPRESENTATIONS USING DECISION VARIABLE CORRELATION

Yu Qian[1] (ericqian@utexas.edu), Wilson Geisler[1], Xuexin Wei[1]; [1]University of Texas at Austin

Previous studies have compared the representations of macaque visual cortex and deep vision-based neural networks. Intriguingly, while some suggest that their representations are highly similar, others argued the opposite. To investigate this question, we develop a method to quantify the trial-by-trial similarity of a neural network and the brain, leveraging the predicted behavior from their internal representations. Our technique is based on decision variable correlation (DVC). DVC was originally developed to infer how correlated the decision variables of two observers are based on their behavior in binary choice tasks. We generalize the method to deal with neural representations. The key idea is to first use an optimal linear classifier to convert the population activity into a decision variable, and then compute the Pearson correlation of the decision variables. To address the under-estimation of DVC caused by noise, we further developed a technique to normalize the estimate using the noise ceiling. We compared ours with a previous used method based on Cohen's kappa. We apply our method to study the similarity of monkey inferior temporal cortex (IT) and deep networks. We use a public dataset of IT collected when monkeys viewing images of objects. We find that the DVC between the monkeys are high. The DVC between the monkey and the networks is generally lower than that between the two monkeys. Interestingly, the similarity of different networks are not as large as previously reported. Additionally, the better performing neural networks appear to be less similar to monkeys, based on the inferred DVC. Our study provides a new way to evaluate the similarity of two neural representations based on their implied behavior under a linear readout. The technique is general and should be applicable to other datasets as well.

53.332 REPRESENTATIONAL GEOMETRY DYNAMICS IN NETWORKS AFTER LONG-RANGE MODULATORY FEEDBACK

Kexin Cindy Luo[1] (kexinluo@g.harvard.edu), George Alvarez[1], Talia Konkle[1]; [1]Harvard University

The human visual system relies on extensive long-range feedback circuitry, where feedforward and feedback connections iteratively refine interpretations through reentrant loops (Di Lollo, 2012). Inspired by this neuroanatomy, a recent model introduced long-range feedback pathways into a convolutional neural network, where late-stage feature channels learn how to influence early-stage channels, to support successful object classification (Konkle & Alvarez, 2023). The model operates in two passes—a feedforward pass, generating initial representations, and a modulated pass, where activations reflect both the feedforward and feedback-modulatory processing. While prior work focused on injecting an external goal signal into the model to leverage feedback connections for category-based attention, here we examine the representational dynamics of this model during its default operation, without any top-down goals. Specifically, we explored how the representational geometry of exemplars and categories changes in the modulated pass, relative to the feedforward pass. We analyzed activations from 100 randomly selected ImageNet categories (300 images each). Local representational structure was evaluated through cluster sizes and k-nearest neighbor analysis, while global representational structure was assessed via prototype shifts and pairwise distances within and between categories. We found that default feedback modulation induced notable changes in representational geometry: Category cluster sizes significantly reduce as exemplar embeddings move closer to category prototypes. Locally, more nearest neighbors fall within the same category, and within-category distances decrease, reflecting tighter clustering. Meanwhile, the distances between categories remain relatively stable. Finally, the larger the prototype shift, the greater the cluster shrinkage, indicating a relationship between internal cohesion and global repositioning. These findings suggest that fixed long-range feedback connections induce an automatic prototype effect in the representational geometry, compacting clusters within categories while preserving global structure. Broadly, these emergent feedback dynamics might naturally induce categorical processing effects by refining local representations without disrupting overall structure, improving downstream category-based task efficiency.

Kempner Institute Graduate Fellowship to Kexin Cindy Luo

53.333 SMALL-SCALE ADVERSARIAL PERTURBATIONS EXPOSE KEY DIFFERENCES BETWEEN ANN-BASED VISION ENCODING MODELS

Nikolas McNeal[1,3], Mainak Deb, Ratan Murty[2,3]; [1]Machine Learning, School of Mathematics, Georgia Tech, [2]Cognition and Brain Science, School of Psychology, Georgia Tech, [3]Center of Excellence in Computational Cognition, Georgia Tech

Artificial neural network (ANN)-based encoding models have emerged as powerful tools in vision neuroscience, offering unprecedented accuracy in predicting responses of neurons, voxels, population patterns, and visual behaviors. But how robust are these vision encoding models to subtle perturbations in stimulus inputs? Are some models better than others? Despite their predictive success, relatively little is known about the susceptibility of vision encoding models to small, targeted stimulus manipulations otherwise expected to leave neural responses unaffected. Here, we focused on a previously untested property of widely used vision encoding models: their susceptibility to targeted stimulus perturbations. To this end, we trained ANN-based encoding models for high-level visual regions using the Natural Scenes Dataset. Consistent with prior reports, all models exhibited high accuracy in predicting responses to held-out images (all $R > .46$, $P<.0001$). Next, we assessed their susceptibility to small-scale "adversarial attacks" ($\varepsilon = 3/255$), ensuring the image changes were imperceptible to the human eye. To our surprise, we found that all encoding models were highly sensitive to these small-scale adversarial attacks, often dramatically changing their response predictions for nearly identical images. We then asked whether adversarial sensitivity could help find more brain-aligned models. Our results showed that adversarial susceptibility discriminated between encoding models more effectively than prediction accuracy alone (normalized variance .002 for accuracy versus .025 for adversarial robustness). Finally, we explored strategies to improve model robustness to targeted noise. Training models specifically for adversarial robustness increased resistance to perturbations but reduced prediction accuracy on brain data. In contrast, using sparse feature-to-brain mappings improved robustness while preserving accuracy (up to 51% median percent change improvement). Together, these findings expose key vulnerabilities in current ANN-based encoding models, introduce adversarial sensitivity as a complementary evaluation metric, and offer new model-to-brain mapping strategies for balancing robustness and predictive accuracy in future vision models.

This work was funded by the NIH Pathway to Independence Award (R00EY032603), and a startup grant from Georgia Tech (to NARM)

53.334 TOWARDS HOLISTIC VISION IN DEEP NEURAL NETWORKS: DISENTANGLING LOCAL AND GLOBAL PROCESSING

Fenil R. Doshi[1,2] *(fenil_doshi@fas.harvard.edu), Thomas Fel*[1,2]*, Talia Konkle*[1,2]*, George A. Alvarez*[1,2]*;* [1]*Harvard University,* [2]*Kempner Institute for the Study of Natural and Artificial Intelligence*

Configural shape processing, the integration of parts into cohesive wholes, is a hallmark of human visual processing. However, current deep neural networks, even those exhibiting shape bias (Geirhos et al., 2019), struggle to capture this critical capacity due to reliance on spurious local features like texture and color (Baker & Elder, 2022). This texture-bias limits their ability to learn robust but more complex shape-based features (Shah et al., 2020). We propose a novel training routine that simultaneously trains two distinct network architectures: a target network, ConfigNet, a standard CNN with progressively expanding receptive fields, and an auxiliary network, BagNet, an all-convolutional architecture with a fixed, restricted receptive field enforcing local featural processing. We introduce Divergence–Variance–Covariance Loss (DVCL), a novel objective function that disentangles local and global processing across the two networks. DVCL enforces orthogonality between models' intermediate representations by decorrelating redundant features while preserving variance. We hypothesize that, because ConfigNet cannot rely on the local strategy employed by Bagnet, it must instead learn a more global encoding strategy. We test this on a colored MNIST dataset where both the shape and color perfectly predict the digit category. BagNet primarily relies on color for classification, while ConfigNet focuses on shape, showing greater robustness to color changes. After training, when colors are shuffled, decoupling them from digit identity, BagNet's performance drops to chance (10%), with misclassifications strongly predicted by shuffled color patterns. In contrast, ConfigNet maintains 42.6% accuracy, with a near-diagonal confusion matrix, showing reliance on shape. When shapes are removed, leaving only color, ConfigNet's accuracy drops to 21.67%, while BagNet achieves 80.17%. This double dissociation highlights ConfigNet's global, shape-based strategy and BagNet's local, color-based strategy. Our findings demonstrate DVCL's potential to steer networks toward global configural processing, fostering robust, shape-based representations and advancing vision architectures toward human-like perception.

NSF PAC COMP-COG 1946308 to GAA; NSF CAREER BCS-1942438 to TK; Kempner Institute Graduate Fellowship to FD

53.335 TRACING THE ORIGINS OF PRIVILEGED AXES IN ARTIFICIAL NEURAL NETWORKS

Sudhanshu Srivastava[1] *(sus021@ucsd.edu), Alex Williams*[2]*, Josh McDermott*[3]*, Nancy Kanwisher*[3]*, Meenakshi Khosla*[1]*;* [1]*UCSD,* [2]*NYU,* [3]*MIT*

High-level perceptual cortices exhibit privileged representational axes (Khosla et al., 2023), meaning that even among representations with equivalent information content and geometry, specific neural tunings are systematically favored, consistently across participants. But why? Privileged axes corresponding to selective responses to faces, places, food, etc. could plausibly reflect the biological importance of these categories to humans. But that account is challenged by the fact that ANNs that know nothing about the significance of these categories nonetheless contain class-selective neurons (Prince et al., 2023; Khosla et al., 2023; Zheng et al., 2024). Here we test the hypothesis that privileged axes enhance task performance under structural constraints, such as nonlinear activation functions, which disrupt rotational symmetry in networks. Using a novel loss function, we trained neural networks for CIFAR-10 classification while explicitly maximizing or minimizing alignment with the representational axes of a reference network. Alignment was quantified using SoftMatch (Khosla & Williams, 2023), a metric that considers both population-level and neuron-level representations. We evaluated two architectures—AlexNet and MyrtleCNN—training 180 models (100 for AlexNet, 80 for MyrtleCNN) across layers with alignment objectives. Each model was trained for 20 epochs. For every layer, networks trained to maximize alignment with the corresponding layer of the reference achieved significantly higher CIFAR-10 classification accuracy than those trained to minimize alignment ($p < 10^{-4}$ for each layer, independent samples t-tests, df = 18). The performance gap widened for deeper layers. These findings support the hypothesis that privileged axes facilitate better task performance under structural constraints like nonlinear activation functions. They highlight the theoretical plausibility that biological activation functions, akin to ReLUs, impart privileged axes to brains, offering a deeper understanding of why specific tuning functions are consistently observed across neural networks, both biological and artificial.

53.336 A UNIFIED COMPUTATIONAL FRAMEWORK FOR VISUAL DYSFUNCTIONS IN PSYCHOSIS

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

Psychosis profoundly affects perception through a wide array of aberrations, ranging from altered contrast sensitivity to impaired facial emotion recognition and vivid visual hallucinations. While these symptoms appear disparate, we demonstrate that they likely emerge from a single computational principle: excessive integration of learned priors with sensory evidence. By leveraging a novel framework that combines image-computable object recognition networks with a generative inference algorithm, we provide the first unified computational account that not only explains normal vision but also bridges the gap between normal perception and psychosis-like visual dysfunctions. Unlike previous approaches where separate models were developed for normal vision and psychosis, our framework uniquely shows how a model of normal vision can transition to psychosis-like states by modulating a single parameter. This unification highlights the continuity between healthy and disordered perceptual processes and provides a novel computational explanation for psychosis symptoms emerging from disruptions in mechanisms supporting robust perception. Our approach makes three key contributions. First, we establish a formal mathematical link between robust object recognition and generative scoring, demonstrating why recognition networks can exhibit generative properties. Second, we trained robust object recognition models on naturalistic stimuli and incorporated feedback mechanisms during inference to simulate the effects of excessive prior integration. Third, we show that a single parameter controlling the integration of learned priors can reproduce the full spectrum of documented visual aberrations in psychosis, including excessive contrast sensitivity, systematic impairments in facial emotion recognition, and the emergence of complex visual hallucinations from spontaneous activity. These findings suggest that perceptual aberrations in psychosis arise from an imbalance in the integration of sensory evidence and learned priors, where excessive reliance on priors disrupts perceptual processing. This computational framework bridges low-level and high-level perceptual deficits, advancing our understanding of visual dysfunctions in psychosis.

T.T. is supported by NIH 1K99EY035357-01. This work was also supported by the Kavli Foundation and Gatsby Charitable Foundation GAT3708.

Object Recognition: Visual preferences

Tuesday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

53.337 A NEURAL SIGNATURE FOR THE PERCEIVED BEAUTY OF VISUAL CATEGORIES

Daniel Kaiser[1,2], Philipp A. Schumann[1], Rico Stecher[1], Martin N. Hebart[2,3]; [1]Justus Liebig University Giessen, [2]Center for Mind, Brain and Behavior, Philipps University Marburg, Justus Liebig University Giessen, and Technical University Darmstadt, [3]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig

The category of an object influences whether or not we find it beautiful. Across many exemplars, some object categories are simply more beautiful than others: we like butterflies and roses more than maggots and potatoes. We hypothesized that such categorical differences in beauty judgments are related to differences in perceptual representations in the visual system. To test this hypothesis, we combined EEG (n=10) and fMRI (n=3) data from the THINGS initiative, which provide spatially and temporally resolved neural responses for 634 visual object categories (12 exemplars each) recorded during oddball detection tasks. We additionally collected beauty ratings from online observers (n=3750) for all object categories (6 exemplars each). To determine how brain responses across space and time covaried with beauty ratings, we performed a model-based EEG-fMRI fusion analysis. Specifically, representational similarity between object categories (collapsed across exemplars) in the EEG was modelled by representational similarity in the fMRI, while determining how much the model fit is predicted by the similarity in beauty ratings. Beauty ratings predicted rapid activations in early visual cortex, with best predictions shortly after 100ms. Predictions for higher-level visual activations were weaker but more temporally sustained. Critically, beauty ratings predicted visual responses even when other properties (e.g., naturalness, arousal, and pleasantness) were controlled for. Our findings suggest that category-level beauty is first assessed during perceptual analysis, even in the absence of a beauty-related task. Besides all semantic associations we have with these object categories, butterflies are thus liked more than maggots because of how they look.

D.K. is supported by the DFG (KA4683/5-1, KA4683/6-1, SFB/TRR 135), an ERC starting grant (ERC-2022-STG 101076057), and “The Adaptive Mind”, funded by the Hessian Ministry of Higher Education, Science, Research and Art.

53.338 A PSYCHOPHYSICAL LOOK AT STIMULI FOR MENTAL HEALTH RESEARCH

Frans Verstraten[1], Renèll Rodrigues[1], Guang Yang[1], David Yu[1], Reuben Rideaux[1], David Alais[1]; [1]The University of Sydney, School of Psychology

In the mental health domain, like substance abuse research, questionnaires are prone to be contaminated by (social desirability) bias. For that reason, there is a trend to replace traditional questionnaires by free-viewing paradigms where for example eye movements and pupil size are registered when looking at naturalistic stimuli. In the realm of (binge)drinking research several stimuli databases of bottles of alcoholic and non-alcoholic drinks have been developed and claim to have reliability and (construct) validity. However, often the saliency of the stimuli is not considered, other than that the products are photographed under the same conditions (e.g. light source, camera location, etc.). We try to find out if differences in the individual stimuli content can account for choice preferences/biases. We calculated the rms-contrast of the stimuli, where the background was not taken into consideration, for luminance (the standard deviation of luminance fluctuations around the mean) as well as the rms-contrast in each color channel (RGB). We ordered the stimuli in increasing rms-contrast over these 4 dimensions separately. In experiment 1, in each trial observers were presented with two stimuli from the database and asked to indicate the most salient one. This resulted in an individual preference order. A correlation approach shows that the rms contrasts have no or very limited predictive value for stimuli preference. However, if we take brightness and saturation into account by using the HSV coordinates (hue, saturation, value), there is a strong correlation between stimuli and choice preference. We next used a noise paradigm where one of 2 noise patterns had a hidden object and became visible when the noise decreased as a function of time. In that case the observer was strongly relying on the object's contours. The results show that stimuli in addiction research need a careful look at possible intervening stimuli characteristics.

53.339 CANONICAL FIELD APPROXIMATION: A METHOD FOR MAPPING PERCEPTUALLY PRIVILEGED VIEWPOINTS AROUND OBJECTS

Athanasios Bourganos[1], Dirk B. Walther[1]; [1]University of Toronto

Canonicalness is a psychological variable that represents viewing preference around a three-dimensional object. Object perspectives with higher canonicalness are perceptually privileged when compared to perspectives with lower canonicalness. Various methods have attempted to quantify canonicalness around objects, employing Likert ratings, best/worst view selection, and Thurstonian Case V scaling. Canonical field approximation (CFA) is an updated, robust method for measuring the canonicalness of a set of discrete perspectives around an object. Using CFA measurements, two-dimensional interpolation facilitates generation of an approximate canonical field around an object. We deployed the CFA methodology on two sets of participants (exploratory and replication) to 1. Validate the CFA method, 2. Attempt replication of observations from past canonicalness research, and 3. Generate and test behavioral hypotheses. Results suggest CFA is a valid method for measuring canonicalness, with high Spearman correlation between canonicalness scores and ordinal perspective ranks and high agreement between participants (within and between samples). Observations from past canonicalness rating studies, relating to observer agreement and rating variance, were replicated. Alongside validation of CFA, two behavioral predictions were generated and tested. Decision time for choosing the more canonical perspective in the 2-alternative forced choice task that underlies CFA is negatively correlated with the absolute difference between the canonicalness ratings of the two perspectives. This is true for both a participant's individual canonicalness ratings and the mean sample canonicalness ratings for predicting decision time, with almost equivalent model parameter estimates. We also predicted and observed an affordance effect on canonicalness, where right- and left-handed participants preferred opposite, mirror-image perspectives of graspable objects with one handle, with the handle oriented towards their dominant hand. This effect was overlooked in past research that employed lower resolution canonicalness rating methods. Ultimately, CFA is a robust canonicalness rating method that provides insights into object perception processes and behavior arising therefrom.

This work was supported by NSERC Discovery Grant (RGPIN-2020-04097) to DBW.

53.340 THE RELATIONSHIP BETWEEN VISUAL PERCEPTION, SUSTAINED ATTENTION, AND VISUAL OBJECT RECOGNITION IN CHILDREN.

Hellen Kyler[1], Gabrielle Shimko[1], Karin James[1]; [1]Indiana University

Multiple mechanisms may facilitate accurate visual object recognition, including visual perception skills, attention, and manual interactions with real 3D objects. Recent research on adults revealed that multisensory control of real 3D objects during encoding facilitated visual object recognition, relative to unisensory passive encoding, which consisted of a replayed 2D video of previous object interactions (Kyler & James, 2024). This multisensory active advantage on recognition has yet to be tested in children, nor related to more general visual perception or sustained attention ability. To further understand the processes involved in active learning, especially within an important developmental period of multisensory integration (Gori et al., 2008), we investigated the role of visual perception skills and sustained attention ability in predicting visual recognition of novel objects made of 3D geons. Eighty children aged 6-8-years-old explored 10 novel 3D objects in either a multisensory (visual-haptic) or unisensory (visual) exploratory condition, and were subsequently tested on a 2D old-new visual object recognition test. We also measured visual perception ability through the Berry VMI and sustained attention by the Kiddie Conner's Continuous Performance Test. Results revealed that sustained attention (detectability) did not predict visual recognition scores, while higher visual perception abilities predicted both accuracy and d', controlling for age and exploratory condition. Detectability predicted visual perception scores, which were inversely related. High visual perception abilities were specifically important for children in the unisensory visual exploration condition, suggesting an importance of visual skills for learning object structure via a computer screen. The finding that object learning was not impacted by children's sustained attention demonstrates that even without active control, 3D, or multisensory object information, the visual only exploratory condition may not exhibit significantly more attentional demands. These results inform research on the development of active learning, visual attention, and memory following real world and computer tasks.

This work was supported by the National Institutes of Health under award number T32HD007475.

Perceptual Organization: Aesthetics

Tuesday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

53.341 2X2 BAR GRAPH READING ACCURACY ROBUST TO LOW-PASS SPATIAL FREQUENCY FILTERING

Nicholas Fencl[1], Nestor Matthews[1]; [1]Denison University

Introduction: College science courses often require that students develop perceptual organization skills to identify statistical effects in 2x2 graphs. For any given 2x2 graph, identifying Factor A's main effect, Factor B's main effect, and the interaction effect entails three different perceptual organization tasks. Here, we tested whether these three distinct perceptual organization tasks differentially rely on a 2x2 bar graph's high-spatial frequency content. Method: Forty-five college aged students from an introductory psychology course viewed black and white 2x2 bar graphs. During training, students received instructions on classifying their randomly assigned target factor (Factor A, Factor B, or Interaction) as either statistically significant or not. During testing, students made their statistical-significance judgments on 2x2 graphs that differed randomly across trials in spatial frequency: full spectrum graphs vs low-pass filtered graphs. The black and white bars had sharp edges in full spectrum graphs, and blurry edges in low-pass filtered graphs. Stimuli from the two spatial frequency conditions had the same root-mean-square contrast. Results: Thirty-nine (86% of) students performed the task with greater-than-chance accuracy (binomial probability $p < 0.001$). Among those students, Wilcoxen signed rank tests for the Factor A, Factor B, and Interaction conditions each showed non-significant differences in accuracy (% correct) between the full spectrum graphs and the low-pass filtered graphs. Median accuracy across the six experimental conditions (three target factors x two spatial frequency conditions) ranged between 85.9 and 92 percent correct. Conclusion: Under the conditions tested here, our data suggest that 2x2 bar graph reading accuracy remains unaffected by the removal of high spatial frequency information. This finding appears to hold regardless of whether the perceptual organization task pertains to Factor A's main effect, Factor B's main effect, or the interaction. Visual mechanisms tuned to low spatial frequencies provide ample spatial information for accurate 2x2 bar graph reading.

Denison University Lisska Center for Intellectual Engagement

53.342 DECODING SENSORY VALUATION IN COMPLEX VISUAL IMAGES

Yang Chang[1], Denise Hsien Wu[1]; [1]National Central University, Taiwan

What makes certain images more pleasurable than others? While much research has focused on low- and mid-level visual features (e.g., symmetry, complexity), the impacts of high-level features (e.g., semantic properties) and information processing characteristics (e.g., perceptual organization) on sensory pleasure remains largely unexplored. Moreover, previous research often employed limited and oversimplified stimulus sets, failing to account for the multidimensional nature of human visual experience. These caveats have led to inconsistent findings across different studies, hindering a generalized view of the factors contributing to sensory value construction. In the present study, we selected nearly 7,000 complex images, including everyday photographs and paintings, from existing datasets. These images are both naturalistic and multidimensional, and have been rated based on various attributes (i.e., beauty, enjoyment, liking, and valence) relevant to pleasure experience. Leveraging on computational analysis and modeling with vision and language-based neural networks for feature extraction, we identified over 200 image features, categorized into three groups: visual statistics, information processing, and semantic properties. Based on these features, we built regression models to predict image pleasure ratings and examined the contribution from each feature category to the prediction. The results demonstrated that for photographs, semantic properties had the strongest influence on pleasure ratings, followed by information processing, while the order was reversed for paintings, with information processing exerting the strongest influence. Visual statistics had the smallest but still significant impacts on both image types. As for the ratings of valence, semantic properties showed a stronger effect than the other feature categories. Further analysis of semantic properties revealed that object identity (nouns) contributes more to sensory pleasure than object relation (verbs) does. In summary, these findings provide a comprehensive understanding of the features driving human preference for high-dimensional images, emphasizing the crucial role of conceptual representation and perceptual organization in shaping sensory valuation.

This work was funded by National Science and Technology Council, Taiwan. grant ID: NSTC111-2410-H-008-060-MY3

53.343 UNRAVELING AESTHETIC EXPERIENCE: NEURAL UNFOLDING OF MULTIPLE COGNITIVE PROCESSES USING REPRESENTATIONAL SIMILARITY ANALYSIS

Yi Lin[1] (yi.lin@kuleuven.be), Johan Wagemans[1], Hans Op de Beeck[1]; [1]KU Leuven

Paintings provide unique aesthetic experiences, yet how visual inputs are processed and how these experiences unfold in the brain remain unclear. To address this, we identified neural networks underlying seven predefined cognitive processes during aesthetic experiences: content recognition, style recognition, emotional valence recognition, perceptual fluency, pleasure, interest, and liking. These processes are thought to occur at distinct moments of the aesthetic experience. We conducted an fMRI study with 34 participants who rated 36 Western paintings—balanced across content, style, and liking, and selected to partially dissociate pleasure and interest—prior to scanning. During the scan, participants viewed the same paintings. Using representational similarity analysis (RSA), we created a fine-grained mapping between neural activity and cognitive processes by comparing representational geometries across both dimensions. Specifically, we integrated fMRI neural data with behavioral ratings (e.g., "how much a painting is liked") or ground-truth categories (e.g., whether a painting is human-focused). Regions with high correlations between neural and cognitive/behavioral data likely reflect their involvement in specific cognitive process networks. Preliminary results revealed distinct neural networks for different cognitive processes. Visual areas, particularly ventral stream regions, were strongly involved in content recognition. Perceptual fluency, reflecting ease of processing, was primarily localized in low-level visual areas. Similarly, style recognition networks were centered in the visual cortex. Emotional valence recognition engaged broader networks spanning temporal, parietal, and frontal regions, including the temporo-parieto-occipital junction and orbital frontal cortices. Although pleasure, interest, and liking are conceptually similar, RSA revealed distinct neural differences. Pleasure networks included the temporo-parieto-occipital junction and dorsolateral prefrontal cortices. The network for interest was mainly confined to the visual cortex, while liking engaged broader regions, including the superior parietal lobule and medial prefrontal cortex. In short, RSA allowed us to unravel the neural mechanisms underlying the complexities of vision-related aesthetic experiences.

53.344 MODELING HUMAN ATTRACTIVENESS JUDGMENTS: THE CONTRIBUTION OF RACE AND GENDER TO CNN PREDICTIONS

Pei-Xuan Luo[1] (sharonluo2024@gmail.com), Denise Hsien Wu[1], Erik Chihhung Chang[1]; [1]National Central University, Taiwan

Whether facial attractiveness is universal or subject to the viewers' own experience and demographic characteristics has been a key focus of research. While some studies demonstrated universal preferences such as higher attractiveness ratings for female than male faces regardless of demographical categories, other studies showed higher consistency of the attractiveness ratings from raters on own-race than other-race faces. Deep learning models, particularly convolutional neural networks (CNNs), have shown remarkable accuracy in predicting human ratings of facial attractiveness. Hence, the current study aims to examine how CNNs make attractiveness judgments on cross-category faces that they were not specifically trained for, offering insights into human facial attractiveness judgments. We fine-tuned two CNN models, GoogLeNet and ResNeXt-50, with facial images and their attractiveness ratings from human participants in the SCUT-FBP5500 (N=5500) and a custom (N=346) dataset. Cares were taken to balance race (Caucasian/East Asian) and gender (male/female) of faces for training. Five-fold cross-validation assessed model performance via Pearson's correlation coefficients between human ratings and CNN predictions. The GoogLeNet and ResNeXt-50 models trained with a specific demographic category achieved average correlations of 0.78 and 0.69, respectively. When applying these models to predict attractiveness of faces in the same or a different demographic category, both models showed higher prediction accuracy for within-category than cross-category faces, indicating a reliance on learned demographic-specific attributes or patterns. Overall, these results demonstrated that CNN models, trained to mimic human judgements, demonstrate higher sensitivity to within-category than between-category preferences when predicting facial attractiveness for novel faces. These findings corroborate the idea that demographic characteristics such as race and gender influence attractiveness judgments, aligning with the role of prior experience in shaping preferences.

This study was supported by grants provided by the National Science and Technology Council, Taiwan (NSTC 111-2410-H-008-060-MY3, NSTC 113-2410-H-008-073).

53.345 UNIVERSAL BEAUTY OF ABSTRACT ARTWORK PREDICTED BY ARTIFICIAL NEURAL NETWORK MODELS

Xinchi Yu[1] (xcyu@umd.edu), Pranav Raghavan[1], Shlok Amit Desai[1], Yan Bao[2,3], Ernst Pöppel[3,2], Weizhen Xie[1]; [1]University of Maryland, [2]Peking University, [3]Ludwig-Maximilians-Universität München

While beauty is often said to be in the eye of the beholder, certain artworks consistently evoke shared aesthetic appreciation across individuals. This raises the question: to what extent does beauty inherently reside within the artworks themselves, independent of personal preferences or prior knowledge? To address this, we investigated whether pixel-level features of abstract artworks could predict consistent aesthetic judgments across people. Using abstract ink blob paintings by LaoZhu, each comprising different color variations (red, green, and gray) of the same painting, we evaluated aesthetic judgments while strictly controlling for confounding factors such as object size, image complexity, and art style. In Study 1, 617 participants rated the beauty of these paintings. Remarkably, participants demonstrated high agreement in their judgments. Interestingly, Western participants (n = 422), who were less familiar with the ink blob style, showed comparable consistency to Eastern participants (n = 195), suggesting that prior knowledge does not fully explain these shared aesthetic judgments. In Study 2, we examined whether image memorability contributed to this consistency using a continuous recognition memory task in 314 participants. The results provided minimal evidence for this, indicating that beauty appreciation is distinct from memory-related processes. To probe further, we analyzed the paintings using a pre-trained deep neural network (DNN, VGG16). We found that activation patterns in intermediate layers with a lower computational processing cost, namely lower mean activation, predicted higher aesthetic ratings. Notably, these DNN features were unrelated to memorability, highlighting distinct mechanisms for aesthetic perception and memory. Overall, our findings suggest that the consistency in perceived beauty across people may emerge from reduced information processing costs of artworks. Guided by these crowdsourced and modeling results, future research with neural recordings may help reveal the neurobiological basis of visual aesthetics.

This study is made possible by funding support from the National Institute of Neurological Disorders and Stroke (R00NS126492, PI: W. X.).

53.346 CHARM IS ESSENTIAL FOR HUMAN PERCEPTION AND AESTHETICS, AND CAN BE IMPLEMENTED IN COMPUTER VISION TOO

Fatemeh Behrad[1] (fatemeh.behrad@kuleuven.be), Tinne Tuytelaars[1], Johan Wagemans[1]; [1]KU Leuven

Human perception is shaped by a complex interaction of bottom-up and top-down processes. Our visual system actively focuses on critical details, preserves spatial relationships, and interprets scenes at multiple scales. This process relies on selective attention and eye movements, which allow us to focus on essential elements while integrating them into cohesive scenes. All of these characteristics of the human visual system are equally important, but they are not (yet) incorporated into computer vision systems, in spite of claims to the contrary. The ability to balance fine details with a broader compositional awareness is key to our aesthetic appreciation as well. Computer vision and machine learning models typically use fixed-size inputs obtained through downscaling or cropping, which ignore the essential ways in which human brains handle large-scale inputs. This standard approach can lead to significant information loss, limiting models' sensitivity to detail and spatial organization, as well as its capacity for aesthetic assessment. We introduce CHARM, a novel approach inspired by principles of human visual perception that preserves Composition, High-resolution, Aspect Ratio, and Multiscale information for Vision Transformers (ViTs). CHARM enables ViTs to mimic the brain's selective processing strategy by retaining high-resolution details in important regions and downscaling less relevant areas. This approach avoids the need for cropping or altering the aspect ratio, providing the model with richer contextual and compositional cues that enhance its performance and generalizability in image aesthetic assessment. In experiments on multiple image aesthetic assessment datasets, CHARM achieves up to an 8.1% performance improvement. CHARM marks an advancement in ViTs' capability to process visual information in a way that mirrors human perceptual efficiency, emphasizing the importance of high-resolution details, compositional integrity, and multiscale processing in both human and artificial vision systems.

53.347 CURVED FOREGROUND ELEMENTS ENHANCE AESTHETIC APPEAL OF INTERIOR SPACES

Claudia Damiano[1], Erick G. Chuquichambi[2], Vasiliki Meletaki[3], Keaton Bruce[4], Na Wei[4], Martin Skov[5], Anjan Chatterjee[3], Dirk B. Walther[1]; [1]University of Toronto, [2]University of the Balearic Islands, [3]University of Pennsylvania, [4]Temple University, [5]Copenhagen University Hospital Hvidovre

Curvature is a prominent visual feature that influences aesthetic experiences across various contexts. Particularly within architectural design, understanding how curvature impacts perceptual and emotional responses can inform both the aesthetics field and practical architectural design applications. This study investigated the effects of background and foreground curvature on aesthetic judgements of interior architectural spaces. In an online study using a 2 (Background: curvy vs. angular) x 2 (Foreground: curvy vs. angular) within-subjects design, participants evaluated 56 digitally rendered indoor spaces on measures of liking, beauty, fascination, coherence, hominess, privateness, and time they would spend in the space. Emotional responses to each space were also assessed using a subset of words from the Positive and Negative Affect Scale (PANAS). Our findings revealed that spaces with curved elements were consistently rated higher across multiple dimensions. For liking, spaces with curvy background elements were rated significantly higher than angular ones, $F(1, 29) = 6.04$, $p = 0.02$. Spaces with curvy foreground elements were also rated higher, $F(1, 29) = 18.95$, $p < 0.001$. Similar effects were observed for beauty, fascination, hominess, and coherence. Interactions between background and foreground curvature highlighted the dominant role of foreground features in shaping aesthetic preferences. Curvy foregrounds consistently drove higher ratings, and their effects were further enhanced when paired with curvy backgrounds, suggesting that foreground curvature plays a leading role in influencing aesthetic responses. Analyses of the PANAS words demonstrated that curvy spaces elicited more positive affect (e.g., "uplifted") and less negative affect (e.g., "stressed"), supporting the biophilic design hypothesis that nature-like visual features, such as curvy elements, in architecture result in positive affective experiences. These results highlight the benefits of curvy elements in architectural design, and offer insights into how visual features of both background and foreground elements interact to shape aesthetic and emotional experiences.

53.348 ALL SET! SET EFFECTS ON SUBJECTIVE RATINGS OF VARIOUS STIMULUS ATTRIBUTES

Johan Wagemans[1] (johan.wagemans@kuleuven.be), Claudia Damiano[2], Eline Van Geert[1]; [1]KU Leuven, Department of Brain & Cognition, [2]University of Toronto, Department of Psychology

Context effects and sequential effects are ubiquitous and well-known in vision science. Lesser-known but also important are set effects: Perception of stimulus attributes is affected by the set within which the stimuli are embedded. This phenomenon has been studied as “Einstellung” in the early days of psychology but has been largely forgotten or ignored by contemporary vision scientists. Here we demonstrate the role of set effects on a wide range of ratings of various kinds of images. 1839 participants rated 810 stimuli (90 stimuli x 9 categories: fractals, patterns, textures, scenes, art photographs, paintings–some figurative, some abstract) on 5 dimensions with 7-point bipolar scales (complexity, order, interest, pleasure, beauty). The critical manipulation was to have 9 homogeneous blocks (1 block of 90 stimuli from 1 category only) and 9 heterogeneous blocks (1 block of 10 stimuli from each of the 9 categories), assigned to different groups of participants. Clear differences were obtained between these conditions: the variance of the ratings was always higher for the heterogeneous blocks than for the homogeneous blocks. For all rating scales, categories with lower ratings on average were scored lower in heterogeneous than in homogeneous blocks, and categories with higher ratings on average were scored higher in heterogeneous than in homogeneous blocks. Heterogeneous blocks thus allowed participants to differentiate the stimuli better, while homogeneous blocks induced regression to the mean. This effect was more pronounced for ratings of more subjective attributes (beauty, pleasure, interest) but still present for more objective attributes (order, complexity). Intercorrelations between the scales were systematically higher for heterogeneous blocks. Subjective complexity ratings correlated more strongly with objective complexity measures for more abstract stimulus categories, and these correlations were generally higher for homogeneous than for heterogeneous blocks. This study clearly demonstrates the presence of set effects on judgments of stimulus attributes.

This research has been funded by an ERC Advanced Grant (no. 101053925, GRAPPA) awarded to JW and an FWO postdoctoral fellowship (no. 1244825N) awarded to EVG.

53.349 WHO DREW THIS? CHILDREN APPRECIATE VISUAL STYLE DIFFERENTLY THAN ADULTS

Shari Liu[1], Chaz Firestone[1], Tal Boger[1]; [1]Johns Hopkins University

Perception often confronts us with the distinction between *content*—what something is—and *form*—how it appears or is represented. For example, the same letter may appear in different typefaces, the same tool may be made of different materials, and the same body may take on different poses. Perhaps the richest example of this distinction arises in visual art: When viewing a painting, for example, we can discern not only what is depicted (e.g., a mountain or a sunset) but also the *manner* in which it is depicted (e.g., an impressionist sketch or a realistic portrayal). What are the origins of our capacity to distinguish content and form? And how might this capacity change throughout development? Artistic style presents an intuitive way to pit content against form, making it a useful case study for these questions. Here, in 3 experiments, we introduced participants to artists who produced various scenes with distinct contents and styles (e.g., a mountain sketched with crayons vs. a beach rendered as a detailed comic). Participants then saw a critical third scene whose content matched one artist’s drawing but whose style matched the other, and were asked which artist produced this critical scene. Whereas adults attributed the critical scene to an artist based on style (responding, e.g., that the crayon artist produced the new crayon scene, even with differing content; Experiment 1), children aged 4-7 years behaved *oppositely*, attributing based on content (responding, e.g., that the mountain artist produced the mountain scene, even with differing style; Experiment 2). We also replicated this pattern on LookIt, an online platform for collecting developmental data (Experiment 3). This work supports two conclusions: (1) The capacity to distinguish content from form arises early; but (2) the way this capacity is applied shifts throughout development.

NSF BCS 2021053, NSF GRFP

53.350 BELIEF IN BEAUTY: SEEING BEAUTY EVEN **WITHOUT HAVING “SEEN”**

Chenxiao Guan[1] (chenxiaoguan@zju.edu.cn), Yongqi Li[1], Mowei Shen[1], Hui Chen[1]; [1]Zhejiang University

Beliefs influence how individuals perceive and evaluate aesthetic experiences, while aesthetic experiences can shape or reinforce beliefs. This interplay highlights the connection between cognitive and perceptual processes. While prior research often examines beliefs related to values or meanings, our study investigates how the belief of having "seen" something affects aesthetic evaluations and how robust this belief becomes once formed. In Experiment 1, participants were paired up and completed a competition task, responding quickly to artworks shown for 3s, followed by "win/lose" feedback. Later, they performed an unexpected recognition task, identifying previously seen artworks, and an aesthetic rating task for both displayed and novel artworks. Artworks were pre-balanced for initial aesthetic ratings from the Vienna Art Picture System. Results showed higher aesthetic ratings for artworks recognized as previously seen in the competition task, even if they had not been displayed. This suggests that when participants believed they had seen an artwork (even without actually having "seen" it), they gave it higher aesthetic evaluations. Experiments 2 and 3 altered the task order and introduced more competitive settings, replicating the findings. Experiment 4 extended the effect to a more general task with only single participant. To test the belief's durability, Experiment 5 separated the competition and recognition/rating tasks by 1-month duration, yet the effect remained. To examine how strong the belief is once it developed, Experiments 6 and 7 manipulated participants' recognition or rating responses by altering the labels for the response groups and then prompted them to give re-evaluations. These manipulations failed to alter the belief, demonstrating its robustness. Our findings show that the belief of having "seen" an artwork strongly influences aesthetic evaluations, persists over time, and resists external manipulation. This research underscores the powerful role of belief in shaping aesthetic experiences and highlights its cognitive and perceptual underpinnings.

This work was supported by grants from National Natural Science Foundation of China (No.32400854 awarded to C.G., No.32171046 awarded to H.C.), Priority-Funded Postdoctoral Research Project, Zhejiang Province (No.ZJ2023052 awarded to C.G.)

Perceptual Organization: Individual differences, events and relations

Tuesday, May 20, 8:30 am – 12:30 pm, Banyan Breezeway

53.351 VISION IN CHILDREN: AS CORRELATED AS WE THINK?

Sandali Liyanagoonawardena[1], Simona Adele Garobbio[1], Michael Herzog[1]; [1]École Polytechnique Fédérale de Lausanne

Visual tests are crucial to assess visual abilities in vision research. Usually, we implicitly assume that the performance in one visual test correlates with that of similar tests. However, despite good test re-test reliability, attempts to show strong correlations between visual tests have proven unsuccessful. Even more surprisingly, visual tests that showed clear-cut differences between young and old observers did not correlate strongly in the older population. These results are surprising because one would have expected that, for example, a mild lens clouding should affect all visual functions, leading to correlations. Yet, observations convey that inter-individual differences are increased during aging. Here, we asked whether children, whose brains are not fully matured, exhibit correlations. We tested 52 children in 8 visual tests and 10 illusions, over 3 visits which were 6 months apart. Each test was measured twice at each visit to examine test-retest reliability. We analyzed the correlations between tests within a single visit and the longitudinal variations in children's vision. Furthermore, we investigated the correlations between performance changes across tests and between age and individual performance. The correlations between the performance of most visual tests remained below 0.4 and failed to reach the moderate range of Spearman normative values. Similarly, insignificant correlations were shown between the performance in each test and the age at each visit. When analyzing the individual performance changes between the three visits, the latent difference scores of visual tests were largely uncorrelated. However, there were significant variances in the individual performance change over time. Our results suggest higher-than-expected inter-individual variability of visual assessments, questioning the validity of visual tests. We encourage the application of a battery of tests, where the tests are carefully chosen, to replace the conventional use of a single visual test.

53.352 MAPPING EARLY PERCEPTUAL PROCESSES IN AUTISM AND PSYCHOSIS: INSIGHTS FROM MAGNO-PARVOCELLULAR PROCESSING AND TEMPORAL CROWDING

Ahmad Abu-Akel[1,2], Roy Ramati[1,3], Shaden Zein Aldeen[1,3], Yaffa Yeshurun[1,3]; [1]School of Psychological Sciences, University of Haifa, [2]The Haifa Brain and Behavior Hub (HBB), University of Haifa, [3]The Institute of Information Processing and Decision Making, University of Haifa

Early perceptual processes provide a crucial foundation for understanding the intersection of autism and psychosis spectrum disorders (ASD and PSD, respectively), offering insights into shared and distinct mechanisms across these dimensions. We report findings from two complementary studies examining early- and mid-level perceptual processing in neurotypical adults, assessing autistic traits and psychosis proneness in tandem. The first study investigates magno-like and parvo-like processing, previously implicated in both ASD and PSD, using the pulsed- and steady-pedestal paradigms combined with luminance contrast discrimination. The second study focuses on temporal crowding, a phenomenon reflecting long-lasting inter-item temporal interference. Three-item sequences were presented centrally with varying SOA (200–400ms); the task required estimating the middle item's orientation and the errors were analyzed using a mixture-modeling approach. Together, these studies explore key perceptual aspects relevant to ASD and PSD while advancing our understanding of perceptual processes within the occipital lobe along a gradient of increasing complexity. Results from polynomial regression with response surface analysis in Study 1 (N=81) found no association between autistic traits or psychosis proneness and parvo-like or magno-like processing. Similar analyses in Study 2 assessed the relationship between autistic traits, psychosis proneness, and temporal crowding (SOA modifications of the model parameters). In the difficult version (N=82), greater bias toward autistic traits corresponded to a larger reduction in guessing rates with longer SOAs, while substitution errors with the preceding distractor decreased in participants with balanced autistic traits and psychosis proneness (e.g., high tendencies on both). In the easier version (N=94), where stimuli were more discernible, we observed a similar reduction in substitution errors for 'balanced' participants with the succeeding distractor. These findings highlight nuanced, task-dependent interactions between autistic traits and psychosis proneness, suggesting that balanced trait dimensions may optimize certain perceptual processes, with implications for understanding sensory integration deficits in ASD and PSD.

53.353 IMAGE-GENERATION MODELS EXHIBITED HIGHER PERCEPTUAL SIMILARITY BUT LOWER ORIGINALITY COMPARED TO HUMANS WHEN GENERATING FROM INCOMPLETE SHAPES

Yaxin Liu[1] (yl1913@georgetown.edu), Maxwell Kay[1], Kibum Moon[1], Roger Beaty[2], Adam Green[1]; [1]Georgetown University, [2]The Pennsylvania State University

Recent advances in generative artificial intelligence (AI) have positioned it as a black-box parallel to the human mind, enabling comparisons between artificial and human cognition. However, much remains unknown about how the creativity of these models compares to human creativity, a hallmark of intelligence. In particular, the ability to generate original outputs that transcend initial perceptual constraints or patterns (i.e., overcoming fixations) remains largely understudied. The present study assessed creative variability by comparing humans and image-generation models using the Multi-Trial Creative Ideation (MTCI) "incomplete shapes" task. In this task, participants generated original doodles by incorporating existing incomplete shapes and lines. Drawings from human participants were compared with outputs from two types of image-generation models: diffusion-based (i.e., Stable Diffusion Img2Img pipeline) and vision transformer-based (i.e., Dalle-E). Both models were configured to simulate line-drawing outputs and were uniformly prompted to generate doodles using the provided incomplete shapes. We analyzed the visual creativity and perceptual similarity of over 6000 drawings. Specifically, we computationally assessed the visual creativity of each drawing by using the Automated Drawing Assessment (AuDra) model, a convolutional neural network (CNN) trained on and validated against human ratings (Patterson et al., 2023). We also quantified the perceptual similarity of drawings using metrics such as the structural similarity index (SSIM) and perceptual hash. Pairwise perceptual similarity scores were averaged across image data. Our findings revealed that human drawings consistently scored higher in visual creativity, as assessed by AuDra, compared to AI drawings. Furthermore, AI drawings exhibited greater perceptual similarity, indicating higher homogeneity compared to those generated by humans. Interestingly, image-generation models were sensitive to changes in hyperparameters. These results provide insight into how humans and generative AI leverage visual inputs to produce creative outputs and may reflect distinct processes in assimilating and generating visual information.

53.354 VISUAL RELATION DETECTION IN HUMANS AND DEEP LEARNING MODELS

Hongjing Lu[1] (hongjing@ucla.edu), Tanaya Jog[2], Zehui Zhao[3], Shuhao Fu[4]; [1]UCLA

Humans can rapidly detect visual relations in images, but how the visual system represents these relations between objects remains unclear. We conducted two experiments to investigate the detection of spatial and agentic relations in both naturalistic and controlled settings. In Experiment 1, participants viewed a sequence of realistic images, each displayed for 67 ms, and performed a Go/No-Go task, pressing a button when detecting a target relation in an image. Participants detected agentic relations more accurately than spatial relations. In Experiment 2, we removed the complex backgrounds and used synthetic images containing two objects either spatially related (e.g., "on top of," "next to") or unrelated. On each trial, participants viewed a pair of images including the same objects in different relational configurations. The two images were separated by a mask image, and participants identified which image appeared first. Humans achieved above-chance performance even with image display durations as brief as 40 ms and reached a performance plateau at durations exceeding 100 ms. We examine whether deep learning models can account for human performance in relation detection. Four models were evaluated: two vision models (ResNet and Vision Transformer) without explicit relational representations, and two relation detection models with explicit relational embeddings—a closed-vocabulary model based on a fixed set of object concepts and an open-vocabulary model without constraints on object concepts. Models computed the similarity of image pairs for each trial. While the two vision models demonstrated some correlation with human performance, the relation model with an open vocabulary of object concepts showed the strongest correlation ($r=0.38$, $p<.001$). In contrast, the relation model with a closed vocabulary failed to account for human performance. These findings suggest that humans' sensitivity to detecting visual relations result from flexible representations of object and relational concepts, as exemplified by the open-vocabulary relational model.

NSF BCS 2142269

53.355 CAUSAL TRANSFER BEYOND FORCE: RAPID INFERENCE OF PROPERTY INHERITANCE FROM AGENT TO PATIENT IN CAUSAL EVENTS

David Schwitzgebel[1], Alon Hafri[2]; [1]École Normale Supérieure—Paris Sciences et Lettres, [2]University of Delaware

Many causal interactions involve the transfer of force from an Agent to a Patient, such as when one object strikes another, causing it to move. These events have been studied extensively in the context of Michottean launching, which has shown how causality shapes visual inferences about a Patient's spatiotemporal properties, including position, direction, and speed. However, causal events often involve more than just force transfer: a wet towel can moisten skin, an ice cube can cool a drink, or a paintbrush can impart its color to a surface. While we can reason about such "property transfers," does the mind make rapid visual inferences about them? Here, we asked whether participants are sensitive to the rapid transfer of a property unrelated to force: color. Participants completed a continuous sequence of trials involving causal launching events with color changes. One object (either the Agent or Patient) started as orange or purple, while the other object (the "probe") began as gray. Upon contact, the probe object changed to either orange or purple. Crucially, the probe's color change was independent of the other object's color. The task was go/no-go: respond only when the probe changed to the participant's pre-assigned target color (before timeout). Response times were fastest when the probe was the Patient, the Agent initially had the target color, and the probe turned that color. Thus, launching events induced a visual expectation that color is selectively "transferred" from Agent to Patient in causal events. This interaction persisted throughout the experiment, suggesting that it was insensitive to task familiarity. Follow-up experiments found that disrupting the timing of the transfer (e.g., adding a temporal gap at contact) disrupted this expectation. These results suggest that the visual system rapidly and automatically infers causal transfer of visual properties like color, extending beyond spatiotemporal properties linked to force.

53.356 THE MISSING SELF IN TIME: DURATION **REPRODUCTIONS DIVERGE WHEN USING THE "SELF"** AS A REFERENCE POINT

Drew Schoenfeld[1], Hee Yeon Im[1], Joan Danielle K. Ongchoco[1]; [1]The University of British Columbia

In order to act in the world — to be in the right place at the right time — visual processing must keep track of time. Psychological time though is malleable, with the same duration seeming longer or shorter, depending on external factors (e.g., how many events occurred) or internal processes (e.g., speed of information processing). But we do not only passively perceive time; we can also make things go faster or slower depending on whether we wait and let time pass, or move things along ourselves. This agency, or the role of the "self" in relation to time, has been often isolated through the following question: "If your

Wednesday meeting is moved forward by two days, when is the new meeting?" If you are *moving* toward the meeting (i.e., "ego-moving"), the answer is Friday. If you are *waiting* for the meeting to approach (i.e., "ego-stationary"), it's Monday. Here we ask whether and how these self-time perspectives change temporal experience. Observers saw an event — a disc flash on a screen. As in typical duration experiments, they reproduced time intervals (i.e., the time elapsed between the start of the trial and the event) via button press. Critically, they also reproduced intervals with their "self" as the reference point (i.e., the time elapsed between now [where you are in time] and the past event). In a large-scale study, ego-moving observers (who reported "Friday") reproduced shorter durations between themselves and the past event, than did ego-stationary observers (who reported "Monday") — while no difference was observed for reproductions of intervals between events independent of the "self." Thus, ego-moving people may be more "ready to act," so past events are represented as having occurred more recently in time — perhaps because these are still deemed relevant for impending future action.

This work was supported by the Hampton Research Grant awarded to J.D.K.O. and the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to H.Y.I.

53.357 THE INFLUENCE OF PHYSICAL STABILITY IN FIGURE-GROUND JUDGMENTS

Akshita Reddy Mavurapu[1] (am1904@usnh.edu), Manish Singh[2], Ömer Dağlar Tanrikulu[1]; [1]University of New Hampshire, [2]Rutgers University, New Brunswick

Humans can report the physical stability of objects with remarkable accuracy, highlighting the predictive power of visual perception (Cholewiak et al., 2013). In this study, we investigated whether physical stability influences figure-ground segmentation by manipulating the relative physical stability of two sets of regions. We created black-and-white repeated-region figure-ground stimuli to investigate how relative physical stability influences figure-ground assignment. The relative stability of adjacent regions was manipulated by using sine waves as boundaries, with variations in the sine waves' amplitude altering the height of the center of mass across the regions, thereby affecting their stability. The regions were carefully designed to eliminate biases from other well-known figure-ground cues, such as symmetry, convexity, and area while maintaining equal widths at the top and bottom to ensure that top-bottom polarity remained constant across stimuli (Hulleman & Humphreys, 2004). Participants viewed each stimulus for 300 ms and indicated whether they perceived the black or white region as the figure. Results showed that participants were more likely to choose regions with a lower center of mass as the figure. By varying the center of mass, we also indirectly altered the critical angle, which observers are known to be sensitive to when assessing an object's stability. The results suggest that stability can influence figure-ground assignment, even when controlling top-bottom polarity. The findings demonstrate that stability influences figure-ground perception, revealing a potential visual processing bias favoring stable objects. While these results introduce stability as a novel cue to figure-ground organization, we propose that stability may represent a more fundamental bias underlying cues previously reported in the literature, such as top-bottom polarity. These cues could be interpreted as exceptional cases of a broader stability-based bias in visual processing.

TUESDAY MORNING POSTERS IN PAVILION

Eye Movements: Cognition

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.401 SEMANTIC INFORMATION SHAPES GAZE PATTERNS DURING NATURALISTIC MOVIE VIEWING

Sophie Su[1], Aditya Upadhyayula[1]; [1]Washington University in Saint Louis

Visual attention in naturalistic scene viewing is guided by high-level knowledge-based, and low-level saliency-based features. Recent work has begun to quantify what these knowledge-based effects are in the context of naturalistic images (Henderson & Hayes, 2017; 2019; 2023). Here we used a state-of-the-art transformer vision model to show that eye movement patterns in naturalistic videos are informed by the underlying semantic knowledge. Participants in our experiment were eye-tracked as they watched videos of everyday activities. Subsequently, the same videos were input to the OpenAI's CLIP transformer model as individual frames to generate an embedding for each frame. Linear regression models using these embeddings as input were further trained to predict the gaze heatmap obtained from (n=101) participants. These models were then tested on unseen gaze data to compute the correlation between the predicted and the observed gaze distributions. To quantify the effect of semantic knowledge in gaze prediction, we also repeated the same procedure with the CLIP embeddings of inverted scenes. Prior work has shown that scene inversion disrupts semantic processing while preserving low-level features (Shore & Klein, 2000). We, therefore, hypothesized that if the gaze prediction is driven by semantic knowledge during the movie viewing, there would be a significant difference in correlations between the model-generated and observed gaze distributions for both the intact and flipped video conditions. A linear mixed effects model showed a significant difference between the correlations for the intact and flipped video conditions (beta = 0.04, t= 6.116, $p < 0.001$). These results suggest that the gaze patterns during naturalistic movie viewing are informed by the underlying semantic knowledge. This work provides crucial groundwork for further exploration of gaze patterns informed by knowledge-based effects and their role in event cognition and memory.

53.402 EFFECTS OF DORSAL POSTERIOR PARIETAL CORTEX LESIONS ON SPATIAL- AND MOTOR-BASED INHIBITION

Julie Ouerfelli-Ethier[1,2], Tristan Jurkiewicz[2], Isabella Comtois Bona[1], Thomas Carrier[1], Aarlenne Z. Khan[1], Laure Pisella[2]; [1]University of Montreal, [2]Université Claude Bernard Lyon 1

Spatial and response inhibition are two different types of inhibition processes. Spatial inhibition refers to the suppression of a specific location whereas response inhibition involves the cancellation of a motor response according to changing contextual demands and is thus motor based. While these two types of inhibition are conceptually distinct, it is unclear if they recruit overlapping or distinct neuronal substrates. Previous findings have pointed to a role of the dorsal posterior parietal cortex (PPC) in spatial rather than response inhibition in an anti-saccade task, involving both types of inhibition. Here we examined the effects of lesions on the dorsal PPC on performance during two saccade tasks separately measuring spatial inhibition (inhibition of return (IOR) task) and response inhibition (stop signal task). To do so, we tested two optic ataxia patients, one unilateral and one bilateral, with lesions to the dorsal PPC, as well as 21 age-matched controls. For our spatial inhibition task, we found the typical IOR effect in our controls, however this was absent in our patients' ataxic hemifields. In contrast, we found no difference in performance between our patients and their respective controls on the stop signal task. These results confirm that motor-based inhibition is preserved following damage to the dorsal PPC, while spatial inhibition is impacted. Our results thus point to a specific role of the dorsal PPC in spatial inhibition, notably related to spatial attentional mechanisms.

This work was supported by the Natural Sciences and Engineer Research of Council of Canada, Unverisité de Lyon Idex mobility fund, Graduate and Postgraduate Studies and School of Optometry of University of Montreal, and Vision Health Research Network.

53.403 A METHODOLOGICAL STUDY: USING EYE TRACKING & COMPREHENSION QUESTIONS WITH L1 AND L2 SPANISH READERS

Julia Cerveira-Bianchi[1] (jcerveirabianch@uri.edu), LeAnne Spino-Seijas[1], Alisa Baron[1]; [1]University of Rhode Island

Eye tracking with a violation paradigm is an online, real-time measure, used within the field of second language acquisition to measure implicit or automatic knowledge. Native speakers (L1) demonstrate implicit knowledge and second language learners (L2) exhibit various levels of developing implicit knowledge, which serves as a measure of proficiency. In a violation paradigm, L1 and L2 speakers read grammatical and ungrammatical stimuli. The eye-tracking data provides insight into participants' sensitivity to grammatical violations. Previous studies have included comprehension questions, an offline measure, to ensure that participants read for meaning. However, eye tracking has rarely been applied to comprehension questions themselves, creating a methodological gap in the research. We fill this gap by investigating the relationships between eye-tracking measures, gaze duration and total duration, with accuracy and response time to comprehension questions. A total of 37 participants, L1 (N = 10) and L2 (N = 27) Spanish speakers, read grammatical and ungrammatical sentences in Spanish across two conditions, temporal reference and adverb placement. Participants read 69 sentences, each followed by a yes or no comprehension question which probed a non-critical region of the sentence, or the area without the grammatical error. The sentences were pseudorandomized to ensure that ungrammatical sentences of the same condition did not occur consecutively and that participants read only one version of all experimental stimuli. Our preliminary results show that L2 speakers exhibited significantly longer gaze duration and total duration than L1 speakers ($p = 0.027$ and $p = 0.025$ respectively). We also found a strong effect of sentence type, with adverb placement yielding a longer gaze duration than the temporal reference condition ($p < 0.001$). Finally, longer response times to comprehension questions were associated with a significant decrease in comprehension accuracy ($p < 0.001$), suggesting that slower responses are reliably linked to reduced comprehension accuracy.

Funding acknowledgment: This project was funded in part by two Project Completion Grants from the Division of Research & Economic Development at the University of Rhode Island (PI: Spino-Seijas; PI: Alisa Baron)

53.404 OCULOMOTOR CORRELATES OF HUMAN VISUAL CATEGORIZATION

Ali M Caron[1], Edward F Ester[1]; [1]University of Nevada, Reno

Categorization describes the process of assigning unique and behaviorally relevant labels to stimuli. Invasive electrophysiological recordings in non-human primates have demonstrated that retinotopically organized brain areas including the superior colliculus, lateral intraparietal area, and frontal eye fields exhibit robust encoding of learned visual categories and are causally involved in abstract category decisions. These same regions are causally involved in the generation and control of voluntary and involuntary eye movements (e.g., microsaccades), raising the possibility that oculomotor behavior can be used to track abstract category decisions. We tested this possibility by tracking gaze position after human volunteers had learned to classify sets of continuous orientation stimuli into discrete groups based on an arbitrary, experimenter-imposed boundary. Importantly, to-be-categorized stimuli were presented foveally and dynamically to minimize voluntary eye movements. In Experiment 1 (N = 25) volunteers reported the category of to-be-classified stimuli as quickly and as accurately as possible. Category information could be decoded from gaze coordinates beginning ~250 ms after stimulus onset, with decoding performance peaking immediately before participants' behavioral responses. Participants in Experiment 2 (N = 24) performed a delayed-match-to-category task that decoupled categorization from response selection by requiring participants to compare the category membership of sample and test stimuli separated by a blank interval. The category membership of the sample stimulus could be decoded from participants' gaze coordinates beginning ~250 ms after its appearance, and decoding performance remained well above chance across the ensuing delay period. These results demonstrate that human oculomotor behavior can be used to reliably track abstract category decisions.

Funding: National Science Foundation Grant 2050833

53.405 FEELING OF AGENCY IN VISUAL ACTIONS? NO EVIDENCE FOR EFFECT BINDING IN MICROSACCADES

Jan-Nikolas Klanke[1,2] (jan-nikolas.klanke@hu-berlin.de), Sven Ohl[1], Martin Rolfs[1,2]; [1]Humboldt-Universität zu Berlin, [2]Berlin School of Mind and Brain

Feeling of agency (FoA)—the experience of control over actions and their outcomes—has been extensively studied for bodily movements, but it remains unclear whether it extends to eye movements. Here we utilized small, ballistic eye movements called microsaccades to examine if such visual actions are characterized by FoA and whether intention mediates this feeling of control. We measured FoA for microsaccades using effect binding: a perceived compression between an action and its effect. In our experiments, we presented a vertically oriented grating, rendered invisible during stable fixation by a rapid temporal phase shift (>60 Hz) that became visible when its retinal motion was slowed by a microsaccade (active condition). The stimulus was presented embedded in a clock face and observers reported perceived stimulus timing in each trial. We compared the perceived timing of microsaccade-contingent stimuli to that of stimuli resulting from replaying a previous microsaccade's retinal consequences (replay condition). Trials without a stimulus were included as a control. To examine the role of intention, we tested this paradigm across two experiments: In Experiment 1, observers were instructed to either saccade or fixate, leading to intended (microsaccade instruction) and unintended (fixation instruction) microsaccades. In Experiment 2, no instruction was administered and microsaccades were labelled as spontaneous. Microsaccades rendered the stimulus perfectly visible, with visual sensitivity substantially higher in trials with generated or replayed microsaccades compared to those without, in particular when microsaccade direction and peak velocity matched direction and speed of the stimulus. Temporal estimates did not differ between active and replay conditions for any microsaccade type. Thus, we found little evidence for temporal binding of an eye movement to its sensory consequence. This data does not support that microsaccades are accompanied by a phenomenally distinct feeling of control, nor that FoA for small eye movements is facilitated by intention.

JNK was supported by the Berlin School of Mind and Brain, Humboldt-Universität zu Berlin. SO was supported by the DFG (OH 274/4-1) and its Heisenberg Programme (OH 274/5-1). MR was supported by the ERC under the Horizon 2020 program (grant agreement 865715), and the DFG (RO3579/8-1 and RO3579/10-1).

53.406 THE EFFECT OF TYPOGLYCEMIA ON READING COMPREHENSION AND EYE TRACKING

Ruggero Micheletto[1] (ruggero@yokohama-cu.ac.jp), Taichi Umikawa[1], Kotaro Oikawa[1], Keiko Tsuchiya[1]; [1]Yokohama City University

Character inversion, also known as “typoglycemia”, involves randomly swapping the letters within a word, while leaving the first and last letters of the word intact. We investigated the effect of reversing the hiragana characters in Japanese words and simultaneously examined the effect of this inversion on eye tracking behavior during reading. The experiment is performed by Japanese native speakers reading two passages containing words that are randomly affected by letter inversion. We used phrases of different complexity, an easier text extracted from a children's book and a more complex one from an adult novel. Regression analysis confirmed that, on average, the typoglycemia phenomenon is less probable than in the more complex text of the novel. Using an eye tracking system, we showed how the subjects skipped over the reversed words recording a heat map of eye movement. Also, analyzing the eye speed and direction of movement, we found that when the typoglycemya happens, the eye movements are faster than normal. Overall, this study shows how the complexity of a text affect the typoglycemia phenomenon, and correlate this with eye tracking behavior and eye movements speed. Our study is also confirming that people rely more heavily on the first and last letters than on the letters in the middle in order to gain reading comprehension even in Japanese characters context. Paradoxically, the analysis of reading speed suggests that letter reversals promote reading speed and text comprehension instead of hindering it.

We thank the Yokohama City University, "Basic Science 2024" grant

53.407 CONTRIBUTION OF EYE MOVEMENT PATTERNS TO REPRESENTATIONAL DRIFT

Yixin Yuan[1], Mikio Aoi[1], John Serences[1]; [1]UC San Diego

Representational drift, the phenomenon where neural activity patterns change in a systematic manner when exposed to the same stimulus, has been observed in mouse and human primary visual area (V1). However, given the complexity of movies and natural images, it's unknown whether the source of this change is intrinsic to neural computation or if animals and humans are covertly attending to a different set of features each time an image is presented. To investigate whether systematic changes in gaze or attention can contribute to representational drift, we ran a longitudinal eye-tracking study with complex visual stimuli repeated over time. 10 healthy adults (8F, 2M) participated in a free-viewing study over 6 sessions that spanned ~2 - 4 weeks. We tracked eye position while participants were presented with naturalistic images that repeated either within, across, or both within and across sessions. Fixation density distribution was estimated for each image to highlight locations that were most commonly foveated within each trial. Based on this density distribution, we summed the normalized fixation likelihoods across all presentations of a given image to form a saliency map. The information accumulation rate for each image was then quantified as the cumulative saliency values from fixation to fixation. This cumulative function can be compared across sessions to investigate differing temporal characteristics of gaze on the same image repeated over time. While there were substantial differences across individuals, when averaged across participants, images presented in earlier sessions had a faster information accumulation rate than those presented in later sessions. This indicates that gaze dynamics change systematically over sessions, suggesting one source of change that may contribute to representational drift. While speculative, these observations suggest that covert top-down cognitive factors such as shifts in attentional focus and eye position may bias neural encoding in a directional manner over time.

Face and Body Perception: Social cognition, behavioural

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.408 RELATIONAL GAZE INFORMATION PREDICTS HUMAN BEHAVIOR AND NEURAL RESPONSES TO COMPLEX SOCIAL SCENES

Wenshuo Qin[1], Manasi Malik[1], Leyla Isik[1]; [1]Johns Hopkins University

Studies have shown that understanding social events requires capturing interpersonal dynamics, which depends on tracking relational visual information between people. Gaze direction is a critical cue to relational information, and is prioritized behaviorally and represented in the superior temporal sulcus (STS), a region implicated in processing social scenes. However, there has been limited work integrating relational gaze information into computational models of social vision. Our current study evaluates simple computational models based on relational gaze features to match human judgments of social events, and compares their performance to state-of-the-art (SOTA) AI vision and language models. Specifically, we tested SocialGNN, a graph neural network that organizes each video frame into a graph structure based on gaze direction, alongside various recurrent neural networks (RNNs), on a dataset of behavior and neural responses to 3s videos of pairs of people engaged in everyday activities. Remarkably, SocialGNN demonstrates exceptional performance in predicting human behavioral ratings of social features, and achieves results on par with SOTA AI vision and language models, despite being trained on significantly less data and having only a fraction of the number of layers and tunable parameters. Notably, it also yields high neural encoding accuracy for STS responses to these videos. Follow-up ablation studies reveal that an even simpler RNN model trained on gaze direction alone is sufficient to achieve the observed alignment with human behavior and neural responses. This suggests that including additional visual features in the GNN framework does not further enhance its performance. These findings underscore the primacy of gaze direction as a relational visual cue in computational models for predicting human social judgments and brain responses. Future work will investigate what additional visual cues can be integrated with gaze direction in computational models to better model human social scene judgments.

53.409 THE EYE OF THE BEHOLDER: ONLY WHITE OBSERVERS HAVE AN ATTENTIONAL BIAS TOWARD LATINO AND SOUTH ASIAN FACES THAN OWN-RACE FACES

Xueqi Ao[1] (aox@mcmaster.ca), M.D. Rutherford[1]; [1]McMaster University

The psychological grouping of individuals by gender, ethnicity or religion influences attentional allocation. Attentional bias—prioritizing attention toward certain stimuli—has been shown to favor specific ethnic groups. However, prior research has predominantly focused on binary comparisons (e.g., Black vs. White faces) with primarily White participants, limiting our understanding of attentional biases across multicultural participant samples and stimulus sets. This study examines attentional biases among a racially diverse sample (N = 104) in response to faces from five ethnic groups: White, Black, East Asian, South Asian, and Latino, four faces each. Methods: Participants completed a dot-probe task, in which they first viewed a fixation cross on the screen for 500 ms, followed by the presentation of two face images, one on the left and one on the right side of the screen, for another 500 ms. These images displayed models' heads to shoulders and were positioned 6.5 degrees apart horizontally. After the facial images disappeared, a dot appeared at the location where the center of one of the two images had been. Participants were instructed to locate the dot. Reaction times (RTs) and accuracy were recorded. Results showed no overall attentional bias toward faces from minority racial groups across the full sample. However, when analyzed by participant ethnicity, White participants displayed a significant attentional bias, showing faster RTs to Latino and South Asian faces and a trend toward faster responses to Black faces compared to own-race faces. East Asian participants exhibited balanced RTs across ethnicities, with a slight preference for own-race faces, while South Asian participants showed a clear bias for own-race faces, responding faster to South Asian faces than to some outgroup faces. Conclusion: These findings show that attentional allocation varies significantly by observer ethnicity, contributing to a more nuanced understanding of attentional biases across racial and ethnic groups.

This research was funded by a Natural Sciences and Engineering Research Council grant to MDR

53.410 EVIDENCE OF DISTINCT RACE AND GENDER CATEGORIES IN FACE PERCEPTION

Nikita Agarwal[1], Tyler Yamato-Chang[1], Stella F. Lourenco[1]; [1]Emory University

Race and gender represent two foundational categories in face perception, shaping how people are recognized, understood, and remembered. When race and gender intersect, however, unique social and perceptual biases may emerge. The theory of 'intersectional invisibility' posits that individuals with intersecting marginalized identities—such as Black women—occupy a liminal space in cognition and perception, wherein they are more likely to be miscategorized than those with single-axis identities, such as Black men or White women. Here we examine the perceptual processes underlying intersectional invisibility, focusing on how race and gender interact in face categorization. Adult participants (N = 48) completed a triplet task—selecting the odd-one-out among three faces—while their choices and reaction times (RTs) were recorded. Participants represented four demographic groups: White men, White women, Black men and Black women (n = 12/group). Across 640 trials, participants compared 32 faces (White and Black, male and female) matched on age and attractiveness. Using non-metric multidimensional scaling (NMDS), we analyzed participants' choices and RTs. Analyses of participants' choices revealed highly reliable and distinct perceptual categories for both race and gender, with Black women forming their own distinct category—comparable to Black men, White men, and White women. NMDS of RTs, using 500 random starting configurations, yielded good fit to a two-dimensional solution (normalized stress = 0.066). Additionally, these analyses revealed that race was weighted more heavily than gender in participants' perceptual categories. More specifically, race emerged as the primary dimension, with gender categories varying to some extent by participant demographics. For example, Black women showed more distinct gender categories than White women. Altogether, these findings challenge claims of perceptual miscategorization for Black women. What is more, they suggest that intersectional invisibility may emerge from conceptual processes independent of perceptual mechanisms.

53.411 REPRESENTATIONS OF DYNAMIC FACIAL EXPRESSIONS ARE SHAPED BY BOTH EMOTIONAL AND SOCIAL FEATURES

Hilal Nizamoğlu[1,2], Fatma Celebi[1], Katharina Dobs[1,2]; [1]Justus Liebig University in Giessen, Germany, [2]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen and Darmstadt

Dynamic facial expressions play a crucial role in daily life by revealing emotions, social signals, and mental states of others. But what dimensions underlie humans' perception of these expressions? To address this question, we measured the perceived similarity of a large-scale set of both emotional and conversational facial expressions and used representational similarity analysis (RSA) to predict this similarity space. Participants (N=19) performed a multi-arrangement similarity judgment task on 48 videos of four actors performing 12 predefined expressions: six emotional (e.g., 'happily surprised'), and, crucially, six conversational (e.g., 'disagree'). Across the stimulus set, we quantified emotional dimensions (valence, arousal, affectiveness), social dimensions (social relevance, friendliness), and motion features (e.g., head and facial part movements) using independent ratings. We then used these features, along with predefined expression and actor identity to predict participants' similarity judgments. We found that all emotional and social dimensions, as well as expression, significantly predicted similarity judgments (FDR, $q < 0.01$), while motion features and identity had limited influence. Among the significant predictors, only valence, social relevance and expression uniquely contributed ($p < 0.001$) to the similarity judgments. These findings highlight that observers rely on both emotional and social dimensions to represent dynamic facial expressions.

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.

53.412 SENSITIVITY TO CUE CONSISTENCY IN MULTI-AGENT CONTEXTS: EFFECTS OF CUE TYPE AND GROUP SIZE

Jessica Savoie[1], Jelena Ristic[1]; [1]McGill University, Montreal, Canada

Humans spontaneously follow where others are looking, with this process in groups modulated by the number of consistent gaze cues. In groups of three, a minority of consistent gaze cues facilitates target responses; in groups of five, a majority of consistent cues is needed for similar response facilitation. Are such evaluative processes unique to gaze cues? We investigated this question in two preregistered experiments in which participants responded to targets cued by a group of three (E1, N=156) or five (E2, N=154) faces or visually matched directional arrows. Participants saw a group of cues at fixation and identified a peripheral target appearing on the left or right of the group. The target location could be cued by 0, 1, 2 or 3 cues in Experiment 1 or by 0, 1, 2, 3, 4 or 5 cues in Experiment 2. In Experiment 1, faster overall responses occurred in response to gaze relative to arrow cues, while reliable target facilitation tracked with increasing consistent cue numerosity for both cue types. In Experiment 2, target facilitation also tracked with increasing cue numerosity for both cue types with no overall differences in speed of responding for gaze and arrow cues. Interestingly, while for gaze cues, targets were reliably facilitated by one, two, and three consistent cues, there were no differences in responses between three and four consistent gaze cues. In contrast, for arrow cues, target facilitation followed a linear trend of reliable facilitation with increasing cue numerosity. These results show that while target responses can be facilitated by the minority of consistent biological and directional cues, responses are also modulated by group size and show nuanced effects across biological cue consistency increases. As such, these results highlight the sensitivity of human perceptual and attentional processes to the complexity of visual information in multi-agent contexts.

NSERC, FRQNT,SSHRC

53.413 EMOTIONAL MODULATION OF GAZE-CUING PROCEEDS IN ABSENCE OF LOWER FACE INFORMATION

Sarah McCrackin[1], Florence Mayrand[1], Jelena Ristic[1]; [1]McGill University

Human attention is spontaneously oriented in the direction of eye gaze. Such gaze following behavior is more pronounced if the person shifts their gaze and reacts emotionally with a facial expression, which is thought to be adaptive for facilitating orienting towards environmentally important events. It remains unknown, however, how the emotional expression on the face is processed along with eye-gaze information to produce the gaze following enhancement. Here we investigated this question by presenting participants with faces that averted their gaze and then either reacted emotionally or remained neutral. They responded to peripheral targets appearing in gaze-congruent or gaze-incongruent locations. Critically, half of the faces were unoccluded, while the other half had their lower part occluded by a surgical mask. Experiment 1 (N=74) presented fear, happy, and neutral expressions. Experiment 2 (N=77) presented disgust, surprise, and neutral expressions (pre-registered: https://osf.io/8uzgf). Thus, facial expressions varied both in eye size (i.e., smaller for disgust, larger for surprise) and in whether their most emotionally diagnostic facial features were visible (e.g. wide eyes for fear remaining unoccluded, wrinkled nose for disgust occluded). In both experiments, participants were overall significantly faster to locate gaze-congruent compared to gaze-incongruent targets, demonstrating classic gaze following behavior. This effect was larger when the faces displayed emotional expressions (fearful and happy vs. neutral in Experiment 1; disgusted and surprised vs. neutral in Experiment 2), but critically did not significantly vary with face occlusion condition. This shows that the information from the eye-region alone appears to contain enough emotional information within the face to drive the emotional enhancement of gaze following, highlighting the powerful nature of the eyes in emotional perception.

This research was supported by the Natural Sciences and Engineering Council of Canada (NSERC; JR & SM), and the G.W. Stairs Grant (SM & JR)

53.414 CUES DRIVING TRAIT IMPRESSIONS IN NATURALISTIC CONTEXTS ARE SPARSE

Ruoying Zheng[1] (r7zheng@ucsd.edu), Chujun Lin[1]; [1]University of California San Diego

Trait impressions are ubiquitous and shape consequential decisions. Prior work investigated what information people used for trait impressions using artificial designs. To advance a naturalistic understanding, we applied novel computational tools to quantify comprehensive cues based on prior theories (facial, bodily, clothing, environmental cues) (Study 1) and manipulate individual cues realistically (Study 2) in naturalistic images. Across two pre-registered studies (N1 = 2,435 U.S. representative; N2 = 569), we found that with rich information available, the cues predicting trait impressions were sparse. We confirmed for a subset of cues that these predictions were causal. Predictive cues carried unique information beyond the consistent information shared with other available cues. Unpredictive cues played a role by shaping the utilization of predictive cues through interactions. Together, our findings suggest that the mind may have evolved to utilize the naturalistic relations between cues to simplify what information to attend to when forming trait impressions.

53.415 CHASING THE SUPERNATURAL: THE PERCEPTION OF ANIMACY FROM MOTION IS RELATED TO SPIRITUAL EXPERIENCES

Dawei Bai[1], Kara Weisman[2], Eleanor Schille-Hudson[3], Elliott Ihm[4], Ann Taves[4], Tanya Marie Luhrmann[3], Brian Scholl[1]; [1]Yale University, [2]University of California, Riverside, [3]Stanford University, [4]University of California, Santa Barbara

When simple geometric shapes move about a display, what do we see? Beyond the obvious lower-level properties (e.g. direction, velocity), we may also spontaneously and even irresistibly perceive seemingly higher-level properties in certain displays -- such as one object *chasing* another. Here we asked whether such percepts might relate to a rather different form of phenomenology: the experience of supernatural or spiritual presences. If some people move through their daily lives perceiving higher degrees of agency, could this also manifest in more reports of sensing the presence of gods or other spiritual beings? Whereas some past work has theorized about such connections, here we isolate visual processing (vs. higher-level interpretation) by exploring correlations with objective visuomotor detection and performance. Subjects viewed displays filled with multiple identical moving discs. One (the 'sheep') was controlled by their own mouse movements, while another (the 'wolf') pursued the sheep. Subjects could only avoid being 'caught' by identifying the wolf's motion amidst the many distractors. Afterwards, subjects also completed two prominent and nuanced measures of spiritual experience (the 'Spiritual Events scale', and the 'Inventory of Non-Ordinary Experiences'). The results were striking: subjects who reported spiritual or supernatural experiences (via both measures) were worse at detecting chasing (and thus avoiding the wolf) -- and this difference in psychophysical functions could not be explained by either demographic variables (e.g. age, education) or more generic measures of religious belief or practice. We explain this by appeal to a type of 'hyperactive agency detection': people who perceive agency even in 'distractors' (and are therefore worse at detecting chasing) are more likely to report spiritual experiences. This work thus highlights a striking connection between a form of basic visual detection and some of the most powerful experiences in many people's lives.

This project was funded by a grant from the Templeton Foundation.

53.416 HUMAN RATINGS OF DISTINCTIVENESS ARE STABLE ACROSS IMAGE VARIATIONS

Isabelle Boutet[1], Ashlyne Rayne[1], Arda Erbayav[1], Alice 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). Many studies on face distinctiveness rely on human evaluations, typically of frontally-viewed faces (Johnston & Edmonds, 2009). In real-life interactions, faces are perceived from images that vary widely, including across different illumination and viewpoint conditions (Burton, 2013). To better reflect this reality, there is growing interest in understanding face perception across image variations. Research on social traits suggests that some evaluations are more consistent across image variations than others (e.g. Gogan et al., 2021; Todorov and Porter, 2014; Sutherland, Youngs, & Rhodes, 2017). We investigated whether distinctiveness evaluations are consistent across different image formats. Methods: Participants (N = 459; aged 17-35) rated a set of eight randomly assigned identities from a sample of 65 male and 64 female identities taken from the FIE database (de Oliveira Junior & Thomaz, 2006). Participants rated each identity across 3 image formats (half of the right profile and two front profiles, one under darker illumination). Faces were rated for distinctiveness (memorable, typical, common) and social traits (e.g., humble, interesting) (Bainbridge et al., 2013; Oosterhof & Todorov, 2008). Results: Distinctiveness ratings were consistent across image formats (r ranging from .63 to .84). For social traits, ratings were most consistent across image formats for attractiveness (all r = .90) followed by humble (all r = .89). Conclusion: Human ratings of distinctiveness are stable across image variations. This suggests that distinctiveness ratings are based on stable cues derived from the face itself, rather than from a combination of the face and specific image conditions. This study lays the foundation for future, more ecologically valid research, on the impact of distinctiveness on face recognition.

Funding: Supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant (2022-03998) to IB.

53.417 IDENTITY-BASED FACE TYPICALITY INFLUENCES THE NORM-BASED FACE SPACE OF FACE IDENTITY

Nitzan Guy[1] (guynitzan@gmail.com), Galit Yovel[1]; [1]Tel Aviv University

A well-established model of face identity posits that faces are represented as points in a multidimensional face space. According to a norm-based version of face space, faces are represented relative to a norm, with distinctive faces farther from the norm. Norm-based theories primarily focused on between-identity representations, where distinctiveness of each identity is evaluated relative to a norm, which is an average of different identity faces. These studies therefore overlooked how the rich representation of familiar faces influences the representation of face identity. In particular, it is unknown whether within-identity typicality – how typical an image of a particular familiar person – contributes to the norm-based face space representation of familiar faces. This study examines how within-identity typicality shapes norm-based face space representations and influences distinctiveness judgments of familiar faces. Our experiment included three rating tasks performed by different participants: (1) familiar distinctiveness: between-identity distinctiveness ratings ("how distinct the face is in a crowd of people"?) by participants familiar with the faces; (2) unfamiliar distinctiveness: between-identity distinctiveness ratings by unfamiliar participants; and (3) within-identity typicality: assessing how a/typical an image is of a specific familiar person. All three measures showed high split-half reliability scores. Results revealed that familiar distinctiveness ratings correlated positively with unfamiliar ratings, indicating shared between-identity perceptual representations. Additionally, images rated as more typical of a familiar identity were judged as more distinctive by participants familiar with them accounting for additional variance in between-identity distinctiveness rating beyond unfamiliar distinctiveness. Our study is the first to show the importance of within-identity representation of familiar faces to the norm-based face space. Within-identity variability is not only integral to representing familiar faces but also shapes between-identity distinctions. The findings have implications for face space frameworks of social perception and deep learning models of face recognition.

53.418 ATTENTIONAL MECHANISMS SHAPE THE RECOGNITION OF OWN- AND OTHER-RACE FACES

Chloé Galinier[1] (galc13@uqo.ca), Justin Duncan[1], Caroline Blais[1], Daniel Fiset[1]; [1]Université du Québec en Outaouais

The other-race effect refers to a recognition disadvantage for other-race compared to own-race faces. Although perceptual and social factors are well-studied, attentional mechanisms are often overlooked. Recent findings using behavioral measures (Duncan et al., VSS 2022) reveal that own-race faces, unlike other-race faces, are recognized automatically. To pinpoint this effect at the electrophysiological level, the exact same dual-task paradigm was used while recording EEG. Twenty-nine White participants completed a dual-task. In each block, they memorized a pair of White (own-race) or East-Asian (other-race) faces. They then performed a tone (Target 1; T1) categorization task followed by a delayed face (Target 2; T2) recognition task whereby they were asked whether T2 more closely matched the left or right face of the memory set. T1 and T2 presentations were separated by a stimulus onset asynchrony (SOA: 150, 300, 600, 1,200ms) to control potential conflict for central attention resources. T2 difficulty was varied by presenting either full signal faces (i.e., 0% Identity 1 - 100% Identity 2), or morphed faces (i.e., 60% Identity 1 - 40% Identity 2).Behavioral results confirmed previous findings (Duncan et al., VSS 2022), showing greater automatization for own-race versus other-race face recognition. For own-race faces, task difficulty effects diminished at shorter SOAs, suggesting cognitive slack absorption and perceptual handling of difficulty. Conversely, for other-race faces, difficulty effects remained consistent across SOAs, indicating post-perceptual processing. All electrophysiological components showed attentional modulation, with the P300 component revealing a marginal interaction between attentional modulation and race. Specifically, the P300 component, reflecting the allocation of limited-capacity attentional resources, showed higher amplitudes for own-race faces under shorter SOA conditions. These findings suggest that the P300 could be a potential locus for the behavioral effect, i.e., the automatization of own-race face recognition.

53.419 THE SELF-ADVANTAGE EFFECT ON THE ATTENTIONAL BIAS OF EMOTIONAL FACES

Jintong Bai[1], Yang Sun[1]; [1]College of Educational Sciences, Shenyang Normal University, Shenyang,110034, China, [2]College of Educational Sciences, Shenyang Normal University, Shenyang,110034, China

In the intricate tapestry of daily interactions, the recognition of facial expressions carries profound social importance. Yet, the mechanisms underpinning how we identify and interpret expressions of self and others remain elusive. To address this, our study employs a self-associative matching paradigm. Participants undergo associative learning by linking two distinct shapes (randomly chosen from circles, horizontal ellipses, and vertical ellipses) with emotional facial expressions (happiness, sadness, and neutrality) tied to two identity labels(self and others) and complete a key-press judgment task. Across three behavioral experiments, we investigate attentional biases toward emotional faces under varying identity conditions.In the first experiment, participants exhibited significantly shorter reaction times when judging self-faces compared to others' faces. Regardless of identity condition, happy expressions elicited the fastest responses.Building on this, the second experiment isolated emotional features by presenting only the eyes or the mouth, with emotions restricted to happiness and sadness. Participants responded faster to self-faces displaying happy mouths or sad eyes. For others' faces, a trend emerged suggesting quicker responses to happy mouths, highlighting potential differences in emotional processing between self and others.The third experiment introduced conflicting emotional cues by removing the influence of overall emotional congruence, such as happy eyes paired with sad mouths and vice versa.Results revealed that participants judged self-faces with sad eyes and happy mouths more quickly, whereas for others' faces, happy eyes and sad mouths elicited faster responses. This suggests distinct processing pathways for self-related and other-related emotional information. These findings highlight the mouth as a crucial region for self-related emotional expression, particularly in assessing happiness. Collectively, the results reinforce the self-advantage effect and provide empirical evidence for distinct cognitive mechanisms in facial expression recognition. These insights deepen our understanding of how individuals process emotional cues and set the stage for further exploration of attentional biases in social cognition.

The research was supported by 2020 Provincial Department of Education Basic Research Project for Colleges and Universities (LJKQR20210142022), Liaoning Province Education Science “14th Five-Year Plan”(JG21CB238).

Face and Body Perception: Social cognition, neural mechanisms

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.420 DISTINCT NEURAL RESPONSES FOR BODY MOVEMENT, INTERPERSONAL PHYSICAL INTERACTIONS, AND COMMUNICATIVE INTERACTIONS ALONG THE LATERAL VISUAL PATHWAY

Yuanfang Zhao[1] (distancejay@gmail.com), Emalie McMahon, Leyla Isik; [1]Johns Hopkins University

Dynamic social vision has been proposed to rely on the lateral visual cortex and superior temporal sulcus (STS), collectively referred to as the “third visual pathway.” However, disentangling distinct neural responses within this pathway has been challenging with hypothesis-driven approaches. To address this, we employed a data-driven voxel decomposition technique (i.e., non-negative matrix factorization) to decompose distinct neural responses in the third pathway in dynamic social vision. Using a large-scale fMRI dataset of 200 three-second video clips depicting two individuals in everyday activities, our analysis identified two components with distinct functional profiles shared across participants. One component, predominantly weighted in the extrastriate body area (EBA) of the lateral visual cortex, responds strongly to videos depicting physical interactions between people, while the other, heavily weighted in the STS, responds strongly to videos depicting communicative interactions. To ensure the robustness of these findings, we replicated the analysis with the BOLD Moments Dataset, analyzing neural responses to 601three-second video clips with one or more people engaging in a broader range of everyday activities. This replication not only confirmed the two components with similar functional profiles, but also revealed a third component, which responds strongly to videos depicting isolated single body movements. This third component is also weighted heavily in the EBA but has a distinct spatial distribution from the component representing interpersonal physical interactions, suggesting a distinction between single and multi-person physical actions. Importantly, none of the components can be explained by motion energy, suggesting that these responses are not driven simply by motion in the videos. Together, our findings suggest that EBA and STS within the third pathway differentially respond to body movement, physical interactions, and communicative interactions, offering new insights into the neural mechanisms underlying social interaction perception in the third visual pathway.

This work is supported by R01 grant: NIH R01MH132826

53.421 NEURAL CORRELATES OF SELF- AND OTHER-RELATED PROCESSING IN ADULTS WITH AND WITHOUT AUTISM

Annabel Nijhof[1], Jan R. Wiersema[1]; [1]Ghent University

Humans show enhanced (neural) responses to self-related stimuli, such as their own face or name. These enhanced responses are thought to be crucial for social functioning, and have been found to be diminished in individuals with autism. However, neural responses to self-related stimuli have rarely been investigated across different stimulus types. Further, it is still being debated precisely which aspects of self-related processing are altered in autism. Therefore, across two studies, we investigated the neural processes underlying own and other face and name processing, and how these may differ between adults with and without autism. In Study 1, 35 neurotypical adults in a 3T MRI scanner passively viewed six runs of face images, and six of visually presented first names, across three categories (Self, Close Other, Stranger). As an attention check, participants were asked to press a button when a fixation cross changed color. In Study 2, the paradigm was the same, but was this time tested in 25 adults with autism, and 24 without autism. For both studies, whole-brain analyses as well as searchlight decoding analyses were employed. Whole-brain analyses revealed stronger activation for familiar faces and names than for a stranger's face/name (but no self-specific enhancement) in several visual areas and in the ACC and precuneus. Searchlight decoding additionally revealed self-specific activation patterns in visual areas for both faces and names, and in the intraparietal sulcus and right STS for faces specifically. This pattern of results was replicated in Study 2. No group differences were found. The patterns of neural activation indicate involvement of modality-specific (i.e., visual) areas, but also areas of the social network, in processing familiar information. Self-specific responses are more subtle and appear stronger for one's own face than name. Finally, results show no differences in this basic form of self-related information processing in autism.

Annabel D. Nijhof was supported by an FWO fellowship (grant number FWO19/PDJ/025).

53.422 FACE RESPONSES IN AMYGDALA AND ENTORHINAL CORTEX ENCODE INTERPERSONAL RELATIONSHIPS WITHIN MACAQUE SOCIAL GROUPS

Ji Young Hwang[1] (jhwang48@jhu.edu), Allen Chen[1], Jerry Jin[1], Heejae Kim[1], Lydia M. Hopper[1], Charles E. Connor[1]; [1]Johns Hopkins University

Face information in ventral visual cortex is primarily structural and encodes face identities by their locations in a structural "face space". Face responsive regions in medial temporal lobe memory structures, including amygdala, associate specific identities with personal characteristics, such as aggressiveness (expression) and social rank. Here, we found that face responsive regions of amygdala and entorhinal cortex also encoded interpersonal relationships between individuals in stable social groups of 5–10 macaque monkeys. This naturally required that individual neurons encode information about multiple individuals, which depended on a radial coding format, centered on individual monkeys and rate encoding the relationship levels of other monkeys to that central individual. Thus, a given neuron might represent how submissive every other monkey was to monkey X. To characterize interpersonal relationships within four macaque social groups, we collected multi-camera surveillance videos in both indoor and outdoor runs across a three-month time frame. We analyzed these videos to measure frequencies of affiliative, aggressive, and submissive behaviors between all pairs of monkeys within each group. Two subject monkeys from the same group were studied with linear array probe recording in amygdala and adjacent 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. We found that many neurons in amygdala and entorhinal cortex encode social knowledge about interpersonal relationships involving either the subject monkey relating to other monkeys (self to other) or relationships not involving the self (other to other). In both cases, information about interpersonal relationships was represented in a radial coding format, in which the self or another monkey was the central node and the responses to other monkeys' faces correlated with a behavioral frequency relative to the central monkey.

53.423 AN EEG-FMRI INVESTIGATION OF THE SPATIOTEMPORAL HIERARCHY OF SOCIAL ACTIONS

Emalie McMahon[1,2], Elizabeth Jiwon Im[3], Michael F. Bonner[1], Leyla Isik[1]; [1]Johns Hopkins University, [2]Massachusetts Institute of Technology, [3]Stanford University

Recent work has argued that in addition to the dorsal and ventral visual streams, there is a third visual stream projecting laterally from early visual cortex to the superior temporal sulcus that is specialized for dynamic social content. A key characteristic of the dorsal and ventral streams is hierarchical computations. Here, we investigate whether the lateral visual stream also has hierarchical representations of social actions by combining EEG and fMRI data, which allows us to investigate the direction of information flow through lateral regions of the brain. Separate participants viewed the same videos while neural responses were recorded with EEG and fMRI. We first performed an EEG decoding analysis and found that the relative latency of decoding visual and social features from EEG progresses from low-level to abstract features, consistent with hierarchical processing. Next, we predicted the average response in functional ROIs from fMRI from the EEG signal across time. We find a short latency in predicting low-level regions and a longer latency in predicting all other regions, suggesting any early-late dissociation in information flow across these regions. Finally, using a novel EEG-fMRI encoding procedure, we track the time course of social action features across lateral stream ROIs. We find that, similar to the whole brain analysis, mid-level features come online in the lateral stream before high-level features of social actions. However, social action information is represented in mid- and high-level regions around the same time, suggesting the lateral pathway may not be organized in a strict feedforward hierarchy. Together these results provide novel insights into the neural processes that support dynamic, social vision.

This work was funded by NIMH R01MH132826 awarded to L.I.

53.424 HIGH-LEVEL VISUAL INFORMATION UNDERLIES SOCIAL AND LANGUAGE PROCESSING IN THE STS DURING NATURAL MOVIE VIEWING

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]Johns Hopkins School of Medicine

Real-world social perception depends on continuously integrating information from both vision and language. However, most prior neuroimaging studies have studied vision and language separately, leaving open critical questions about how these distinct social signals are integrated in the human brain. To address this gap, we investigate how rich social visual and verbal semantic signals are processed simultaneously using controlled and naturalistic fMRI paradigms. Focusing on the superior temporal sulcus (STS), previously shown to be sensitive to both visual social and language signals, we first localized visual social interaction perception and language regions in each participant (n=19) using controlled stimuli from prior work. We show for the first time that social interaction and language voxels in the STS are largely non-overlapping. We then investigate how these regions process a 45 minute naturalistic movie by combining vision (Alexnet) and language (sBERT) deep neural network embeddings with a voxel-wise encoding approach. We find that social interaction selective regions are best described by vision model embeddings of the frames of the movie and, to a lesser extent, language model embeddings of the spoken content. Surprisingly, language regions are equally well described by language and vision model embeddings, despite the lack of correlation between these features in the movie. Both regions were best explained by the last two layers of the vision model, suggesting sensitivity to high-level visual information. Follow-up analyses suggest that the most predictive vision model features are similar in social interaction and language regions, but different from the most predictive vision model features in low-level vision region MT. Together, these results suggest that social interaction and language-selective brain regions respond not only to spoken language content, but also to semantic information in the visual scene. This work highlights the importance of combining controlled and naturalistic approaches to study multimodal social processing.

53.425 A NOVEL FMRI DATASET TO STUDY THE NEURAL AND COMPUTATIONAL BASIS OF SOCIAL SCENE UNDERSTANDING

Manasi Malik[1] (mmalik16@jhu.edu), Shari Liu[1], Tianmin Shu[1], Leyla Isik[1]; [1]Johns Hopkins University

The ability to interpret social information from visual scenes is critical to human cognition, yet the neural computations underlying this ability remain poorly characterized. We present a novel fMRI dataset using procedurally generated stimuli to investigate these computations. We collected fMRI data from thirty participants as they watched animated videos from the PHASE dataset, depicting two agents interacting in ways that resemble real-life social behaviors. Participants rated the agents' relationships as "friendly," "neutral," or "adversarial." Participants also completed standard localizer tasks to identify brain regions associated with theory-of-mind, social interaction perception, and physical reasoning. Additionally, we collected individual social ratings for each video, Autism Spectrum Questionnaire, and demographic data. Our dataset offers two significant advantages. First, it provides a unique opportunity to compare neural data with computational models. Prior work has identified two theoretically distinct models that uniquely explain human social scene understanding - a bottom-up graph neural network based on visual information and a generative inverse planning model grounded in mental state inference. However, generative inverse planning models have rarely been compared to neural representations, largely because existing datasets lack stimuli designed with physical simulators. Our dataset addresses this limitation by using stimuli derived from a physical simulator, thus allowing for generative models to be built to work with them. This pairing enables unique comparisons between neural representations and both neural network-based and generative inverse planning models of social scene recognition. Second, the procedural generation of stimuli provides ground truth information about visual features (e.g., agent size, trajectories) and higher-level physical and social knowledge (e.g., agent goals, strength). This allows for systematic exploration of the role these features play in social scene understanding in the brain. This richly annotated fMRI dataset collected using procedurally designed stimuli, will advance our understanding of the neural basis of human social scene understanding.

This work was funded by NIMH R01MH132826 awarded to L.I.

Face and Body Perception: Features

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.426 RE-MODELING THE INVERTED FACE EFFECT FOR UNFAMILIAR FACES

Garrison Cottrell[1], Mrigankshi Kapoor[2]; [1]UCSD, [2]Computer Science and Engineering

Subjects perform poorly at recognizing upside-down faces. Previously, we presented an anatomically-inspired model with a foveated retina and the log-polar transform from the visual field to V1, followed by a standard CNN. The log-polar transformation causes changes in scale to appear as horizontal translations, while rotation in the image plane leads to vertical translations. When fed into a standard convnet, this provides rotation and scale invariance. However, because V1 is not a torus, for rotation, features "fall off" the top of the planar representation and reappear underneath, disrupting the configuration of the features, while preserving the features themselves, leading to the IFE. A standard CNN, fails to model the effect, being overly disrupted by inversion. Because the model was trained on these faces, this represents the IFE for familiar faces. For unfamiliar faces, in order to model the same effect, we created a simple memory model by storing noisy vector representations of the novel study faces from the penultimate layer of the network, without any weight updates. By adding noise at storage and/or inspection, the model will make a few errors on upright faces it has "studied" in this way. However, inverted faces also show small errors, as the representations still only differ slightly due to the noise. This failure to model the inversion effect on unfamiliar faces caused us to rethink the model. In order to bias the model towards upright faces, we trained the penultimate layer to be an attractor network for upright faces. Now, inverted faces with noise added do not match as well to novel inverted face activations after processing by the attractor network, restoring the inverted face effect. It is therefore important to consider dynamical representations of faces – in the form of attractor networks – in order to faithfully model the behavioral data.

This work was supported by NSF CRCNS grant #2208362

53.427 ORIENTATION TUNING OF FACE PROCESSING IN HUMAN V1

Mrittika Dey[1,2] (mrittika.dey@uclouvain.be), Jolien P Schuurmans[1,2], Valerie Goffaux[1,2]; [1]Psychological Sciences Research Institute, UCLouvain, [2]Institute of Neuroscience, UCLouvain

Human observers identify faces best based on horizontal cues. This horizontal tuning is disrupted when the face is inverted, highlighting the importance of horizontal information for the specialized processing of the upright face stimulus. Face-preferring regions like the fusiform face area (FFA) are selectively tuned to horizontal content of upright face information. Orientation selectivity is a fundamental characteristic of neural coding in the primary visual cortex (V1). It is also shown to interact recurrently with high-level visual regions to refine visual representations. The present fMRI study investigates whether high-level face-specialized mechanisms influence orientation-selective coding in V1. Using correlation, we compared orientation selectivity in V1 and FFA for upright, inverted, and scrambled faces, filtered in restricted orientation bands (0°,30°,60°,90°,120°,150°). In a separate identity-matching task, we measured face identification performance at the same orientations. For all stimulus types, V1 univariate response follows a negatively-curved profile dipping in the horizontal range, suggesting response suppression in this range. FFA response peaked in the horizontal range for upright faces only. This horizontal tuning was eliminated by inversion, mirroring face identification performance. The correlation between FFA and V1 responses also peaked in the horizontal range for upright faces only, suggesting an increased V1/FFA communication for the face-specialized processing of horizontal information. Identification behavior and V1 response correlated more negatively in the horizontal range and with inversion, suggesting a potential influence of V1 coding on the horizontal tuning of recognition behavior. Multivoxel pattern analysis (MVPA) revealed that stimulus class significantly impacts V1 orientation selectivity, as orientation decoding was best for cardinal orientations for faces but not scrambled images. These findings suggest that V1 orientation selectivity is influenced by downstream face-specialized processing stages. This supports the view of V1 as an active blackboard, updating its orientation-selective coding under the influence of higher-level visual regions to facilitate perceptual behavior.

53.428 DYNAMIC AND FLEXIBLE FEATURE ROUTING IN BRAIN PATHWAYS FOR DIFFERENT FACE PERCEPTIONS

Yuening Yan[1] (yuening.yan@glasgow.ac.uk), Jiayu Zhan[2], Hui Yu[1], Chen Zhou[1], Oliver G. B. Garrod[1], Robin A.A. Ince[1], Rachael E. Jack[1], Philippe G. Schyns[1]; [1]School of Psychology and Neuroscience, University of Glasgow, [2]School of Psychological and Cognitive Sciences, Peking University

Faces can be perceived differently based on distinct features: static 3D shape/complexion for identity (e.g. 'Mary') and transient movements (Action Units, AUs) for emotion (e.g. 'Happy'). However, how the brain dynamically routes these features for identity recognition and emotion categorizations remains unknown. Using a generative model of the human face, we independently manipulated 3Dshape/complexion of six face identities and the AU features of six basic emotions (happy, surprise, fear, disgust, anger, sad). Participants first learned to identify the six identities and classify the six emotions with 100% accuracy. In the neuroimaging phase, participants viewed 3,6000 facial animations and categorized them according to identity, emotion or both (identity+emotion; N = 8 participants per condition). We recorded their MEG/behavioral responses during the task. All participants viewed identical facial animations. Information-theoretical analyses revealed where and when MEG source amplitudes represented identity and emotion features (replicated across all individual participants, $p<0.05$, FWER-corrected): 1. Emotion: Social pathway selectively routed representations of dynamic emotions and their individual AUs laterally to Superior Temporal Gyrus, with task-irrelevant identities briefly represented in Occipital Cortex (OC) 2. Identity: Occipito-ventral pathway selectively routed static 3D identity representations to Inferior Temporal Gyrus, with task-irrelevant emotions briefly represented in OC 3. Dual-task: Identities and emotions were routed separately via ventral and social pathways, demonstrating task-specific flexibility in feature processing These findings show that the brain dynamically and flexibly routes specific static and dynamic facial features via separate pathways depending on the perceptual task. When the same features are task-irrelevant, their representations are limited to early visual areas. Our study offers a novel framework to understand how the brain computes 4D social information critical for socio-emotional perception and decision-making.

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) (P.G.S); ERC [FACESYNTAX; 759796] (R.E.J); the Wellcome Trust [214120/Z/18/Z] (R.I.).

53.429 RACE FAMILIARITY MODULATES NEURAL FACE CATEGORIZATION IN CHILDREN: EVIDENCE FROM FAST PERIODIC VISUAL STIMULATION (FPVS)

Grace Wallsinger[1] (gwallsinger@ufl.edu), Maeve R. Boylan[1], Ryan Barry-Anwar[1], Lisa S. Scott[1]; [1]University of Florida

Preschool-aged children have been found to demonstrate superior recognition of own-race faces. However, the developmental trajectory of the categorization of face sex and race during the preschool years is not well understood. The present study examined neural categorization of sex and race in children aged 3-5 years using Fast Periodic Visual Stimulation (FPVS). Continuous EEG was recorded while faces from one category were presented at a frequency of 6 Hz (the standard category). Every 5th face was a face from a different category (the oddball category presented at 1.2 Hz). We tested 4 types of categorization: 1) race categorization within male faces, 2) race categorization within female faces, 3) sex categorization within familiar-race faces, and 4) sex categorization within unfamiliar-race faces. Race group familiarity was coded for each participant based on a parent-completed questionnaire about the racial identities of people their child frequently interacts with. It was predicted that children would show robust evidence of neural categorization within the most commonly experienced groups (female faces and familiar-race faces) and that older children would show stronger effects than younger children. Results reveal a main effect of race group familiarity for the 6 Hz response, with increased amplitudes for familiar-race faces compared to unfamiliar-race faces. Additionally, a robust 1.2 Hz response was observed across all conditions, demonstrating children's ability to differentiate between face categories. This 1.2 Hz response was also modulated by face familiarity, with larger responses for familiar-race faces. Preliminary topographic analyses of the 1.2 Hz response show significant medial and right occipitotemporal activity - further highlighting the role of familiarity in shaping neural responses to faces. These findings suggest that prior familiarity with certain face group categories influences both the general visuocortical face processing response and the neural categorization of faces during childhood.

53.430 CORTICAL AND SUBCORTICAL CORRELATES OF EMOTIONAL EXPRESSION IN FACE PAREIDOLIA

Susan Wardle[1], Jessica Taubert[1,2], Malcolm Udeozor[1,3], Chris Baker[1]; [1]National Institute of Mental Health, [2]University of Queensland, [3]University of Cincinnati

Face pareidolia (the perception of illusory faces in objects) engages cortical areas in the fusiform gyrus and on the lateral occipital surface in the human brain (Wardle et al., 2020, Nature Communications). In the macaque, face pareidolia also activates the amygdala (Taubert et al., 2022, Social Cognitive & Affective Neuroscience). Although emotional expressions are attributed to illusory faces by human observers (Wardle, Paranjape et al., 2022, PNAS), it is unknown whether the human amygdala responds to face pareidolia, and whether its response is mediated by emotional expression. Here we measured the BOLD response to human faces and illusory faces in cortical and subcortical human brain regions using 7 Tesla fMRI (1x1x1mm voxels, TR = 2s). Participants (N = 21) viewed photographs of 21 human faces with natural and candid emotional expressions (happy, angry, and neutral), and 21 examples of face pareidolia with the same expressions. For each illusory face, we included a matched control image of a similar object that did not elicit any face perception (21 images). Each of the 63 images was shown once per run for 300ms with a 5.7s ISI; participants completed 6-7 runs. We used a 1-back task to maintain attention in the scanner. We defined face-responsive regions of the amygdala and cortical face areas in occipitotemporal cortex using independent localizer runs containing different emotional face and object photographs. Human faces produced the strongest response in the amygdala, regardless of their emotional expression. The response to illusory faces was stronger than that of similar objects without a face, and in the amygdala this was modulated by the perceived emotional expression. Together, the results demonstrate a subcortical response to face pareidolia in the human brain and show that illusory faces share emotional expression processing mechanisms with real human faces.

This research was supported (in part) by the Intramural Research Program of the NIMH (ZIAMH002909).

53.431 EARLY REPETITION SUPPRESSION FOR FACES IS BASED ON TASK-RELEVANT FEATURES

Jérémy Lamontagne[1] (lamj54@uqo.ca), Francis Gingras[1,2], Laurianne Côté[1], Alexis Bellerose[1], Chanelle Demeule[1], Caroline Blais[1], Daniel Fiset; [1]Université du Québec en Outaouais, [2]Université du Québec à Montréal

A previous study from our laboratory revealed the N170 amplitude correlates with diagnostic information for face recognition. The present study aims to replicate these results and ascertain the task specificity of the effect by extending it to a repetition suppression design. Six participants (3 women) completed 6000 trials of a same/different identity task while EEG data was being recorded. The stimuli were grayscale faces of 84 celebrities (42 women), with 3 different images per identity. Each trial presented two successive faces,one of which was filtered using Bubbles, both presented for 1s and preceded by a fixation cross shown for 1-1.4s. A classification image (CI) representing the information leading to accurate categorization was computed, and trials were then grouped in 6 levels of 1000 trials, based on their amount of available diagnostic information. Group analyses reveal a linear positive relationship between the quantity of diagnostic information and the N170 amplitude on the first face, $F(4) = 72.16$, $p = .001$ as well as a negative relationship between the diagnostic information group and the amplitude of the N170 for the second face, $F(4) = 15.52$, $p = .017$ and $F(4)=6.56$, $p = .062$ respectively for same and different identities. Additionally, group CIs showed that the N170 amplitude was driven by the presence of the eyes and mouth in trials containing highly diagnostic information. This study suggests a direct link between featural information processing and the facial repetition suppression found on the N170. The results also suggest that the N170 is sensitive to identity, demonstrated by a better fit of the regression model for trials in which the identity was repeated.

53.432 TRACING THE FLOW OF PERCEPTUAL FEATURES IN THE OCCIPITAL-VENTRAL PATHWAY FOR FACIAL IDENTITY ANALYSIS

Yinuo Yang[1], Yuening Yan[2], Philippe Schyns[2], Jiayu Zhan[1]; [1]Peking University, [2]University of Glasgow

Human faces comprise a high-dimensional information space, characterized by rich and complex variations in both 3D shape and complexion. Identifying individual faces thus requires feature-level analysis sufficiently detailed and efficiently versatile to support accurate and rapid recognition. While previous research has identified brain regions selectively involved in face processing (as distinct from those for place or tool processing), the feature-level representations of face-related brain activity remain poorly understood. To address this gap, we asked participants to evaluate the perceptual similarity between randomly generated face identities, each parametrized with 200 shape features using a generative face model. In each trial, two of these faces were randomly paired, each displayed a fixed duration (500 ms) and separated by a blank interval (800-1200 ms); Behavioral responses were allowed only after the second face disappeared. Concurrently, we measured the single trial source-reconstructed MEG activity on 8196 sources throughout the task. To trace the representational dynamics, we computed the mutual information between feature samplings and brain responses of each source on every 4 ms during the first face presentation. As expected, feature representations started early in the occipital region (~80 ms post-stimulus) and propagated along the ventral pathway towards the fusiform gyrus. Upon reaching the fusiform gyrus, the representational dynamics revealed a global-to-local processing pattern: the first peak, around 120 ms post-stimulus, represented the global contour information defining the overall length and width of the face; A second peak, around 240 ms post-stimulus, represented more localized feature information, including the shape details of eyes, nose and cheekbones. Notably, this global-to-local processing pattern was not observed in the occipital region. Our results provide the first detailed account of the dynamic feature representation of 3D faces along the occipito-ventral pathway in the human brain.

53.433 UNDERSTANDING VISUAL SYMMETRY-RELATED RESPONSES IN BRAINS AND MACHINES

Fernando Ramirez[1] (fernando.ramirez@nih.gov), Harish Katti[1], Peter Bandettini[1]; [1]NIMH

Bilateral symmetry is a fundamental property of organisms that actively navigate the world they inhabit. This property has been implicated in the estimation of face and body orientation, as well as exploited by prominent models of face identification to achieve representations invariant to head rotations. In line with these models, single-cell recordings from the macaque anterior lateral (AL) face-patch have shown a high-concentration of neurons exhibiting reflection invariance, or, in other words, neurons that exhibit similar responses to mirror-symmetric views of a face, like its left- and right-profile views, albeit distinct responses to different face-identities. However, mixed evidence of mirror-symmetry has been reported by neuroimaging studies of human high-level visual areas, and the precise source of findings of mirror-symmetry in monkeys and humans alike remains unclear. A key interpretational challenge regards the confounding role of low-level image properties associated with head rotations [Revsine et al., 2024 J Neurosci]. Nonetheless, a recent study probing convolutional neural networks [Farzmahdi et al., 2024 eLife] concluded that reflection invariant face-representations emerge in fully-connected network layers by pooling of reflection equivariant responses from earlier processing stages, and these properties could not be attributed to low-level image confounds. Here, we show that confounds consistent with those previously described in human neuroimaging studies are present in the images used in this study. More important, we show current methods to control for low-level confounds by equalizing the spectra of images presented to humans and/or artificial networks are limited; cortical magnification of the fovea in primates and a centrality bias of diagnostic features in pre-trained neural networks render spectral equalization ineffective as a control for low-level image confounds. As a way forward, we propose to embrace the variability—instead of controlling it—and suggest explicit experimental manipulations and analyses to identify genuine mirror-symmetric responses within a cross-validated model-comparison framework.

53.434 CATEGORY-SELECTIVE NEURAL ACTIVITY DECREASE IN THE HUMAN VENTRAL OCCIPITO-TEMPORAL CORTEX IN INTRACEREBRAL RECORDINGS

Corentin Jacques[1] (corentin.jacques@univ-lorraine.fr), Jacques Jonas[1,2], Sophie Colnat-Coulbois[1,3], Bruno Rossion[1,2]; [1]Université de Lorraine, CNRS, F-54000 Nancy, France, [2]Université de Lorraine, CHRU-Nancy, Service de Neurologie, F-54000 Nancy, France, [3]Université de Lorraine, CHRU-Nancy, Service de Neurochirurgie, F-54000 Nancy, France

Visual object recognition is a fundamental human brain function supported by a bilateral network of brain regions extending ventrally from the occipital pole to the anterior temporal cortex. Within this network, category-selective brain regions exhibit differential activity to their preferred category (e.g., faces) relative to other categories, as measured with various neuroimaging methods. While category-selectivity is generally defined as an increased neural activity for the preferred category, activity decreases to a specific category have also sometimes been described. Here, we investigate selective neural activity increase and decrease to face relative non-face stimuli with intracerebral recordings across the whole ventral occipitotemporal cortex (VOTC) of a large sample of participants (N=140). Objectively tagged face-selective responses in the high-frequency broadband range (HFB: 30-160 Hz) were found in 10% of recording sites distributed over the whole VOTC, with regional peaks of activity around and along the fusiform gyrus, leading to four key observations. First, face-selective HFB neural activity is characterized both by response increase (signal+ : 57% of sites) and decrease (signal- : 43% of sites) relative to other non-face categories, with a clear lateral-medial anatomical distinction along the VOTC for signal+ and signal- face-selective sites. Second, signal+ activity manifests clearer markers of face-selectivity: higher amplitude and face-selectivity, right hemispheric dominance,... Third, while increases and decreases exhibited mostly overlapping time-courses, slightly shorter response durations were found for response decrease. Four, signal- sites reflect either a lower response to faces than nonface objects (~70% of signal-sites) or an active suppression in response to face images (~30% of sites). Overall, our observations suggest a potential role for category-selective decreases in neural activity in visual (face) recognition.

Funding: ERC AdG HUMANFACE 101055175

53.435 GENDER STEREOTYPE ASYMMETRY IN WORD-FACE AND FACE-WORD PRIMING: AN ERP STUDY

Francesca Pesciarelli[1,2], Luana Serafini[1]; [1]University of Modena and Reggio Emilia, [2]San Diego State University

While implicit and automatic gender stereotyping and its neural correlates have been widely investigated in language, their functioning in the human face processing domain remains largely unexplored. We recorded Event-Related Potentials (ERPs) to a target face (male, female) preceded by stereotypically associated words (e.g., conducente "driver", badante "caregiver"), or to stereotypically associated target words preceded by a face (male, female). Participants performed a gender categorization task on target faces and a lexical decision task on target words/non-words. ERPs to both face and word targets showed a gender stereotype asymmetry. Female faces elicited a larger N200 when preceded by stereotypically gender-congruent than -incongruent prime words, and larger anterior P300 and LPP when preceded by stereotypically gender-incongruent than -congruent prime words; male faces elicited a larger N400 (limited to men) when preceded by stereotypically gender-incongruent than -congruent prime words. Similarly, stereotypically female words elicited a larger midline N400 when preceded by a gender-incongruent than congruent face, and stereotypically male words elicited a larger P300 when preceded by a gender-incongruent than congruent face. Results show that faces are potent triggers of gender stereotyping.

53.436 PARALLEL AND INTEGRATED PROCESSING OF SHAPE AND TEXTURE IN PRIMATE INFEROTEMPORAL CORTEX

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

Object images consist of both shape and texture, two fundamental components of visual recognition. Previous behavioral studies have identified shape as the primary feature for object classification in humans, with evidence indicating that the inferotemporal (IT) cortex is organized primarily around object shape, underscoring its crucial role in classification. However, the mechanisms by which the IT cortex extracts and utilizes texture information for object classification remain poorly understood. To address this gap, we conducted functional magnetic resonance imaging (fMRI) studies on macaques to identify regions within the IT cortex that are selective for animate shape and texture. For texture, we contrasted animate textures (e.g., fur) with inanimate textures (e.g., wood), while for shape, we contrasted silhouette images of animate versus inanimate objects. Both contrasts revealed several bilateral regions extending from the posterior to the anterior IT cortex. Interestingly, the shape- and texture-selective regions in the middle IT cortex were anatomically adjacent but exhibited minimal overlap, whereas in the anterior IT cortex, these two regions showed significant overlap. Electrophysiological recordings using Neuropixels further confirmed the presence of two distinct neuronal populations in the middle IT cortex of macaques, each with a strong preference for either animate texture or shape. This clear separation was not observed in the anterior IT cortex, where neurons showed a combined preference for both shape and texture. These findings suggest that object classification in the middle IT cortex involves parallel processing pathways for shape and texture information, with distinct neuronal clusters dedicated to each feature type. In the anterior IT cortex, however, these features are integrated, resulting in a more comprehensive representation. Our results reveal a novel hierarchical structure within the IT cortex: object information is decomposed in the early IT, but integrated at a later stage.

Face and Body Perception: Development, clinical

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.437 PREFERENCE FOR EYES DECREASES IN THE FIRST YEAR OF LIFE IN INFANTS WITH A FAMILIAL HISTORY OF AUTISM

M.D. Rutherford[1] (rutherm@mcmaster.ca), Esin Gürcan[1]; [1]McMaster University

Autistic individuals have difficulty processing information in the eye region, and autistic traits are more common in their family members. In this longitudinal study, we examined looking to eyes and faces in infants with and without relatives with autism, using faces with open or closed eyes. Infants were recruited at 3, 6, 9, and 12 months, and later assessed for autism. A linear mixed model was used to predict preferential looking toward the eyes. There were main effects for open eyes at both 9 ($t(335) = -2.70$, $p = 0.007$; Std.beta = -0.57) and 12 months ($t(335) = -2.03$, $p = 0.043$; Std.beta = -0.43), and relatedness to autism ($t(335) = -2.99$, $p = 0.003$; Std.beta = -0.89). The interactions between age and relatedness to autism were significant at all age points: 6 months ($t(335) = 2.02$, $p = 0.044$; Std.beta = 0.73), 9 months ($t(335) = 2.82$, $p = 0.005$; Std.beta = 0.99), and 12 months ($t(335) = 2.31$, $p = 0.021$; Std.beta = 0.81). Infants preferred looked to the eyes rather than the mouth at 6 months ($t(239) = -2.00$, $p = 0.047$; Std.beta = -0.51) and 12 months ($t(239) = -3.18$, $p = 0.002$; Std.beta = -0.84). Relatedness to autism was marginally predicted this preference ($p = 0.063$). The interactions were significant at 6 ($t(239) = 2.04$, $p = 0.042$; Std.beta = 0.74) and 12 months ($t(239) = 2.48$, $p = 0.014$; Std.beta = 0.95). The comparison group initially preferred the eyes less but developed a stronger preference over time, surpassing the high-risk group. At 9 months, they preferred open eyes more, though this wasn't seen in the eyes vs. mouth task, possibly reflecting typical language-related attention to the mouth.

This research was funded by Natural Sciences and Engineering Research Council of Canada grant to M.D. Rutherford

53.438 AGE-RELATED DIFFERENCES IN PREFERRED FIXATED REGION (PFR) FOR FACE PERCEPTION

M. Eric Cui[1,2], Björn Herrmann[1,2], Allison B. Sekuler[1,2,3]; [1]Rotman Research Institute, Baycrest Academy for Research and Education, [2]Department of Psychology, University of Toronto,, [3]Department of Psychology, Neuroscience, and Behavioural Neuroscience, McMaster University

Face regions differ in their information value for face discrimination, with more information in eye compared to nose. Despite this information difference, individuals vary in their preferred fixation region (PFR) when viewing faces: There are eye-lookers and nose-lookers. This could reflect individual differences in optimal processing regions for face perception. Peterson & Eckstein (2013) found nose-lookers performed better when viewing the nose rather than eyes, although results vary across studies (cf. Arizpe et al., 2016). Older adults as a group fixate on the lower halves of faces more than younger adults (e.g., Firestone et al. 2007), and older adults show reduced face perception accuracy (e.g., Konar et al., 2013). Here, we ask: Are nose-lookers more prevalent among older adults? and Is the link between PFR and optimal processing the same in older and younger adults? We investigated PFR in 27 younger and 27 older adults, examining face perception accuracy under free-viewing and region-restricted viewing (forehead, eye, nose, and mouth), following the method of Peterson & Eckstein (2013). Preliminary findings revealed a higher proportion of nose-lookers among older (50%) than younger adults (20%). Older observers generally showed worse performance accuracy than younger adults, but performance between eye-lookers and nose-lookers did not differ for each age group. Overall patterns of performance were similar across age groups for eye-lookers, with peak performance for the eye, and relatively poor performance for forehead and mouth. Older nose-lookers showed a clear benefit for the nose, but younger nose-lookers showed strong performance for both eyes and nose. All nose-lookers showed strong performance for the mouth. These findings suggest that age-related changes in PFR do not fully explain reduced face perception in older adults. We currently are analyzing EEG measures to learn how neural activities associated with face detection (N170) and identification (N250) relate to PFR across adulthood.

ABS's NSERC and BH's CRC

53.439 HUMAN-ANIMAL INTERACTIONS RECALIBRATE THE FACE-SELECTIVE NETWORK IN CHILDREN.

Jessica Taubert[1], Cate MacColl[1], Callyn Farrell[1], Amanda K. Robinson[1], Virginia Slaughter[1], Jason B. Mattingley[1]; [1]The University of Queensland

Our ability to recognize faces is calibrated by experience. For example, if children are exposed to monkey faces, they retain the ability to recognize monkey faces, unlike children that are not exposed to monkey faces (Pascalis et al., 2002, Science). It remains unclear, however, how experience infleunces the underlying neural representation of faces. Here we measured the BOLD response to human faces and dog faces in children between 8 and 12 years of age (N = 71) using 3 Tesla fMRI. During each run, participants viewed 40 images of human faces, 40 images of dog faces, and 40 images of non-face objects (vehicles). To maintain their attention, we used a "Catch the Pokémon" task whereby participants reported the presentation of Pokémon characters as quickly as possible. All participants completed 4-8 runs. The children were divided into two groups (35 neurotypical children and 36 neurodiverse children) based on an official diagnosis of attention-deficit hyperactivity disorder and/or autism spectrum disorder. In addition, caregivers reported whether they had a pet dog at home and the quality of the interactions between the child and the dog. We defined face-selective areas in occipitotemporal cortex using the contrast between human faces and objects. Dog faces elicited more activity in face-selective areas for children in the neurodiverse group than the neurotypical group. For both groups, the magnitude of the response in face-selective areas was correlated with the number of months the child had lived with a pet dog at home. We are currently using these data to investigate the possibility that human-animal interactions, and increased attention to the faces of pets, can "reboot" the face-selective network in neurodiverse children. In sum, our findings suggest that having a pet dog at home during childhood can change the tuning properties of face-selective brain areas in children.

Funded by 2022 NARSAD Young Investigator award (BBRF 31020)

53.440 INHIBITION BETWEEN FAMILIAR FACE REPRESENTATIONS BUILDS UP OVER TIME

Holger Wiese[1], Linda H. Lidborg[1], A. Mike Burton[2]; [1]Durham University, [2]University of York

Humans know literally thousands of faces. Accordingly, the question arises why we so rarely mix them up. As a potential answer, the representation of a familiar face, once activated, may inhibit potentially competing representations, therefore focusing activation on the best match. In contrast to this suggestion, we recently demonstrated that two face representations can be activated simultaneously, and without apparent cost, when the stimuli are presented at different locations. Here, we tested whether simultaneous activation or inhibition is observed when two faces are presented sequentially, but in the same location. We used event-related brain potentials in a repetition priming paradigm and presented two prime stimuli in sequence, each shown for 200 ms, followed by a fixation cross (600 ms) and the target (1,000 ms). While we never repeated specific images, either the first, the second, or neither of the two primes could show the same identity as the target. In Experiment 1, using only familiar faces, we found more negative amplitudes for repetitions of the second prime relative to non-repetitions at occipito-temporal channels from approximately 300 ms after target presentation. No corresponding priming effect was observed for the repetition of the first prime, suggesting inhibition by the second stimulus. In Experiment 2, primes could be either familiar or unfamiliar, while all targets were familiar faces. Again, no priming was observed for the first prime. Critically, amplitudes to familiar first-prime repetitions were less negative when the second prime was familiar relative to unfamiliar, suggesting stronger interference in the former relative to the latter condition. We conclude that inhibition of competing familiar face representations does occur, potentially focusing activation on a single best match. The comparison to our previous results further suggests that such inhibition builds up over time and is effective with a 200 ms stimulus asynchrony.

Funded by Economic and Social Research Council (ESRC, UK) grant ES/X014657/1 to HW and AMB.

53.441 TWO CASES OF PROSOPOMETAMORPHOPSIA WHOSE DISTORTIONS ARE FACE-SPECIFIC IN DAILY LIFE AND IN TESTING WITH 510 OBJECT CATEGORIES

Sarah Kerns[1] (sarah.h.kerns.gr@dartmouth.edu), Antonio Mello[1], Bradley Duchaine[1]; [1]Dartmouth College

For more than 50 years, neuroscientists have investigated the visual system for the presence of specialized face processing mechanisms. Face-specific deficits in prosopagnosia have provided key support for the existence of face-specific processes. However, the number of object categories tested alongside faces has been limited by practical constraints, and prosopagnosic self-assessment of object recognition in daily life is challenging. As a result, clear dissociation between impaired face perception and intact object perception has been documented in only a small number of object categories.To expand the number of face/object comparisons, we assess face and object perception in participants with prosopometamorphopsia (PMO), a perceptual deficit characterized by distortions of the shape, feature position, texture, and color of faces. As with prosopagnosiacs, some PMO participants have perceptual deficits for both objects and faces. Here, five PMO participants were presented with 100 photographs of faces and 1020 paired photographs of objects from the THINGS database (Hebart et al., 2019). To increase the chances of finding object distortions, a computational model of fusiform face area activation (Murty et al., 2021) was used for image selection. Participants indicated whether they saw distortion in each image. Three participants reported daily-life distortions in both faces and objects, and also demonstrated face and object distortion in testing, with 2.8%, 27.4%, and 36.7% of object images and 42%, 90%, and 95% of face images, respectively, distorting. In contrast, two participants, A.S. and Aurora, both reported universal face distortion in daily life, with zero object distortion. Their self-report and test scores were consistent, with 100% of faces and 0% of objects distorted during testing. These dissociations in face and object perception, both in daily life and formal testing, provide a strong demonstration of face-specific deficits and compelling support for face-specific mechanisms in the human visual system.

53.442 THE PHENOMENOLOGY OF FACE BLINDNESS: HOW DO INDIVIDUALS WITH DEVELOPMENTAL PROSOPAGNOSIA EXPERIENCE FACES?

Erling Nørkær[1], Tone Roald[1], Randi Starrfelt[1]; [1]University of Copenhagen

Developmental prosopagnosia (DP) is a neurodevelopmental condition of unknown origin, characterized by lifelong difficulties in recognizing faces. Research in DP has increased starkly over the last two decades, but the increased interest has not resulted in consensus on how to characterize the condition. Most studies have taken an experimental approach based on cognitive neuropsychology and neuroscience, aiming to understand how cognition is affected and relate findings to models of neurotypical face processing. Here, we took a methodologically alternative approach, using in-depth interviews and phenomenological analyses to investigate what individuals with DP experience when they perceive and attempt to recognize familiar faces. Participants were six adults with DP as established by a standard diagnostic procedure. We analyzed our interview data using Giorgi's descriptive phenomenological method in order to delineate a general structure of the DP experience. We found that face perception in DP is characterized by an unusual temporal and spatial instability of the percept; faces as whole gestalts quickly fade into facial fragments, and different visual instances of the same face can only be matched to one facial identity through conscious effort. The context in which the face appears takes the stage as perceptual figure, while faces are perceived as ground. Face recognition in DP is an effortful problem-solving process in which contextual and facial information is pieced together like clues to a mystery. Our findings suggest that faces hold an altogether different epistemological value for individuals with DP. Rather than being readily available, stable, identity marking gestalts, faces are fragmented, unstable and enigmatic visual objects that in themselves carry little or no information about who someone is. Our description of the phenomenology of DP may inspire new ways of investigating face processing in both experimental research and through clinical assessment tools of face recognition ability.

This research was supported by a grant from the Independent Research Fund Denmark (grant number DFF-1024-00139B)

53.443 VISUAL PERCEPTUAL DISTORTIONS REVEAL AN OBJECT-CENTERED REFERENCE FRAME FOR FACES AND A STIMULUS-CENTERED REFERENCE FRAME FOR HANDS

Antônio Mello[1], Chandana Kodiweera[1], Daniel Stehr[1], Krzysztof Bujarski[1], Brad Duchaine[1]; [1]Dartmouth College

A reference frame (RF) is the spatial coordinate system in which a visual stimulus is represented. Many theories of object recognition have proposed the existence of object-centered RFs in the human ventral visual pathway, but evidence for them is limited. Here, we investigate the RFs underlying the representation of high-level categories in a rare case of hemi-prosopometamorphopsia. Nagel is a 40-year-old man who sees distortions on the right side of faces and, less frequently, on the same side of hands. We presented Nagel with 432 photographs of faces and hands (216 of each) that varied in visual field presentation, the part of the stimulus visible, picture-plane orientation (0°, 90°, 180°, 270°), and viewpoint. He reported what he perceived for each image. Analyses using Fisher's exact test (Bonferroni-corrected) revealed no significant difference in the frequency of distortions across visual fields for either faces or hands ($p = 1.00$). Nagel also consistently perceived distortions on the right half-faces and right half-hands but not on their left counterparts ($p = .025$). Together, these findings demonstrate that his condition does not result from disruptions to representations encoded within retino-centered RFs. Moreover, distortions were almost exclusively confined to the right half of faces, regardless of orientation or viewpoint (i.e., distorting the same set of facial features), indicating disruptions to an orientation- and view-independent object-centered RF – specifically, a face-centered RF. In contrast, distortions affecting hands were primarily localized to the right-hand side of stimuli across the same conditions (i.e., distorting the right side of the stimulus, even when hands were rotated), suggesting disruptions to a stimulus-centered RF instead. Overall, these findings provide additional evidence for the existence of face-centered RFs in human face processing and indicate that the ventral visual system likely employs multiple RFs to encode high-level object representations.

We thank the Hitchcock Foundation for funding this research.

53.444 THE ROLE OF EARLY VISUAL EXPERIENCE ON FACIAL EXPRESSION AND IDENTITY TUNING IN MACAQUE IT CORTEX

Saloni Sharma[1] (saloni_sharma@hms.harvard.edu), Margaret Livingstone[1]; [1]Harvard Medical School

Faces convey essential information about identity and expression, which are central to social communication. Decades of research have shown that face-selective regions in the primate inferotemporal (IT) cortex are involved in such aspects of visual face processing. Previous work has shown that monkeys raised without seeing faces instead develop a behavioral preference and IT neural selectivity for the socially relevant and experienced stimuli - hands. This is dramatically different from control monkeys who, like humans, prioritize looking at faces and show neural face-only selectivity. This divergence raises an important question: how does early sensory deprivation impact core IT functions, such as facial expression and identity processing? Here, we conducted behavioral and electrophysiological experiments with three such monkeys 2, 3 or 5 years after re-exposure to faces and three typically reared monkeys. To measure behavioral preferences, monkeys freely viewed image pairs of faces, hands or objects. To investigate neural selectivity, we presented the monkeys with images of faces, hands and objects while monkeys passively fixated. Finally, to investigate facial expression and identity tuning, we presented 18 monkey identities and two expressions – threat and neutral. Behaviorally, the monkeys with 2-3 years of face experience still retain a hand bias when freely viewing images, while the monkey with five years of face experience recovered from the hand bias. However, the IT neurons of all three monkeys never developed normal face selectivity, instead showing mixed selectivity for faces and hands. Further, while we could successfully decode facial identity information across expression and vice-versa in control monkeys, all three face-deprived monkeys failed to show significant expression or identity tuning. These findings suggest that early sensory deprivation disrupts the normal development of neural mechanisms underlying face processing in higher-order visual cortex, highlighting the importance of early experiences in shaping the social brain.

R01 EY025670; William Randolph Hearst Fund

Object Recognition: Features and parts

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.445 THE REPRESENTATION OF SPIKY OBJECTS IN THE HUMAN VISUAL CORTEX

Boyang Hu[1], Marvin Chun[1], Yaoda Xu[1]; [1]Yale University

The hallmark of primate vision is its ability to extract and represent different visual features from the surrounding environment, allowing us to easily recognize and interact with the objects in the natural world. Subregions of the macaque inferotemporal (IT) cortex have been shown to respond differentially to spiky and stubby objects (Bao et al., 2020). Extending this result, we recently unveiled several distinct regions preferring spiky over stubby objects in the human IT cortex for both animate and inanimate stimuli. To understand what is driving this spiky-object preference, we carried out two fMRI experiments to examine how the relative body-to-limb aspect ratio (Experiment 1) and the number of limbs on a body (Experiment 2) affect responses in areas preferring spiky objects. Using the output of a Convolutional Neural Network previously shown to mirror macaque IT's representations of stubby and spiky objects, in Experiment 1, we collected a set of four-legged mammal stimuli varying in body-to-limb aspect ratio, from stubbier mammals with large bodies and short limbs (e.g., wombat) to spikier mammals with skinny bodies and long legs (e.g., greyhound). Preliminary results showed that in visual areas preferring spiky objects, fMRI responses somewhat tracked the body-to-limb aspect ratio, being greater for greyhounds than wombats, but also high for animals with intermediate ratios, such as puma. In Experiment 2, using images of ceiling fans with 3, 5, 8, or 12 blades to vary the number of limbs on the hub, preliminary results showed that the areas preferring spikiness exhibited a U-shaped response profile, with responses being higher for 3- and 12-blade fans than for 5- and 8-blade ones. Our study constitutes a first step toward a better understanding of what aspects of an object's shape drive the responses in areas preferring spiky objects in the human visual cortex.

This research was supported by NIH grant R01EY030854 to YX.

53.446 A NEURAL COMPUTATIONAL FRAMEWORK FOR VISUAL OBJECT CODING IN THE HUMAN BRAIN

Runnan Cao[1] (rncao90@gmail.com), Jie Zhang[1], Jie Zheng[2], Yue Wang[1], Peter Brunner[3], Jon Willie[3], Shuo Wang[1,3]; [1]Washington University in St. Louis, [2]University of California Davis, [3]Washington University in St. Louis

A critical question in cognitive neuroscience is how unified category representations emerge from visual inputs that undergo high-dimensional visual changes, such as visual variations in shape, color, and texture. Two distinct hypotheses have been proposed. On the one hand, non-human primate studies suggest that single neurons in the inferotemporal cortex, which is homologous to the human VTC, encode specific feature axes. On the other extreme, neurons in the human MTL are reported to encode abstract concepts related to individual persons or places, demonstrating a highly selective and sparse code. Little is known about how the perceptual-driven representations in the VTC are transformed into memory-driven representations in the MTL. It is also unclear whether the human VTC encodes visual objects using an axis-based code. To address these questions, we recorded intracranial EEG activity across the VTC and MTL areas, along with single-neuron activity in the MTL. Indeed, the human VTC exhibited strong axis-based feature coding, as shown in primate studies. By constructing a neural feature space using the VTC neural axes, we observed that MTL neurons encode receptive fields within the VTC neural feature space, exhibiting region-based feature coding. This region-based coding of MTL neurons at the low-dimensional feature space, which was derived from the VTC, may serve as an intermediate step that connects dense visual representations to sparse semantic representations. Notably, we uncovered the physiological basis for this coding transformation, showing that the VTC axis-coding channels exhibited stronger synchronization with the MTL. These findings propose a new computational framework for object recognition, advancing our understanding of both visual perception and visual memory.

This research was supported by the NIH (K99EY036650, R01MH129426, R01MH120194, R01EB026439, U24NS109103, U01NS108916, U01NS128612, R21NS128307, P41EB018783), AFOSR (FA9550-21-1-0088), NSF (BCS-1945230), and McDonnell Center for Systems Neuroscience.

53.447 COGNITIVE STRATEGIES IN THE MENTAL ROTATION TASK REVEALED BY EYE TRACKING

Brandon Eich[2] (beich@arizona.edu), Tom Tomshe[1], Raevan Hanan[1], Chloe Kindell[1], Hanane Ramzaoui[1], Heather Lucas[1], Melissa Beck[1]; [1]Louisiana State University, [2]University of Arizona

Mental rotation is a spatial reasoning skill that is predictive of general intelligence, creativity, and STEM performance. In a common mental rotation task, two images from different angles are viewed side by side and are either identical (same trial) or mirrored versions of each other (different trial). Typically, there is a linear increase in response time (RT) as the angle of disparity (AoD) increases between stimuli. This increase in RT is thought to represent a strategy of using mental imagery to rotate, like the physical action of rotating, one of the images to see if it can match the other, with this process taking longer the larger the AoD between stimuli. However, there are other more analytic strategies that may be used as well. The mental rotation strategy used between trials with small versus large AoDs was examined in a large sample of undergraduate-student participants while their eye-movements were tracked. Results suggest that for smaller AoDs (20°, 40°, & 60°), RT increased as AoD increased, suggesting that rotation of visual imagery strategies were used. However, for larger AoDs (120°, 140°, & 160°), RTs plateaued and did not increase, suggesting a different strategy was used. Consistently, the number of fixations and saccade amplitude within an object, and the number of gaze changes between objects increased as angular disparity increased (i.e., from 20° to 60°), but then plateaus at larger AoDs. Together, these results suggest that at smaller AoDs a strategy of mentally rotating stimuli was used, while at larger AoDs, a piecemeal strategy was used, where parts of the object were used for comparison. This difference in strategy led to an increase in RT and decrease in accuracy up until the larger AoDs, where participants made similar eye-movements across the larger AoDs and produced similar RTs and accuracy.

This project was supported by an NSF grant (AM230273) titled Enhancing STEM Success: A Multi-modal Investigating of Spatial Reasoning and Training in Undergraduate Education

53.448 MODELING TREE SHREW HIGH-LEVEL VISUAL BEHAVIORS

Emily E Meyer[1], Lingqi Zhang[1], Nicolas P Cottaris[1], David H Brainard[1], Michael J Arcaro[1]; [1]University of Pennsylvania

A hallmark of primate vision is the ability to quickly recognize objects despite considerable variations in how an object is projected onto the retina. However, the evolutionary origins of this behavior remain poorly understood. Among the closest relatives to primates, tree shrews (Tupaia belangeri) offer unique insights into the evolution of visual processing. Their extensive extrastriate cortex and visually guided behaviors represent key adaptations that may have supported advanced object recognition in primates. We trained three adult tree shrews on a match-to-sample task using stimuli previously used to demonstrate complex object recognition in humans, macaques, and marmosets. Like primates, tree shrews successfully identified objects across variations in position, size, and orientation, and when embedded in complex scenes. Moreover, behavioral performance was correlated across shrews, suggesting they utilize a common shape representation. To gain deeper insight into the representations driving their behavior, we compared tree shrew performance with predictions from visual processing models. We accounted for tree shrew optics using a front-end visual system model, ISETBio, then employed deep convolutional neural networks (DCNN) to probe the visual representations emerging from core features of the primate visual system—hierarchical connectivity and convolutional processing. We analyzed the correspondence between DCNN layer representations and tree shrew behavioral performance, finding that layers best predicting tree shrew performance varied with task complexity. While this provides insights into the depth of processing, it does not reveal which specific stimulus features drive tree shrew behavior. Moreover, the most diagnostic stimulus features for tree shrew behavior may not be captured by DCNNs. Therefore, we are testing models representing specific aspects of processing, including local texture (Gabor-jet model), structural shape (skeletal model), and visual saliency (SALICON). These findings help establish tree shrews as a model for high-level processing and offer insights not just about whether, but how they discriminate complex objects.

53.449 PERCEPTIONS OF SEMANTIC SIMILARITIES OF OBJECTS BASED ON VISUAL CUES IN CEREBRAL VISUAL IMPAIRMENT

Lotfi B. Merabet[1] (lotfi_merabet@meei.harvard.edu), Claire E. Manley[1,2], Peter J. Bex[2]; [1]Massachusetts Eye & Ear, Harvard Medical School, [2]Department Psychology, Northeastern University

The semantic relationships among real objects have been quantified in the THINGS image database. We used this database to investigate visual semantic behavior in individuals with cerebral (cortical) visual impairment (CVI), a brain-based visual disorder associated with higher-order visual processing deficits, including object recognition. Participants (8 CVI, mean age=20.13 years ±5.96 SD; 9 control, mean age=24.89 years ±11.57 SD) viewed a series of image pairs: one baseline pair and one experimental pair per trial, and determined which pair was more similar. In the baseline pair, one image was selected at random, and the second image differed by a pedestal of 0.1, 0.2, or 0.3 in semantic units. In the experimental pair, one image was selected at random, and the second image differed by an amount controlled by a 3-up-1-down staircase. Threshold semantic difference was estimated from the 75% point of a psychometric function fit to the proportion of 'more different' responses as a function of semantic distance across 60 trials. There was a significant pedestal effect [F(1,48)=38.82, p=1.204e-0] and group effect [F(1,49)=7.30, p=0.0095] on discrimination thresholds level, with semantic difference thresholds increasing as pedestal difference increased and thresholds being higher for the CVI group (0.1: 0.35±0.27 SD, 0.2: 0.54±0.24 SD, 0.3: 0.65±0.15 SD) than the control group (0.1: 0.12±0.08 SD, 0.2: 0.42±0.07 SD, 0.3: 0.50±0.05 SD). There was no evidence of internal noise effects. These findings suggest that individuals with CVI have greater difficulty distinguishing images of common objects, indicating a potential deficit in mental representations to interpret their visual environment.

NIH/NEI R01 EY030973

53.450 THE CONTRIBUTION OF CONFIGURAL SHAPE TO OBJECT RECOGNITION IS PROCESSED BY A LATE-

ONSET MECHANISM LIKELY LOCALIZED IN RIGHT TEMPORAL CORTEX

Shaya S Samet[1,2] (shayasam@my.yorku.ca), Jasman Kahlon[1,2], Nicholas Baker[3], Erez Freud[1,2], James H Elder[1,2], Peter J Kohler[1,2]; [1]York University, [2]Center for Vision Research, [3]Loyola University

Object shape perception depends on both local shape features, such as contour curvature, and configural shape information, which arises from the spatial relationships of these features. Specialized stimuli and behavioural methods have dissociated local and configural shape contributions. Here we used high-density EEG to investigate cortical mechanisms of local and configural shape perception, with an SSVEP paradigm that isolates differential brain responses between image-pairs. Stimuli included intact natural animal silhouettes, silhouettes where configural shape had been disrupted ('Frankensteins'; Baker & Elder, 2022), and synthetic curvature-matched-controls where local curvature was progressively matched to natural silhouettes (CMCs; Elder et al., 2018). In Study 1 (n=32), we compared pairs of CMC classes and intact animals to assess effects of local curvature and global configural constraints. Results revealed shape-contingent activity in occipital and temporal cortices peaking ~150-250ms post-stimulus, influenced by both local and configural information. Local curvature had measurable effects on brain responses, but even when paired with CMCs with identical local curvature statistics, natural silhouettes produced robust differential responses. The latter response was subject to an inversion effect, reflecting semantic and holistic processing, but using inverted silhouettes did not fully abolish differential responses, suggesting that some configural processing survives inversion. In Study 2 (n=30), we compared Frankenstein and intact shapes for shapes where 'Frankensteining' had strong or weak behavioural effects on object recognition. Responses to Frankenstein stimuli were smaller overall, peaked later (~350ms), and were localized over right temporal cortex. Importantly, these responses were abolished by inversion, and were only found for stimulus pairs that produced strong behavioural effects, establishing the role of the observed brain activity in configural shape perception. These studies demonstrate that local and configural shapes produce distinct brain responses, and that our approach can isolate brain mechanisms that process configural information to support object recognition.

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 Peter J Kohler.

53.451 WHAT DOES PERCEIVED SIMILARITY MEASURE? A SYSTEMATIC COMPARISON OF EIGHT SIMILARITY TASKS

Martin N Hebart[1,2,3], Malin Styrnal[2], Laura Stoinski[1], Philipp Kaniuth[1,2]; [1]Justus Liebig University Giessen, [2]Max Planck Institute for Human Cognitive and Brain Sciences Leipzig, [3]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen and Darmstadt

Similarity tasks are widely used in vision science and cognitive science for studying mental representations, object recognition, semantic processing, or categorization. Despite the many variants of different similarity tasks, to date little is known about how these tasks relate, the degree to which they measure the same underlying construct, and how efficient they are at measuring them when taking into account reliability. To address these questions, here we systematically compared eight different similarity tasks: (1) pairwise ratings, (2) pile sorting, (3) single and (4) multiple arrangement, (5) triplet odd-one-out judgments, (6) sequential forced choice similarity judgments, (7) speeded visual search and (8) speeded same-different judgments. We collected data from 100 online participants for each of the tasks using three sets of stimuli, ranging from natural objects embedded in scenes to abstract shapes. For the natural object images, we analyzed the similarity estimates between tasks and their alignment with deep neural network representations and a semantic embedding, respectively. The results showed that these tasks can be grouped into three types: (1) tasks that primarily capture visual features (response time tasks), (2) tasks that primarily capture semantic features (sorting tasks) and (3) tasks that capture both visual and semantic features (choice tasks). Within each group of tasks, we additionally determined their reliability and efficiency to support researchers with their choice of paradigm. For more abstract stimuli, differences between tasks became smaller, indicating that they tend to measure a similar construct, with less striking differences between tasks. Thus, when choosing a task one should consider not only what features are of interest, but also what stimuli are being used. Together, this work reveals the nature of the representations measured by different similarity tasks, provides suggestions for choosing one task over another, and highlights the role of task in the assessment of perceived similarity.

53.452 IDENTIFYING FEATURES FOR SUPERORDINATE OBJECT CLASSIFICATION THROUGH CREATIVE DRAWINGS

Filipp Schmidt[1,2] (filipp.schmidt@psychol.uni-giessen.de), Emily A-Izzeddin[1], Henning Tiedemann[1], Christian Houborg[1], Roland W. Fleming[1,2]; [1]Justus Liebig University Giessen, [2]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt

Visually classifying objects into superordinate classes, such as "plant" or "animal," presents a significant computational challenge as radically different items, like "jellyfish" and "flamingo," must be grouped together. Here, we probed observers' visual intuitions about the key features that unite members of superordinate classes by asking them to generate novel items. Sixteen participants drew new (i.e., original, unfamiliar) members of nine specified classes (animal, building, clothing, furniture, household appliance, musical instrument, plant, tool, and vehicle) as well as for a general "object" class. Another 16 participants assigned each drawing to one of the nine class labels (or suggested a new class), yielding average accuracy of 69% (compared to 10% chance and 89% for familiar control drawings). Virtually all drawings were grouped into one of the nine specified classes, including most created for the general "object" class. Performance was particularly high for unfamiliar animals, plants, buildings and vehicles. They also rated the typicality of each drawing on a 10-point scale, yielding a mean typicality of 4.9 for unfamiliar drawings, vs 8.2 for familiar control drawings. Our results suggest people can generate novel, unfamiliar drawings that capture key features of superordinate classes. To identify the visual features driving these classifications, we asked another group of 35 participants to classify the drawings, as well as mark and label the drawings' defining "parts". The most frequent labels were consistent with signature features of each class (e.g. "leg"

and “eye” for animals), while both the overlap of labels between classes—and consistency of participants’ labels—was related to classification performance. Together, our results suggest that observers learn key features shared by highly diverse members of superordinate classes, and can ‘remix’ these to create new examples—a process that can be uniquely probed through analysis of creative drawings.

Supported by German Research Foundation (222641018–SFB/TRR 135 TP C1); the European Research Council (ERC) Advanced Grant ‘‘STUFF’’ (ERC-2022-AdG-101098225); and Research Cluster ‘‘The Adaptive Mind’’ funded by the Hessian Ministry for Higher Education, Research, Science and the Arts.

53.453 THE EFFECT OF COLOR FILTERS ON FOOD RECOGNITION DEPENDS ON FOOD NEOPHOBIA AND FOOD DISGUST

Melina Mueller[1] (melina.o.mueller@vanderbilt.edu), Isabel Gauthier[1]; [1]Vanderbilt University

Color can influence experiences with food. The presence of color moderates the relationship between food recognition ability and affective responses to food. Here we investigated how specific color channels influence food recognition, as a function of two traits reflecting attitudes towards food: food neophobia (FN, fear of new foods) and food disgust (FD, aversion to harmful or offensive foods). We randomly assigned participants to perform food recognition in three color conditions: normal color, red-green filtered, or blue-yellow filtered. We measured FN and FD in non-colorblind adults ($n = 562$), followed by a food oddball recognition task. FN was negatively correlated with food recognition, with a stronger effect for filtered images (red-green: $r = -.311$, blue-yellow: $r = -.329$) compared to normal images ($r = -.124$; $z = 2.29$, $p = .02$). FD negatively impacted food recognition when red-green information was absent ($r = -.413$) compared to normal color ($r = -.043$; $z = 3.61$, $p < .001$) or blue-yellow filtered images ($r = -.156$; $z = -2.49$, $p = .01$). Color channels influence food recognition performance on a task without explicit affective judgment. FN is associated with poor food recognition ability regardless of color content, possibly because FN limits perceptual experiences with food. With high FN, unusual colors may reduce familiarity, further impairing recognition. In contrast, FD does not limit food recognition for foods in normal color. Those with high FD performed most poorly when recognizing food images lacking red-green information. This could be due to an arousal effect for foods perceived as disgusting or, food recognition in those with high FD may rely on red-green information more than for those with low FD. Color did not significantly impact performance for those with low FN and FD, suggesting that color elicits stronger affective responses for those with high trait levels.

This work was supported by the David K. Wilson Chair Research Fund (Vanderbilt University).

3D Processing: Space, coordinate frames, virtual environments

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.454 A PLATFORM-INDEPENDENT METHOD FOR STUDYING VISION SCIENCE IN AR/VR ENVIRONMENTS

Zoltan Derzsi[1,2], Robert Volcic[1,2,3]; [1]New York University Abu Dhabi, [2]Center for Artificial Intelligence and Robotics, New York University Abu Dhabi, [3]Center for Brain and Health, New York University Abu Dhabi

In vision science, experimental software development has been shaped by packages like Psychtoolbox or PsychoPy, which have become de-facto standards. While these open-source tools may benefit from active development communities, often only a handful of individuals bear the burden of support and must navigate changes beyond their control. The growing adoption of augmented and virtual reality (AR/VR) in vision science has introduced greater development challenges and shortened the product life cycle drastically. While computer graphics back-ends may have a life cycle of decades, similar support for virtual reality devices is often much shorter. To circumvent this problem, two approaches have been proposed in the past: back-porting game engine features to computer graphics software, and implementing psychophysics features into game engines as add-ons. As there is a massive hardware diversity in AR and VR hardware, both of these approaches lack general solutions that can be easily ported to different hardware platforms, rendering them out of date disproportionately quickly. Here, we introduce a universal, platform-independent communication framework for integrating VR hardware with the preferred software. We treat the virtual reality hardware and the game engine as a separate interactive volumetric display instead of an extension of other software packages, and we use simple and standardized communication with external devices. This approach allows researchers to create stimuli and control experiments in AR/VR while keeping the software they are already familiar with. We demonstrate the flexibility of our method by running the very same code on different platforms simultaneously in the same virtual space. Additionally, we highlight the ease of integrating well-known software with various hardware (motion trackers, game controllers and even custom electronics), while maintaining adaptability for future technologies.

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.

53.455 ADVANTAGE OF THE FOREGROUND FOR ABSOLUTE DISTANCE PERCEPTION IN THE INTERMEDIATE DISTANCE RANGE

Lizhu Yan[1] (lizhu.yan@louisville.edu), Lingling Bai[1], Teng Leng Ooi[2], Zijiang He[1]; [1]University of Louisville, [2]The Ohio State University

Reliable ground surface representation is vital for accurate absolute distance judgment. We proposed the ground surface is represented by a Sequential Surface Integration Process (SSIP) (Wu et al., 2004, Nature), wherein the ground around the observer’s feet is represented before the farther regions. This way, the visual system can use the

near depth cues to form a reliable ground representation, which is then used to sequentially integrate with texture gradient cues in the more distant ground regions to form a global ground representation. We further tested the SSIP hypothesis by varying the spatiotemporal availability of external depth cues. Four conditions were tested with binocular and monocular viewing. (i) Dark: no texture. (ii) Full-texture: a 2x5 parallel texture array (4 s) spanning 2-8 m in the dark. (iii) Near-to-far sequence: the texture pair nearest to the observer at 2 m was first presented (4 s) and followed sequentially by the 3.5m, 5m, 6.5m, and 8m pairs (0.75 s each). The presentation sequence was run twice. (iv) Far-to-near sequence: the texture presentation sequence was the reverse of the Near-to-far condition. The test target (0.20°, 2 s) was presented in the dark 1 s after the texture removal. The observers' (n=8) task was to judge the target location (4.5, 5.75, 7.0m @ 0.14m height and 5.75m @ 0.5m height) using the blind walking-gesturing paradigm. Judged distances in the Near-to-far and Full-texture conditions were similar and most accurate. Confirming the SSIP hypothesis, judged distances were significantly more accurate in the Near-to-far than the Far-to-near conditions ($p<0.001$). The data trend was similar with binocular and monocular viewing ($p>0.05$). No significant difference was observed between the Far-to-near and Dark conditions ($p>0.05$). Judged angular declination was similar among the four conditions ($p>0.05$), suggesting observers maintained reliable angular declination even as distances could not be accurately judged.

NIH R01EY033190

53.456 ANCHORING OF THE INTRINSIC BIAS FOR CODING DISTANCE DEPENDS ON ENVIRONMENTAL SCENE AND COGNITIVE DEMAND

Zijiang He[1], Lizhu Yan[1], Lingling Bai[1], Teng Leng Ooi[2]; [1]University of Louisville, [2]The Ohio State University

The intrinsic bias is the visual system's internal model of the ground surface (Ooi et al, 2001 Nature). It is anchored at the feet's location of a static observer tested in the dark. However, it is unknown if the anchored location shifts away from the static observer under more natural circumstances such as in the view of a walkable ground surface, and if the relocation/shift requires attentional effort. This is a significant issue as the intrinsic bias, and where it is anchored, is the basis for creating the ground surface reference frame for spatial coding of object locations. Eight observers were tasked to judge the target (0.16 cd/m2, 0.2 degree, 2 seconds) location (4.5, 5.75, 7 m @ 0.14 m height) in the dark using the blind walking-gesturing paradigm. Ten seconds prior to judging, they were exposed for 10 seconds to one of six conditions: 2 viewing environments (dark vs. impoverished) x 3 cognitive demands (neutral, count-number-backward, and lateral gaze-shift). The dark and impoverished environments had the observers exposed, respectively, to total darkness and an array of 2x8 parallel texture elements (0.04 cd/m2) on the floor (spanning 2.25-9.25 m). During the exposure, the observer either remained neutral, counted numbers backward, or directed their gaze alternately to the left and right shoulders (~0.25 Hz). Judged distances were longer in the impoverished-with-neutral condition than the dark-with-neutral condition ($p<0.001$). This suggests the visual system relocated the intrinsic bias in front of the observer when there was visual texture to define the ground surface. Additionally, data from the impoverished and dark environments became similar ($p>0.05$) when observers performed number-counting or gaze-shifting, suggesting that attention was required to relocate the intrinsic bias. In the dark environment, judged distances were similar under the three cognitive demands indicating the intrinsic bias remained anchored at the feet's location.

NIH R01EY033190

53.457 FAILURES OF DEPTH MAGNITUDE ESTIMATION IN VIRTUAL BUT NOT PHYSICAL STIMULI

Arleen Aksay[1] (aksay@yorku.ca), Laurie M. Wilcox[1]; [1]York University, Toronto, Canada

When asked to estimate depth magnitude in naturalistic, cluttered stimuli observers lacking experience with virtual reality (VR) exhibit little to no scaling of depth with binocular disparity. Here, we asked whether this failure reflects a general inability to generate metric depth estimates for these types of stimuli or if it is related to cue conflicts that limit the use of vergence to estimate viewing distance in VR. To this end, we rendered low ('branch') and high ('thicket') complexity stimuli in VR and 3D printed exact physical replicas. In the branch condition, two branches were presented on either side of a central reference branch. The thicket condition consisted of two mirrored clusters of overlapping branches centred around the reference. The total disparity ranged from 0.33 to 2.10 degrees. Separate groups of novice observers estimated the 3D volume of branches (N=28) or thickets (N=25) using a virtual ruler. Both types of stimuli were tested in VR and physically; the physical stimuli were tested monocularly and binocularly. Our VR results replicated previous outcomes: novice observers showed limited depth scaling irrespective of complexity. Their binocular functions were relatively flat and similar to those obtained monocularly. Very different results were obtained using physical stimuli depth, for which estimates scaled well with disparity. These results support our hypothesis that novice observers' poor depth scaling is due to unreliable distance information in VR, which is needed to scale binocular disparity. Importantly, experienced observers show reasonable depth scaling for virtual stimuli, suggesting that they are able to ignore cue conflicts. These findings have important implications, both for cue integration studies, which typically involve experienced observers, and for the use of VR in studies of depth perception.

Natural Sciences Research Council of Canada (NSERC); CF-REF program Vision Sciences to Applications (VISTA)

53.458 THE TEMPORAL FEATURES OF SIZE CONSTANCY IN TWO- AND THREE-DIMENSIONAL STIMULI REVEALS A REAL-WORLD ADVANTAGE

Mirko Tommasini[1], Sara Battisti[1], Romeo M. Minutolo[1], Giulia Tonielli[1], Simona Noviello[1], Juan Chen[3], Melvyn A. Goodale[2], Irene Sperandio[1]; [1]University of Trento, Trento, Italy, [2]University of Western Ontario, [3]South China Normal University, Guangzhou, China

Size constancy is the ability to maintain a stable percept of object size despite variations in the retinal image due to changes in viewing distance. Recent research using real-world objects at real distances has demonstrated that this phenomenon emerges at the earliest stages of cortical processing. It remains unclear, however, whether these findings are applicable to both 3D and 2D stimuli. Here,

participants were presented with either 3D or 2D stimuli placed at different distances and asked to perform a manual size estimation using their right thumb and index finger. The stimulus physical size was scaled with respect to distance to yield a constant retinal angle. Concurrently, electroencephalographic (EEG) data were recorded using a 64-channel scalp electrode array. Results revealed an advantage for real objects in the computation of size constancy, as indicated by an earlier difference in neural responses to small versus large stimuli, observable in the first positive-going component, peaking at ~80 ms after stimulus onset. In contrast, size constancy for 2D stimuli emerged approximately 150 ms after stimulus onset. Furthermore, stimulus predictability played a role in enabling faster size-distance integration. These findings provide electrophysiological evidence for a 'real-object advantage' in size constancy. This advantage may be partially explained by top-down mechanisms, such as affordance—the potential to physically interact with an object—which could enhance the perceptual processing of real 3D objects relative to 2D representations. Additionally, 3D objects may provide visual cues to distance that are not available in 2D stimuli, further contributing to size constancy.

53.459 THE ROLE OF PREDICTION IN CONTINUOUS MANUAL TRACKING OF 3D TRAJECTORIES

Bita Manouchehri[1], Stephanie M Shields[1], Kathryn Bonnen[1]; [1]Indiana University

Humans are able to predict and follow the trajectories of moving targets. Previous research has studied manual tracking of targets moving in unpredictable Brownian trajectories, trajectories with abrupt changes in direction, and highly predictable sine wave trajectories. Here, we test the hypothesis that participants' ability to follow a target depends systematically on the predictability of the target's trajectory. We generated 3D pink noise (1/f) trajectories and manipulated their predictability by applying a bandpass filter with a fixed lower frequency and a variable upper frequency. We presented targets dichoptically using a PROPixx projector (VPixx Technologies), we recorded responses with a LeapMotion device (UltraLeap), which allowed participants to move a cursor in 3D by moving their hand. Their task was to follow the target's motion with that cursor. We evaluated tracking performance by calculating cross-correlograms (CCGs) of the target and response velocities in the horizontal, vertical, and depth dimensions. We then quantified response latency by finding the latency of each CCG's peak correlation. Overall, we found that participants' response latency decreased as predictability increased, suggesting that more predictable trajectories were indeed easier for participants to follow. For the most predictable trajectories (with near-sinusoidal motion), response latencies were close to zero (0.038 seconds). Consistent with existing literature, we found that performance tended to be worse for motion-in-depth than for horizontal and vertical motion: Peak correlation values tended to be lower, and latencies tended to be greater. Notably, however, the motion-in-depth latency of multiple participants decreased to near-zero, essentially catching up to the horizontal and vertical latencies. Such an effect could potentially result from depth tracking operating over a longer temporal integration window. Hence, our results support the hypothesis that tracking performance improves as target trajectories become more predictable and suggest interesting directions for future research.

Color, light and materials: Neural mechanisms, clinical

TUESDAY, MAY 20, 8:30 AM – 12:30 PM, PAVILION

53.460 ASYMMETRIES IN THE STRENGTH OF CHROMATIC INDUCTION ALONG AN L-M AXIS OF COLOR SPACE

Michael Rudd[1] (mrudd@unr.edu), Sean O'Neil[1], Natasha Staneva[1], Idris Shareef[1], Michael Crognale[1]; [1]University of Nevada, Reno

We adapted a psychophysical technique previously used to measure asymmetries in light/dark induction strength (Rudd, 2013) to here measure the relative strengths of chromatic induction along an L-M opponent axis in color space. Two achromatic disks (1.16 deg diam, 40 cd/m2) were presented on opposite sides of a Display++ monitor (Cambridge Research Systems, Kent, UK), each surrounded by an achromatic annulus (80 cd/m2). The annuli had widths 0.19 and 1.78 deg, with the display side containing the thinner annulus randomized across trials. The chromatic background was 60 cd/m2. Across trials, the L/M contrast of the background was randomly varied. Four participants adjusted the L/M content of the disk surrounded by the thinner annulus to match the two disks in hue. The background induced a contrasting hue in both disks and annuli. These hue changes were stronger on the side with the thinner annulus. We wondered if, when changing saturation of the background, the rate of change in the match settings would differ across the ranges of background saturations for which L was either < M or > M. If so, this might imply the existence of separate underlying halfwave-rectified L-M and M-L mechanisms (corresponding to separate half axes of the L-M color dimension) characterized by different gains and producing different chromatic induction strengths. In fact, we did find differences across the two background ranges. The induction strength, as measured by the rate of change in the settings as the background varied, was highly reliable across observers when the disks appeared reddish (more L), but differed markedly across observers when the disks appeared greenish (more M). The implications of our results for an edge integration theory of color (Rudd, 2010) will be discussed.

NIGMS NIH P30 GM145646

53.461 CENTRAL FOVEA HIGH ACUITY: NO LUMINANCE ADVANTAGE, BIG INDIVIDUAL CHROMATIC DIFFERENCES

Lingyu Gan[1] (lingyug1@uci.edu), George Sperling[1]; [1]University of California, Irvine

Isoluminance describes a condition where lights of different spectral compositions exhibit equal luminance, allowing researchers to study visual processing independently of luminance cues. Typically, isoluminance is determined using conditions that predominantly stimulate the magnocellular pathway, such as flicker photometry, but it is often applied in conditions involving the parvocellular pathway. The magnocellular and parvocellular pathways serve distinct roles: the magnocellular pathway is sensitive to rapid changes in luminance, whereas the parvocellular pathway is tuned to high spatial resolution and color processing. This study specifically attempts to measure parvocellular isoluminance. The stimuli were the highest resolvable spatial-frequency gratings (28.6-35.8 cpd) presented within .15-.19 deg within the extreme central fovea, discernible only by the parvocellular system. Subjects briefly viewed high spatial-frequency yellow/red and yellow/green gratings and reported grating orientation. The yellow stripes were constant, and the intensity of the red or green stripes varied over the full available range to guarantee achieving isoluminance somewhere in the range. Results: (1) None of the 10 male and 10 female participants experienced a performance reduction to chance for any red or green stripe value (which included all possible isoluminant values). (2) Minimal visibility points in the central fovea generally did not coincide with isoluminant points determined by a motion-reversal paradigm. (3) Large between-subject differences in accuracy when judging the orientation of yellow/red versus yellow/green gratings formed three distinct response patterns: red-dominant, green-dominant, and red-green approximately equally dominant. For green-dominant and red-dominant subjects, luminance information was insignificant for grating acuity in gratings containing the dominant wavelength. Conclusions: (1) In the central fovea, the luminance system does not provide any advantage over the color system for grating acuity. (2) The concept of isoluminance for the multichromatic parvo system is fundamentally different from that of the monochromatic magno system, as typically determined by flicker photometry and similar methods.

53.462 CONTRAST MATCHING BETWEEN SVSLM AND LVSM SIGNALS USING STEADY STATE VISUAL EVOKED POTENTIALS

Jesse Macyczko[1], Camile Franke[1], Michael Crognale[1], Michael Webster[1]; [1]University of Nevada, Reno

At early visual stages, information about color is carried by two “cardinal” dimensions comparing signals in the LvsM cones or SvsLM cones. Equating the strength of signals along these two dimensions is important for designing and interpreting many experiments, but is inherently problematic and task-specific, and can range in metrics from cone contrasts to multiples of threshold. We developed an objective measure of the equivalent contrasts along the two axes at early cortical levels, based on steady-state visual evoked potentials. Stimuli were 1 c/deg chromatic gratings defined by the LvsM and SvsLM axes, with equiluminance determined empirically with a minimum motion task. The two color axes were shown as onset patterns and alternated at 3 or 4 Hz , with responses recorded by a single electrode positioned at OZ. During presentation the contrast of the SvsLM grating remained fixed, while the LvsM contrast was varied over a wide range across successive trials, in order to estimate the relative contrast at which the alternation response (based on summing the to the fundamental frequency and its unique harmonics) reached a minimum. The procedure is thus similar in principle to nulling paradigms such as flicker photometry. Observers exhibited nulls that were well-fit by a modified Naka-Rushton function where the portion of the function below the cross point was mirrored. This allowed us to capture the contrast match point between the SvsLM and LvsM gratings, which was found to be generally consistent with a multiple-of-thresholds scaling of the axes. This direct measure of cortical contrast can be used to track color coding in different populations, such as during development or with color deficiencies. In additional measurements we are applying the technique to compare chromatic responses in color-normal observers and anomalous trichromats in order to test for potential compensatory neural gains in the anomalous observers.

Supported by EY010834

53.463 COLOR AND RETINAL OPTICS IMPROVE THE PREDICTION OF OCCIPITAL FMRI RESPONSES TO NATURAL SCENES

Jenny Bosten[1], Chris Racey[1], Kendrick Kay[2], Ian Pennock[1]; [1]University of Sussex, [2]University of Minnesota

Color is an integral part of visual perception, yet it is often omitted from encoding models. Similarly, the retinal image deviates substantially from the RGB images typically used to train encoding models, yet the effect of retinal optics on the stimuli is often ignored. We built a Gabor wavelet pyramid-based encoding model that includes biologically plausible color and retinal optics generated using ISETBio (Cottaris et al., 2019, Journal of Vision). We fit this model using fractional ridge regression (Rokem & Kay, 2020, GigaScience), and predicted occipital BOLD responses to the color-calibrated natural scene stimuli in the 7T fMRI Natural Scenes Dataset (Allen et al., 2022, Nature Neuroscience). We created a luminance (L+M)-only model without retinal optics as a baseline model. For V1 voxels, adding color information from the ‘cardinal’ retinogeniculate color channels (S/(L+M) and L/(L+M)) increased average prediction accuracy (Pearson's r) by about 20%. Including optical factors using ISETBio to estimate retinal images increased average prediction accuracy a further 5%. Our results show that incorporating color and the retinal image into Gabor wavelet-based encoding models improves prediction of BOLD responses to natural scene stimuli. We conclude that these well-established biologically relevant features of low-level visual signals are important, and can be successfully incorporated in encoding models.

The work was funded by ERC grant 949242 COLOURCODE to JMB

53.464 FROM COLOR ENCODING TO COLOR PERCEPTION: THE TEMPORAL DYNAMICS OF COLOR VISION IN THE HUMAN BRAIN

Lina Teichmann[1], Ka Chun Lam[1], Amaia Benitez-Andonegui[1], Danny Garside[2,3], Sebastian Montesinos[1], Bin Guan[2], Francisco Pereira[1], Chris I. Baker[1], Bevil R. Conway[2]; [1]National Institute of Mental Health, [2]National Eye Institute, [3]University of Sussex

Color perception depends on the ability to extract differences in the spectral composition of light. In humans with normal color vision this process is initiated by the absorption of light by three classes of cone photoreceptors (L, M, S) followed by cone-opponent operations carried out by bipolar cells. But these color encoding mechanisms do not fully predict color appearance. Many attempts have been made to construct formal color spaces that capture uniform perceptual similarity relationships among colors, but discrepancies remain between the models and individual perception. Combining Magnetoencephalography (MEG) and psychophysical data obtained in the same set of participants, we examined the extent to which these discrepancies and individual differences can be accounted for by the geometry of the neural representation of color and its evolution over time. Unlike previous work, we collected neural responses from eight individuals who each viewed hundreds of unique colors presented in 16,200 color trials. The dense sampling of color space allowed us to reconstruct a fine-grained geometry of the neural representation of color in individual participants, with millisecond accuracy. We related these neurophysiological measurements to similarity judgements collected in the same participants. Using Variational Interpretable Concept Embeddings, we extracted similarity embeddings from the behavioral data to model the neural signal directly using regression models. Consistent with published results, we find that color information is present in the neural signal from approximately 70 ms onwards. Importantly, we discovered that the neural color-space geometries unfolded non-uniformly over time. In particular, the representational structures read out with MEG are best characterized across all participants by models of cone-opponent encoding mechanisms, especially early in the timeseries. These findings highlight the gap between theoretical color spaces and color perception and provide a novel avenue to gain insights into the subjective nature of perception.

53.465 FMRI RESPONSES TO COLOR AND OBJECTS IN THE VENTRAL VISUAL PATHWAY

Erin Goddard[1], Kateryna Marchenko[1], Colin Clifford[1]; [1]UNSW, Sydney

The human ventral visual cortex contains a mosaic of visual areas, including areas with higher level responses to objects, places and faces. It also contains several regions that are highly responsive to color. However, characterising how these areas vary in their response specialisations for basic visual features is complicated by the multidimensional nature of their responses. Recent work (Goddard & Mullen, Neuroimage, 2020), using Representational Similarity Analysis (RSA) of fMRI data showed two independent changes in response properties along the ventral visual pathway (V1, V2, V3, hV4, VO1, VO2): a decreasing dependence on spatial frequency, and an increasing dependence on the presence of color contrast. This previous work used simple stimuli (radial gratings) designed to preferentially stimulate different subcortical inputs to cortex. Here, we collected fMRI data (n=20) while participants viewed complex stimuli (filtered natural images) to target areas with higher-level responses. Across 12 stimulus conditions, all images included a range of spatial frequencies (either low, high or very high), were either colored or greyscale, and their content was either intact or phase randomised. Intact color images were either naturally colored (original) or unnaturally colored (color contrast reversed). We combined pairwise classification analyses with RSA to compare functional response preferences across the entire ventral occipital-temporal cortex, using the cortical parcellation of Glasser et al. (2016, Nature). Our data replicate the trends found previously for simple stimuli: along the ventral visual pathway, responses to filtered images showed decreasing dependence on spatial frequency, and increasing dependence on color. Across all areas, responses were modulated by whether the images were intact or phase-scrambled. However, there was an interaction between color responses and image structure, with higher-level visual areas showing greater evidence of color responses for intact than for phase-scrambled images. No regions showed clear evidence of responses that varied with color naturalness.

This work was supported by the Australian Research Council (DP220100747 to EG)

53.466 A POTENTIAL SPECTRAL CODE FOR HALLUCINATORY COLOR PERCEPTION

Lemona Xinxuan Zhang[1], Abigail Flowers[1], Wumeng Wang[1], Lucy P Somers[1], Jenny M Bosten[1]; [1]University of Sussex

Color experiences, typically elicited by colored light, can also arise in the absence of physically colored stimuli. Uniform-field (ganzfeld) intense flickering white light can evoke hallucinatory colors and forms. The mechanisms underlying this phenomenon, known as flicker-induced colors, are not well understood, but may provide insights into the neural basis of color perception. We aimed to investigate the correspondences between specific flicker frequencies and hallucinatory colors. In Experiment 1, participants completed a 2-session dichoptic color matching task. To one eye, we presented intense unstructured flicker using a custom LED stroboscope. To the other eye, we presented an adjustable colored disk on an LCD display. In session 1, participants explored the flicker frequency range (3-38 Hz), and adjusted the hue and saturation of the disk to match any (and all) hallucinatory colors they experienced at each frequency. In session 2, participants were presented with flicker at their previously matched frequencies, and were asked to match perceived colors again. A permutation-based analysis revealed that observed between-session color differences for same-frequency color matches were significantly smaller than permuted between-session color differences at randomly selected (unmatched) frequencies ($p < .01$). In Experiment 2, we investigated the frequency-specificity of flicker-induced colors by verbal report, avoiding dichoptic presentation. Over two sessions, participants reported hallucinatory colors while the flicker (3-38Hz) was presented binocularly. In concordance with the results of Experiment 1, we found a significant within-individual correspondence between flicker frequency and hallucinatory colors ($p < .01$). Our findings demonstrate reliable correspondences between the frequencies of flickering white light and induced illusory color experiences, compatible with a frequency-based component in the cortical coding of color.

This work was funded by the ERC grant 949242 COLOURCODE to JMB.

53.467 VISUAL RESPONSIVENESS IN AUTISM EXAMINED WITH FMRI, EEG AND PSYCHOPHYSICS

Daniela L Seczon[1], Hannah M Rea[1], Kristin M Woodard[1], Tamar Kolodny[2], Sara J Webb[1,3], Scott O Murray[1]; [1]University of Washington, [2]Ben-Gurion University of the Negev, [3]Seattle Children's Research Institute

Visual sensory processing in autism is often characterized by heightened sensitivity, which may be a product of increased neural responses to visual stimuli. To investigate these behavioral and neural responses, we integrated three measurement techniques—fMRI, EEG, and psychophysics. Twenty-seven autistic and thirty-one non-autistic adults completed experiments that used a consistent visual stimulus across all methods; bilaterally presented circular counterphase 6Hz-flickering checkerboards, in high (100%) and low (2%) contrast conditions. For fMRI and EEG sessions, participants fixated on the center of the screen and the bilateral checkerboards were presented following a design that alternated 10 seconds of stimulus with 20 seconds of blank-fixation. For fMRI, responses were measured in retinotopically-defined regions of early visual cortex including V1, V2, and V3. For EEG, steady-state visual evoked potentials (SSVEPs) amplitudes were extracted and analyzed for peak amplitudes across a pre-selected set of electrodes. During the psychophysics experiment, participants underwent three versions of a detection task: no adapter, low contrast adapter, and high contrast adapter conditions. Adapter conditions presented the flickering checkerboards before a subsequent lateralized target checkerboard in one of the two (left or right) hemifields. The target's contrast was adjusted following a staircase procedure, and behavioral data were extracted to estimate individual contrast detection thresholds. Stimulus contrast had a significant effect across experiments; the high contrast condition elicited greater fMRI and SSVEP responses, and the high contrast adapter led to higher contrast detection thresholds. Significant group differences emerged in SSVEP responses, with autistic individuals exhibiting higher amplitudes across both contrast conditions. Contrastingly, groups had comparable BOLD responses and contrast detection thresholds. This study provides some evidence that autistic individuals show heightened neural activity to visual input, as measured by SSVEPs, but these differences do not translate to differences in behavioral or hemodynamic responses.

NIH R01MH118847 and SFARI 647415 to SOM and SJW

53.468 HYPERSPECTRAL IMAGING OF WHOLE IMAGES FOR CLINICAL DIAGNOSIS: TEST CASE OF EDEMA

Cameron May[1] (cameron.may@nih.gov), Karthik Kasi, Leo Kobayashi, Bevil Conway, Joseph Pare; [1]National Institutes of Health

Hyperspectral imaging (HSI) is emerging as a powerful tool in medical diagnostics for its ability to capture and analyze data from across the electromagnetic spectrum. In contrast with traditional RGB imaging which is limited to just three-color channels, HSI allows for more color channels across a broader spectrum. HSI allows for the identification of unique spectral signatures based on how human tissues interact with different wavelengths of light. In medicine, HSI has been applied to distinguish between pathological and non-pathological states such as kidney stones or cancers. However, most studies focus on pixel-level classification (e.g., detecting cancer in specific pixels), rather than whole-image classification methods commonly used in convolutional machine learning (e.g., determining if an entire image shows signs of disease). We address this question by determining the extent to which whole image classification can discriminate between edematous or non-edematous legs as a proof of principle. Edematous legs are associated with different levels of perfusion which should manifest in different spectra observed with HSI. Here, we test this prediction. We used a push-broom hyperspectral camera (SOC-710, Surface Optics Corporation) to capture images and assess the classification potential of this method. 68 Images were captured from 27 emergency room patients over 4 months. We generated 5x5 pixel patches labeled as edematous or non-edematous. Using principal components analysis for unsupervised dimensionality reduction, we achieved a 66% classification accuracy with support vector machines, utilizing a linear kernel. This analysis demonstrates HSI's potential to expand access to medical imaging by offering a more affordable and portable alternative to conventional methods, particularly benefiting under-resourced clinics.

53.469 UNDERSTANDING THE IMPACT OF LIGHT STRESS ON DAILY LIFE: IMPLICATIONS FOR CATARACT TREATMENT

Jacob Harth[1] (jbh46589@uga.edu), Cameron Wysocky[1], Lauren Hacker[2], Billy Hammond[1], Lisa Renzi-Hammond[1]; [1]University of Georgia, [2]Emerging Vision Explorations (EVE) Research

Introduction: Age-related cataracts are a leading cause of vision problems in the U.S., and the primary treatment is surgery. Cataracts gradually impair vision, which can make everyday activities more difficult and dangerous, increasing the risk of accidents and limiting personal independence. While cataract surgery is often considered when individuals report difficulties in daily tasks, these reports usually do not account for the challenges people face in specific environments, such as bright sunlight or when driving at night with headlights. This study aimed to better understand how people with cataracts experience visual discomfort in real-world conditions and when cataract surgery might be needed. Methods: Data were aggregated from clinical trials conducted at the University of Georgia between 2017-2022. Participants (N=232; 18-78 years; M=42.39±14.65 years, 76% Female, 33% non-White) were divided into 4 groups: those without cataract (n=102), with early cataract (n=67), with advanced cataract (n=40), and those who recently had cataract surgery (n=23). Participants had otherwise good ocular health, as determined by an eye care provider. Glare discomfort was measured objectively by bioimaging the squint response and subjectively by participant rating of discomfort. Results: Participants with advanced cataracts had a significantly stronger squint response (M=8.16mm±4.75mm) than those without cataracts (M=4.10mm ±2.34mm, t=-6.79, p<.001), with early cataracts (M=4.55mm±2.60mm, t=-5.07, p<.001), and post-surgery (M=2.13mm±1.81mm, t=- 5.83, p<.001). Measures of subjective and objective glare discomfort were significantly correlated when controlling for age (r(209)=.34, p<.001) and iris lightness (r(209)=.33, p<.001). Conclusions: People with more severe squint responses to light may benefit most from cataract surgery, as it may reduce their heightened discomfort from glare (e.g., sunlight or headlights). This research expands our practical understanding of when cataract surgery may be necessary to improve a person's quality of life.

Temporal Processing: Neural mechanisms, models

Tuesday, May 20, 8:30 am – 12:30 pm, Pavilion

53.470 ENCODING FIDELITY OF FLICKER FREQUENCY DOES NOT DIFFER BY POST-RECEPTORAL DIRECTION OR PHOTOPIC LIGHT LEVEL

Ruby S. Bouhassira[1], Alan A. Stocker[2], Geoffrey K. Aguirre[1]; [1]Department of Neurology, University of Pennsylvania, [2]Department of Psychology, University of Pennsylvania

Temporal sensitivity functions recorded from ganglion cells and post-retinal sites differ for achromatic and chromatic modulations, and at high and low light levels. Visual experience statistics, however, suggest that the 1/f^2 distribution of wide-field temporal variation has the same form for achromatic and chromatic modulations (JH van Hateren et al 2002 J Neurosci; data not available across light levels). Efficient coding models therefore predict similar encoding fidelity across frequency for achromatic and chromatic modulations, despite differences in neural tuning. We tested this idea in a delayed flicker frequency estimation task, using luminance and L–M modulations against high and low photopic backgrounds. METHODS: On each of 1200 trials five participants were shown a 2s "reference" flicker (30° wide, uniform-field, bimodal sinusoidal temporal flicker, 1–32Hz uniform log random sampling). After a 2s delay, the participant adjusted the frequency of a "test" flicker to match their memory of the reference (starting point randomly selected within ±7dB of the reference). Positive feedback was provided for responses within ±1.5dB of the reference. Blocks of trials within a session alternated between LMS (33% contrast) and L–M (7.5%) modulations, and the six sessions alternated between a high (3400cd/m^2) and low (2cd/m^2) luminance background. RESULTS: Variance in estimation responses was a constant proportion of the reference frequency (except for low luminance L–M >20Hz), and did not differ with light level or post-receptoral direction (1.6dB). Responses were biased across frequency, with a negative slope and an intercept at ~2 Hz (i.e., rapid flicker was reported as slower); this bias did not differ with stimulus condition. These results are consistent with a Bayesian observer model constrained by efficient coding, reflecting a common 1/f^2 temporal frequency prior across our four stimulus conditions. Our results demonstrate a distinction between neural tuning functions and encoding for temporal variation.

R01EY036255

53.471 AS THE HEAD TURNS: DECODING HEAD ORIENTATION IN DYNAMIC CONTEXTS

Sebastian Montesinos[1], Lina Teichmann[2], Shruti Japee[3], Chris Baker[4]; [1]National Institute of Mental Health

Studies of visual perception frequently use rapidly presented static images, even though natural visual input is dynamic. The current study investigates the extent to which representations evoked during the presentation of static stimuli generalize to dynamic movies that involve those same stimuli. To address this question, we used magnetoencephalography (MEG) to compare the time course of brain activity as participants viewed static images and dynamic movies of human faces. The static images depicted faces with varying head orientations, while the movies showed these faces transitioning between orientations, passing through frames presented during the static image trials. We used time-resolved multivariate analysis approaches (i.e., classification and regression) to compare MEG signal patterns evoked by the different face orientations in both contexts. Results from both analysis approaches indicate that head orientation information can be reliably detected during static image viewing 100 ms after stimulus onset, with peak performance of the models occurring around 120–140 ms. Training models on MEG data evoked by static trials and testing them on MEG data evoked by movie trials, we found that head orientation information generalizes from static images to movies. However, we observed a temporal asynchrony between these trial types, with models trained on later parts of the static trials (300-400ms) best generalizing to movie trials. Together, these findings illuminate the similarities and differences in the temporal dynamics of processing across static and dynamic contexts, and allow for testing of predictive processing models of perception.

53.472 A DYNAMIC SPATIOTEMPORAL NORMALIZATION MODEL FOR CONTINUOUS VISION

Angus Chapman[1], Rachel Denison[1]; [1]Boston University

Motivation: How does the visual system process dynamic inputs? Both neural responses to and perception of a given stimulus are affected by temporal context from both past and future stimuli. Some effects of temporal context have been modeled using temporal normalization—the divisive suppression of neural activity by that occurring at other points in time. However, existing models do not compute responses in real-time, limiting biological feasibility, and do not pool suppression across neurons as is common in static normalization models. Here we ask whether the effects of temporal context on neural responses can be captured by a unified spatiotemporal receptive field structure that implements divisive normalization across space and time. Methods: We developed a dynamic spatiotemporal normalization model (D-STAN) that implements temporal normalization through excitatory and suppressive drives that depend on the recent history of stimulus input, controlled by separate exponential temporal windows. D-STAN extends on previous models, using a recursive neural network architecture with real-time simulation of sensory processing and decision-making, with spatiotemporal pooling of suppressive drives that allows for effects of temporal context both forward and backward in time. Results: Reverse correlation analysis of D-STAN's sensory responses uncovered effective temporal receptive fields that followed a half "Mexican hat" profile, similar to empirical findings. This response profile was not built directly into D-STAN but emerged from the interaction between the excitatory and suppressive windows and the normalization computation. D-STAN also reproduced several non-linear properties of neural responses that depend on temporal context, including subadditivity, repetition suppression, and backward masking. Finally, D-STAN predicted changes in perception, capturing bidirectional contrast-dependent suppression between stimuli at different times. Conclusions: Temporal normalization within a population of neurons with spatiotemporal feature tuning can account for a wide range of neural and behavioral effects. D-STAN is a step toward dynamic movie-computable models for continuous vision.

Startup funds from Boston University to R.D.

53.473 SERIAL DEPENDENCE OCCURS AFTER MOTION REPULSION

HOYEON YOON[1] (hyyoon@unist.ac.kr), JONGMIN MOON[2], OH-SANG KWON[1]; [1]Ulsan national institute of science and technology, [2]The University of Texas at Austin

Despite numerous studies on serial dependence in visual perception, the locus of the serial dependence within the perceptual process remains unclear. Moon and Kwon (2022) proposed that repulsive serial dependence occurs at an early encoding stage, while attractive serial dependence arises at a later estimation stage of visual processing. This hypothesis was based on their finding that the estimation of the current stimulus is repelled away from the previous stimulus but attracted toward the previous response. In contrast, Cicchini et al. (2021) demonstrated that attractive serial dependence occurs prior to surround tilt illusion, suggesting that attractive serial dependence occurs at an early stage of processing. Here, we investigated serial dependence using a motion repulsion stimulus, in which two directions of random dot motion move transparently. Motion repulsion exaggerates the perceived difference between adjacent motions and is known to occur at an early stage of motion processing (Grunewald, 2004). To distinguish the two sets of random dots, we used color. On each trial, participants were informed of the target color before stimulus presentation and were asked to report the motion direction of the target. A total of 14 participants completed 1533 trials. As expected, we observed robust motion repulsion and attractive serial dependence. More importantly, we tested whether serial dependence occurs before or after motion repulsion. This was achieved by examining whether the center of serial dependence depends on the magnitude of motion repulsion across participants and conditions. Statistically significant positive correlations ($r=0.2878$, $p=0.0157$) were found between the magnitude of motion repulsion and the pattern of serial dependence, indicating that serial dependence occurs after motion repulsion. These results align with the hypothesis that repulsive serial dependence reflects efficient coding at early stages, while attractive serial dependence reflects perceptual inference at later stages of visual processing (Moon & Kwon, 2022).

This work was supported by National Research Foundation of Korea: [Grant Number NRF-2020S1A3A2A02097375]

53.474 INVESTIGATING TEMPORAL INTEGRATION IN PRIMATE VISUAL CORTEX USING NATURALISTIC VIDEO STIMULI

*Nathan C L Kong[*1], Akshay V Jagadeesh[*2], Margaret S Livingstone[2]; [1]University of Pennsylvania, [2]Harvard Medical School*
[*]Equal Contribution

Animal behavior is supported by dynamic neural representations that encode information from continuous visual experience. The vast majority of object vision research has examined representations of static images and behaviors that can be performed with static stimuli, such as object categorization. However, numerous behaviors, such as action recognition, causal attribution, or invariant object representation learning, may require temporally integrating information over continuous visual experience. It remains unknown whether the ventral visual cortex, widely thought to support object vision, is capable of such temporal integration. Here, we investigate the role of ventral visual cortex in encoding and integrating information over time. Using chronically-implanted microelectrode arrays in visual cortex of macaque monkeys, we collected 30 hours of neural responses from inferior temporal (IT) cortex while subjects viewed 960 three-second videos from the Moments in Time dataset. These naturalistic videos contain a wide variety of objects and actions and were presented either in their original form, in reverse, or statically (single video frame). We found that action decoding on time-averaged neural responses to original videos was significantly above chance, but not significantly different from decoding performance using responses to static frames, suggesting decoding was driven primarily by visual feature differences. To assess temporal integration in IT cortex, we evaluated how well responses at each timepoint predict every other timepoint and found that responses were only significantly predictive of each other within a relatively narrow temporal band. Finally, we evaluated a variety of computer vision models, including static image models and video models that integrate over video frames, and found substantial gaps between all models' neural predictivity and the noise ceiling. Our results provide preliminary evidence that temporal integration in IT is limited and narrow.

53.475 UNRAVELING THE GEOMETRY OF NEURAL REPRESENTATIONAL DYNAMICS IN RAPID VISUAL PROCESSING

Zirui Chen[1] (zchen160@jhu.edu), Leyla Isik[1], Michael F. Bonner[1]; [1]Johns Hopkins University

Human vision excels at processing rapidly changing inputs while maintaining coherent representations of the preceding content. Previous studies have shown that stimulus features can be decoded from the brain for hundreds of milliseconds after presentation even as new visual inputs arrive, but decoders fit at a certain time point do not generalize across the entire decodable range. This suggests that there are dynamic shifts in the representational structure over time. However, the nature of these transformations remains poorly understood. To address this gap, we analyzed the THINGS EEG2 dataset, a large-scale collection of human neural responses to rapidly presented visual objects, and we investigated how representations of behaviorally relevant object features unfold over time. Specifically, we decoded these features via regression at each time point and tracked the temporal evolution of decoder weight vectors, allowing us to characterize how latent axes that represent stimulus features transform over time. Our analyses revealed two key findings. First, the transformations in feature representations are not merely the result of signal suppression by new visual inputs. Instead, these transformations reflect rotations of latent axes that are consistent across trials, enabling the same feature to be encoded along orthogonal axes at different time points. Second, we found that there are common representational transformations over time that are shared by multiple clusters of distinct features. These findings suggest that the human brain employs structured and reliable modes of temporal transformation to encode information from a rapid succession of visual inputs. This work offers promising directions for understanding representational dynamics in biological and artificial systems through time.

53.476 SPATIOTEMPORAL INTEGRATION IN THE HUMAN FOVEA

Jingyi He[1], Maxwell J. Greene[1], William S. Tuten[1]; [1]Herbert Wertheim School of Optometry and Vision Science, University of California, Berkeley, Berkeley, CA, USA

Adaptive optics platforms enable the delivery of small, diffraction-limited stimuli to known retinal locations. Attempts to relate the resultant cone activations to perception would benefit from a deeper understanding of how the intervening post-receptoral mechanism(s) integrate photoreceptor signals over space and time. The double-pulse technique has been widely used to estimate the temporal impulse response function (IRF) of the human visual system. In this study, we examined how the detectability of pulse pairs depended on stimulus-onset asynchrony (SOA). An adaptive optics scanning laser ophthalmoscope with two narrowband primaries (543 nm and 680 nm; "green" and "red", respectively) was used to deliver increment flashes with higher-order and chromatic aberrations corrected. Stimulus pulses subtended either 6'×3' (1 ms duration) or 24'×12' (4 ms duration) and were presented in either isochromatic ("green-green") or heterochromatic ("red-green") configurations. Heterochromatic pulse pairs were presented with the intensities of the red and green primaries scaled according to single-pulse threshold measurements. A white background (2.2°×1.1°) was provided by an auxiliary display. Detection thresholds for SOAs ranging from 0 to 200 ms were obtained using a 2IFC task in three subjects. For the smaller stimulus size, threshold-versus-SOA curves were similar for both color conditions, exhibiting near-complete summation at the shortest SOA and rising monotonically to an intensity corresponding to ~67% of the single-flash threshold. By contrast, for the 24'×12' isochromatic pulse pair, the highest thresholds were observed at an SOA of 50 ms, suggesting the emergence of an inhibitory lobe in the IRF as stimulus size increases. No such inhibition was observed for the red-green 24'×12' pulse pair. Our results are broadly consistent with prior studies demonstrating that the form of the IRF depends on the spatio-chromatic properties of the stimulus, and can be incorporated into computational models of early visual processing in the human fovea.

This work was supported by the Air Force Office of Scientific Research under award numbers FA9550-20-1-0195 and FA9550-21-1-0230

53.477 MODELING THE FMRI NEURAL TEMPORAL RESPONSE ESTIMATION PROCEDURE

Christopher Tyler[1]; [1]Smith-Kettlewell Eye Research Institute

The procedure for neural population receptive field (pRF) estimation from non-invasive functional Magnetic Resonance Imaging signals (Dumoulin & Wandell, 2008) is now well established in many fields of application. We proposed parallel procedure for estimating the local neural temporal population response (Tyler & Likova, 2009), which allows estimation local neural response to its native temporal resolution from fMRI signals. The procedure involves modeling the temporal structure of the neural population response (NPR). This procedure relies on the fact that neural responses are inherently rectifying, carrying no response for negative, or hyperpolarizing, intracellular neural signals. In the population, therefore, the negative lobes of the neural temporal responses are carried by a positive signal in off-cells. Since both on- and off-cells (or Gabor RFs having all phases of the carrier wave) have the same sign of energy requirements, the BOLD metabolic response to local neural activation is inherently rectifying, providing the nonlinearity required for estimation of the temporal NPR kernel from BOLD signals as a function of stimulus duration. We model the sensitivity of the procedure as a function of the NPR parametrization and fMRI signal-to-noise ratio. This shows that the larger the difference in time constant between the BOLD MRK and the NPR kernel, the better the NPR kernel estimation. The estimable parameters are the neural transient time constant and gain, the transient/sustained response ratio, and the off-response polarity and gain. The results show that the typical fMRI signal-to-noise ratio is sufficient to determine the time constant for a typical neural population transient of 0-100 ms to a resolution of ~25 ms. However, there is no simple relationship between the summation curve property and the underlying nonlinearities and time constants. To determine the specific parameter values, the parametrized model NPRK was optimized to the full dataset for V1 and hMT+.

53.478 FORWARD AND BACKWARD MASKING OF NATURALISTIC TEXTURE

Laura Palmieri[1] (lp3006@nyu.edu), Timothy D. Oleskiw[1,2], J. Anthony Movshon[1]; [1]New York University, [2]University of Regina

The detection and discrimination of spatial stimuli is affected by overlaid masks presented before or after the target stimulus. When the masking involves contrast detection or simple discrimination, masking is thought to depend on contrast signals like those in primary visual cortex. We have now extended this approach to signals that rely on downstream processing. For this we used the discrimination of naturalistic texture statistics, which is known from neuronal recordings in macaque to depend on signals in areas downstream from V1, notably V2 and V4. Subjects discriminated "naturalistic" textures created from the Portilla-Simoncelli model from phase-randomized "noise" textures. We interpolated the model parameters between matched noise and naturalistic textures to vary "texture coherence" – 0% for noise textures and 100% for textures with all the P-S statistics. We used a 3-choice oddity task to measure performance. Subjects discriminated textures of variable coherence, set using a staircase procedure, to define coherence threshold for brief (40 ms) targets. These were presented with or without a zero-coherence 100 ms mask that either preceded or followed the target with a gap between 0 and 50 ms. When the target followed the mask ("forward masking"), coherence thresholds were modestly elevated (1.2–2x); masking was strongest for the 0-gap condition. When the target preceded the mask ("backward masking"), threshold elevation was greater (2.5-4x), and was maximal for a gap of 25 ms. Performance was similar for a simultaneous 3-choice task using near-peripheral viewing, and a sequential 3-choice task using central viewing. Our results can be accounted for using a delayed normalization model to describe neural response dynamics, combined with temporal crowding to capture backward masking. They are also compatible with data we have obtained using a similar paradigm from single- and multi-unit neuronal recordings from area V4 of awake, fixating macaque monkeys.

53.479 TESTING PREDICTING CODING USING MEG-BASED DYNAMIC RSA

Marisa Birk[1] (marisa.birk@unitn.it), Annika Tesio[1], Ingmar de Vries[1,2], Moritz Wurm[1]; [1]CIMeC - Center for Mind/Brain Sciences, University of Trento, Italy, [2]Donders Institute, Radboud University, Nijmegen, The Netherlands

According to the theory of predictive coding, the brain constantly generates predictions about sensory input. These predictions are compared against the incoming signal, with accurate predictions suppressing expected bottom-up information. In this framework, stimuli that are more predictable lead to stronger predictive representations and more suppressed sensory input. By contrast, unpredictable stimuli result in weaker top-down predictions and stronger, less suppressed representation of the bottom-up input. Here, we tested these hypotheses using dynamic Representational Similarity Analysis (dRSA). This approach uses temporally variable models that capture the representational content of a dynamic event at each individual time point. This allows testing when the brain represents a stimulus at a given time point in relation to its real-time occurrence, i.e., in a predictive or lagged manner. This provides a means to disentangle top-down predictions and bottom-up sensory input. During magnetoencephalography (MEG), 22 participants watched videos of a moving dot. We manipulated the predictability of the dot's trajectory by varying the probability that the dot would change its direction, creating four conditions that ranged from fully predictable to highly unpredictable. dRSA was applied to capture neural representations of the dot's position and movement direction. For all predictability levels, we found that dot position was represented with a consistent lag of 120 ms. By contrast, the dot's direction was represented in both predictive (-1100 to -300 ms) and lagged (900 to 1700 ms) manners. Importantly, the strength of dRSA for both predictive and lagged representations systematically increased with higher predictability. Thus, while top-down predictions became more pronounced with increasing predictability, we did not observe a suppression of lagged (potentially bottom-up) representations, challenging a key assumption of predictive coding. Instead, our findings suggest that increased predictability sharpens the representation of bottom-up sensory input by enhancing the selectivity for relevant stimulus features.

53.480 LUMINANCE FLICKER BOOSTS SENSITIVITY TO SMALL, DIFFRACTION-LIMITED, INCREMENT SPOTS

JT Pirog[1] (jpirog@berkeley.edu), Jingye He[1], William Tuten[1]; [1]University of California, Berkeley

Psychophysical evidence shows that exposure to luminance flicker reduces sensitivity to subsequent luminance modulations, likely reflecting contrast adaptation that has been observed physiologically in subcortical magnocellular neurons. Here, we investigate how luminance flicker influences the visibility of small, diffraction-limited flashes, subtending <10 foveal cones. We used an adaptive optics scanning laser ophthalmoscope (AOSLO) to present narrowband (543 nm, "green") increments of various diameters (1 to 23 arcmin) for 200 ms to the fovea following the delivery of isochromatic luminance (orange) flicker comprised of cycling the red (680 nm) and green AOSLO primaries in phase at 3.75 Hz. The initial adaptation period spanned 30 seconds, with one second of top-up adaptation inserted between each one-second trial. We measured detection thresholds across a range of stimulus onset asynchronies (SOAs; 33 to 500 ms) in 3 subjects. Control measurements were also obtained without exposure to flicker. We found that the visibility of large spots (23 arcmin) was reduced at all tested SOAs following flicker adaptation. With small spots (1 arcmin), a different pattern was observed: visibility after adaptation to identical flickering conditions was similar to control at the shortest and longest SOAs, and sensitivity was improved by ~25% from control over a range of intermediate SOA values (100 to 200 ms). Our results are corroborated by recent findings from retinal physiology showing opposing patterns of SOA-dependent sensitivity changes in magnocellular and parvocellular neurons following exposure to luminance flicker. Future work will include measurements of chromatic identification thresholds following flicker to further parse the contributions of the chromatic and luminance channels to processing very small stimuli approaching the neural grain of the foveal receptor array.

This work was supported by the Air Force Office of Scientific Research (FA9550-20-1-0195, FA9550-21-1-0230), the National Institutes of Health (R01EY02359, T32EY007043), the Hellman Fellows Fund, and the Alcon Research Institute.

53.481 REACTIVATION OR ACCUMULATION? **EXPLORING WORKING MEMORY'S ROLE IN EVENT** SEGMENTATION

Berna Güler[1] (bernaaguler@gmail.com), Sümeyye Karahamza[2], Yağmur Damla Şentürk[3], Oded Bein[4], David Clewett[5], Eren Günseli[6]; [1]Sabanci University, [2]Sabanci University, [3]Sabanci University, [4]Weill Cornell Medical College, [5]University of California, UCLA, [6]Sabanci University

Working memory (WM) plays a critical role in constructing the segmented structure of episodic memories, a process known as event segmentation, where experiences are divided into meaningful units. However, its specific contributions remain unclear. This study evaluated two competing perspectives: (1) the accumulation view, which posits that sequences of items are stored in WM until an event boundary is encountered, and (2) the reactivation view, which suggests that WM stores items temporarily but reactivates them at boundaries. Participants viewed sequences of images and sounds organized into alternating mini-blocks: six items from one category followed by two from another (e.g., 6-2-6-2-6-2), with category switches marking event boundaries. After each block of 36 items, participants completed temporal order and sequential memory tasks to assess event segmentation. The sequential memory accuracy was higher for within-event items compared to across-event items, reflecting event segmentation. The EEG indices of WM storage, which were alpha power negativity and contralateral delay activity (CDA), provided stronger support for the reactivation view: Alpha power was more negative at boundaries signaling the end of main events (6 items) than boundary events (2 items), indicating larger reactivated memory loads following longer events. Evidence for the accumulation view was mixed: there was moderate evidence for a CDA increase and no evidence for an alpha negativity increase across items within events. These results suggest that while WM may accumulate items to some extent during events, reactivation at event boundaries plays a more prominent role. This dynamic interplay between WM and long-term memory processes appears critical for constructing the segmented temporal structure of episodic memories.

53.482 APERIODIC EEG ACTIVITY CORRELATES WITH OCCIPITAL GLUTAMATE FROM 7 TESLA MRS

*Aislin Sheldon[*1] (sheld173@umn.edu), Hannah Moser[*1], Kamar Abdullahi[1], Karly Allison[1], Carter Mulder[1], Kyle Killebrew[1], Samantha Montoya[1], Scott Sponheim[1,3], Małgorzata Marjańska[2], Michael-Paul Schallmo[1]; [1]Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis, MN, [2]Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, Minneapolis, MN, [3]Veterans Affairs Medical Center, Minneapolis, MN [*]Equal Contribution*

Oscillatory activity across the cortex is linked to multiple cognitive processes, including visual perception. However, recent studies which separate oscillatory and aperiodic components of electrophysiological activity have highlighted the functional role of aperiodic activity. Studies in animal models suggest the aperiodic slope of electrophysiological power spectra reflects the local ratio of excitatory:inhibitory synaptic transmission. We hypothesized that shallower aperiodic slopes from human EEG, measured over the occipital lobe, would correlate with higher MRS measures of occipital glutamate, reflecting cortical excitation. EEG was collected from 26 healthy adults during eyes-open rest. Power density spectra from occipital electrodes were analyzed using the fitting oscillations and one over f (FOOOF) toolbox, yielding a measure of the slope of the aperiodic (1/f-like) component for each participant. In the same participants, we separately acquired magnetic resonance spectroscopy (1H MRS) data at 7 tesla with a stimulated echo acquisition mode (STEAM) sequence (TE = 8 ms, TR = 5 s). Occipital MRS data were collected in a single voxel (30x30x18mm3) centered over the calcarine sulcus while participants completed a central fixation task on a gray screen. MRS data were processed using the matspec toolbox in MATLAB and metabolites were quantified using linear combination of model (LCModel) software. Across participants, we observed a significant negative correlation between aperiodic slopes and glutamate concentrations (r24 = -0.47, p = 0.016), whereas the correlation with GABA levels (a marker for neuronal inhibition) was not significant (r24 = 0.17, p = 0.4), nor NAA - a metabolite that is assumed to remain unchanged with visual stimulation, NAA (r24 = -0.18, p = 0.4). Our results show that occipital glutamate concentrations relate to aperiodic activity during fixation, and support the use of aperiodic electrophysiological activity as a marker for occipital excitation.

K01 MH120278

TUESDAY AFTERNOON POSTERS IN BANYAN BREEZEWAY

Spatial Vision: Models

Tuesday, May 20, 2:45 – 6:45 pm, Banyan Breezeway

56.301 GLOBAL PERCEPTUAL ORGANIZATION ATTACHES TARGET POSITION TO THE REFERENCE FRAME IN THE FRAME EFFECT.

Xu Jiaheng[1,2,3], Zuo Zhentao[1,2], Zhang Ruifan[1,2,3], Xue Rong[1,2], Zhou Tiangang[1,2,3]; [1]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, China

Human visuospatial perception is profoundly influenced by the surrounding frame. For instance, in the frame effect (FE), the perceived positions of two physically aligned targets are shifted apart as if anchored in the moving frame (Özkan et al., 2021; Cavanagh et al., 2022). However, it remains unclear how the visual system attaches a target to the reference frame. Here, we hypothesize that the target is bound to the reference frame with the perceptual organization. In Experiment 1, a rectangular frame (width: 15°) moved back and forth below the fixation point, and two targets (white disks) flashed at the same spot but inside or outside the frame, depending on the motion path (horizontally covering 0° or -7.5° or -15° ± path length/2). This generated three organization conditions, both/either/neither inside the frame, across different path lengths. FE was quantified as the reported separation between two flashed targets. We found a systematically maximal FE in both-inside condition. Moreover, Experiment 2 yielded a maximum FE when the targets both flashed on a bar-shaped frame, compared with either-on and neither-on conditions, which collectively revealed a facilitation by the closure and connectedness organizations. To control the target's distance to frame's center (Shams et al., 2024), Experiment 3 varied distances while maintaining perceptual organization, and Experiment 4 fixed the distance but altered organization conditions. Both results confirmed that perceptual organization, rather than a certain distance, enhanced FE. In Experiment 5, we examined whether organization lacking a global configuration, e.g., color similarity, influence FE equally. The results failed to reproduce the pattern of the maximal FE when the targets were both isochromatic to the frame. Overall, these findings suggest that global perceptual organization, such as closure and connectedness, plays a significant role in frame-induced position shifts by binding the targets to the reference frame.

This project supported by Ministry of Science and Technology of China (2023YFF0714204, 2022ZD0211901), National Key Laboratory of Human-Machine Hybrid Augmented Intelligence, Xi'an Jiaotong University (No.HMHAI-202414), and the CAS (2021091, YSBR-068, 2024GZL003).

56.302 STIMULUS-DEPENDENT VARIABILITY IN POPULATION RECEPTIVE FIELD MAPPING

Christian Windischberger[1], David Linhardt[1], Siddharth Mittal[1], Michael Woletz[1]; [1]High Field MR Center, Center for Medical Physics and Biomedical Engineering, Medical University of Vienna

Population receptive field (pRF) modelling quantifies the organizational properties of visual cortex areas by mapping locations in the visual field to corresponding regions in the cortex. While this method is powerful and reproducible (Linhardt et al, 2022), measurement parameters and analysis choices can significantly influence the results. One crucial factor is the choice of visual stimulus. Our laboratory has developed a fast pRF analysis framework that enables parallel simulation of thousands of artificial time courses (Mittal et al. 2024). Using this framework, we simulated the variability of two commonly used stimuli (moving bar and wedge/ring) and compared them to experimental results (Linhardt et al. 2021). We simulated ground-truth positions distributed across the visual field and generated noisy time-courses using Gaussian white noise. For each position, we obtained pRF parameters (x, y, sigma) and analyzed their distributions. To evaluate parameter estimation accuracy, we calculated root-mean-square errors across different locations in the visual field. Error distributions revealed systematic patterns across the visual field. Both stimulus variants showed lower errors near the fovea and increased errors toward stimulus boundaries, an effect attributable to reduced encoding power when pRFs extend beyond the stimulus edge. The bar stimulus produced higher eccentricity errors in peripheral regions compared to the wedge/ring, while the latter showed higher errors in central regions. Interestingly, pRF size parameter errors followed an opposite trend, with the wedge stimulus showing higher peripheral errors and the bar stimulus showing higher central errors. These patterns suggest that each stimulus type has distinct advantages for different regions of the visual field. These simulation results provide an objective method for comparing visual stimulation paradigms in pRF mapping by quantifying parameter variability across the visual field. This approach offers valuable insights for optimizing experimental design in visual neuroscience research and understanding the inherent limitations of different stimulus configurations.

This research was funded in whole or in part by the Austrian Science Fund (FWF) [https://www.doi.org/10.55776/P35583]

56.303 VALIDATING RUN-TO-RUN VARIABILITY SIMULATIONS FOR POPULATION RECEPTIVE FIELDS (PRF) MAPPING

Siddharth Mittal[1], Michael Woletz[1], David Linhardt[1], Christian Windischberger[1]; [1]Medical University of Vienna

Population receptive field (pRF) modelling is a widely used technique for fMRI-based visual field mapping. While pRF mapping has been shown to be a reproducible method, inter-run variability is still present and not consistent across the visual field. This study aims to generate pRF estimations from simulated data and validate the variability observed by demonstrating its similarity to empirical data. Empirical retinotopy data were collected using a Siemens 3T Prisma Fit scanner from two subjects across five sessions, with six runs per session, resulting in 30 pRF estimations per subject using a 2D Gaussian model. The empirical pRF estimations exhibited systematic distributions dependent on visual field location (μ_x, μ_y, and σ). While these distributions showed a circular Gaussian-like form in foveal regions, distributions in more peripheral locations were ellipsoidal with increased width in eccentricity direction. To replicate this behaviour, we simulated visual receptive fields uniformly distributed across 51x51 spatial locations (x, y) with varying sizes (σ). Using the same stimulus as employed in the empirical data, we generated 5000 simulated BOLD responses per position, incorporating white noise to achieve a contrast-to-noise ratio (CNR) of 1 to 4. This large-scale dataset was processed using GPU-empowered pRF mapping software (GEM-pRF; Mittal et al., 2024) to estimate the probability distribution of pRF parameters for each simulated field. Maximum likelihood estimation was employed to find the best matching simulated distribution to each empirical voxel's estimated pRF results. The significance of these matches was validated using the multivariate Kolmogorov-Smirnov (Naaman, 2021) test, confirming that simulated variabilities reliably reflect empirical variability. This work demonstrates that simulations can accurately replicate the run-to-run variability of pRF parameter estimations. These findings represent a significant step forward in enabling the development of optimised visual stimuli to minimise variability.

This research was funded in whole or in part by the Austrian Science Fund (FWF) [https://www.doi.org/10.55776/P35583]

56.304 THE SPATIAL AREA THAT MASKS A STIMULUS IS LARGELY INDEPENDENT OF SPATIAL FREQUENCY AND ECCENTRICITY

Aqsa Hassan[1], Heiko H. Schütt; [1]University of Luxembourg

Masking experiments are an important basis of our understanding of early visual processing. Masking data allows us to constrain the non-linear interactions between spatial frequency and orientation-tuned channels like divisive normalization. It is well established how these interactions work across spatial frequencies and orientations. However, it is not well understood how these interactions work across space, and we lack data for different mask shapes and sizes. Therefore, we set up a contrast masking experiment to test how strongly different sizes of masks increase the contrast threshold. We use small Gabor targets (SD=0.5 cycles) centrally superimposed over a 15% contrast grating mask of the same spatial frequency and orientation. The circular aperture of the mask is composed of a flat top and a gradual border with the same fall off as the Gabor target. We vary the size of the flat top of the mask for changing its size. The spatial frequencies of both stimuli are 1, 2, or 4 cycles per degree, which are displayed at 5° and 10° eccentricity. The observer's task is to judge on which side the test stimulus is superimposed on the mask. In our preliminary experiments, contrast thresholds increase with increasing mask size and saturate at some point as expected. However, the shape of the masking curve is surprisingly independent of eccentricity and spatial frequency, saturating always at a radius of about 1° visual angle. We plan to extend our results with a foveal condition and masks elongated along or orthogonal to the target grating by the time of the conference. Our results have important consequences for our understanding of early visual processing as the prevalent expectation based on neural recordings is that normalization pools are proportional to receptive field sizes and should thus be strongly eccentricity and spatial frequency dependent.

56.305 ORIENTATION TUNING OF CONTRAST-SENSITIVE MECHANISMS: INSIGHTS FROM INDIVIDUAL DIFFERENCES

David H. Peterzell[1,2] (davidpeterzell@mac.com), Omar Bachtoula[3], Ichasus Llamas-Cornejo[3], Ignacio Serrano-Pedraza[3]; [1]Fielding Graduate University, Santa Barbara, California, [2]JFK School of Psychology, National University, Pleasant Hill, California, [3]Department of Experimental Psychology. Faculty of Psychology. Complutense University of Madrid, Madrid, Spain

Orientation selectivity for luminance-varying patterns is well-supported by anatomical, physiological, and psychophysical studies. Most psychophysical evidence has come from subthreshold summation, adaptation, and masking experiments, which suggest the existence of narrowly tuned orientation mechanisms with bandwidths ranging from 6° to 30°. In this study, we used factor-analytic dimension reduction techniques to investigate the presence of independent orientation mechanisms. Using Bayesian staircases, 36 participants performed a contrast detection task with gratings of 1.5 c/deg at ten orientations (0° to 90°), spatially windowed by a Butterworth 2D function with a diameter of 11°. Preliminary results revealed no significant differences in contrast thresholds across orientations. Instead, contrast thresholds exhibited strong intercorrelations, suggesting the presence of a broadly tuned factor or factors mediating sensitivity across orientations. A two-factor analysis (PCA with oblique Promax rotation) identified two highly intercorrelated factors: one peaking at vertical-to-oblique gratings (F1: 0° to ~50°) and the other at oblique-to-horizontal gratings (F2: ~50° to 90°). These findings are consistent with the presence of broad orientation-tuning mechanisms but do not support the existence of independent, narrowly tuned mechanisms. Split-half and test-retest analyses replicated these results, although test-retest reliability was not robust. More broadly, these findings challenge the traditional view that contrast sensitivity is determined by narrowly tuned, independent orientation-selective mechanisms. Instead, they support the idea that orientation and spatial frequency sensitivity are mediated by overlapping, interdependent mechanisms with broad tuning. However, the limited sample size and moderate reliability highlight the need for caution in interpreting these results. Future research with larger samples and enhanced reliability testing is essential to validate and extend these findings.

Supported by grant PDI2021-122245NB-I00 from Ministerio de Ciencia e Innovación (Spain) to ISP

56.306 **A "BACK-POCKET" MODEL OF V1 ORIENTATION** PROCESSING PREDICTS CURVATURE BLINDNESS

Steven Dakin[1] (s.dakin@auckland.ac.nz), Selassie Tagoh[2]; [1]University of Auckland, [2]University of Auckland

Takahashi's popular "curvature blindness" illusion (2017, i-Perception 8(6): 1–12) reveals that modulation of luminance along a sinusoidal-shaped line can dramatically alter its perceived shape, (notably) causing smooth contours to appear to have sharp corners. We asked how well a simple model of V1 orientation processing might predict this striking illusion. We presented a series of test stimuli: sinusoidal lines whose luminance was modulated along the line according to a square wave with a similar period to the sinusoid but with a variety of phase offsets (compared to the line shape). Six observers then adjusted the shape of a larger, white sinusoidal reference line to match the appearance of a given test. Observers did this by adjusting the amplitude and "sharpness" (0: pure sinusoid, 1: pure triangle wave) of the reference. The sharpness could be adjusted independently for the portions of the reference to the left and right of its peak (allowing us to capture perceived asymmetry). As well as replicating the original finding – that luminance transitions at the peaks/troughs of the curve (90 deg phase offset) make it appear much sharper – we report that these transitions increase the perceived amplitude of the sinusoid (by at least 25%). Stimuli with 45 and 135 deg phase-offsets induce asymmetry in sharpness around the peaks and troughs of the test patterns. A simple "back-pocket" model of V1 orientation processing– using oriented Gabor filters operating in a winner-take-all manner at a given spatial location - predicts perceived changes in sharpness, asymmetry and amplitude measured in the matching experiment. We conclude that this powerful shape illusion emerges naturally from V1 processing without the need for dedicated explanations involving e.g. segmentation.

Supported by the Leitl Charitable Trust

56.307 ALPHA-BAND PHASE MODULATES PERCEPTUAL SENSITIVITY BY REDUCING INTERNAL NOISE

April Pilipenko[1], Jason Samaha[1], Alexandra Mcgowan[1], Jacob Chaudhry[1], Vrishab Nukala[1], Emily Lincoln[1], Aishu Narayanan[1]; [1]UC Santa Cruz

Trial-to-trial variability in alpha-band neural oscillations (8–13 Hz) have previously been linked to variability in conscious visual perception. However, the functional role of alpha phase in visual perception remains debated, with prior studies reporting effects of pre-stimulus phase on detection (i.e., hit rates) but also null results. Methodological differences such as stimulus properties (e.g., duration, size, location), and low trial counts may contribute to these inconsistencies. Additionally, previous studies have not distinguished whether alpha phase modulates perceptual sensitivity (d') or criterion, using signal detection theory (SDT). To address these gaps, we used a Yes/No detection paradigm with a 8ms vertical Gabor target (2 cycles per degree) embedded in filtered noise. On half of the trials only the noise patch was presented. Participants reported whether they perceived the target in the noise or only the noise. Each participant (n=5) completed 6,020 trials across 4-5 EEG sessions. We observed a significant effect of pre-stimulus alpha phase on d' for most individual subjects and at the group level with no accompanying effect on criterion. Examining detection across the alpha cycle indicates that these sensitivity modulations arise from an inverse phasic relationship between hit and false alarm rates, consistent with a reduction in internal noise during optimal alpha phases. This contrasts with a multiplicative gain model, which would predict a constant false alarm rate across phases and only a boost in hits. These findings suggest that alpha phase enhances perceptual sensitivity by decreasing internal noise, rather than amplifying the signal. Lastly, we created classification images of the optimal and suboptimal phases. These images showed small boosts in the sharpness of spatial frequency and orientation tuning during the optimal alpha phase, suggesting that alpha phase may modulate sensory tuning towards relevant stimulus properties.

56.308 SPATIOFEATURAL RECEPTIVE FIELD MODELING OF PRIMATE IT CORTEX NEURONS

Akshay V Jagadeesh[1], Sohrab Najafian[1], Michael J Arcaro[2], Margaret S Livingstone[1]; [1]Harvard Medical School, [2]University of Pennsylvania

Neurons in primate visual cortex encode visual information within localized regions of visual space, i.e., receptive fields. In early visual areas, spatial receptive fields are well-characterized using simple stimuli like bars or gratings. However, in high-level areas such as inferior temporal (IT) cortex, the complex feature selectivity of neurons makes mapping their spatial receptive fields more challenging. To understand how spatial and featural selectivity are integrated in IT cortex, we measured multiunit electrophysiological responses throughout macaque IT cortex. Monkeys passively viewed a set of naturalistic images, including faces, face parts, hands, objects, and scenes, presented in a grid spanning the central 17 degrees of visual space. Using a Gaussian receptive field model, we estimated reliable receptive field positions for face-selective units and observed shifts in receptive field position based on stimulus type. These shifts were explained by spatially non-uniform contributions of face parts (e.g., eyes, mouth), emphasizing the necessity of jointly modeling spatial and featural selectivity for accurately characterizing IT neuron response properties. To address this, we developed a novel patch-based deep neural network model that simultaneously estimates spatial and featural tuning of neurons. When tested on neural responses to natural scene images, this model identified diverse receptive field structures with spatially-dependent feature tuning, evidenced by differential feature encoding across receptive field subregions. Even for neurons in anterior IT cortex, classically thought to exhibit position invariance, the model uncovered differential tuning for specific face features at distinct spatial locations, such as preference for eyes in the upper visual field and mouths in the lower visual field. This approach demonstrates evidence for precise spatial coding of features in high-level visual cortex and provides a formal modeling framework to characterize the interaction between spatial and featural coding in brain areas with complex tuning properties.

Spatial Vision: Neural mechanisms

Tuesday, May 20, 2:45 – 6:45 pm, Banyan Breezeway

56.309 MODULATION OF BASELINE ACTIVITY AND ASSOCIATED RECEPTIVE FIELD DYNAMICS IN THE HUMAN VISUAL CORTEX VIA TRANSCRANIAL DIRECT CURRENT STIMULATION

Jeongyeol Ahn[1,2], Juhyoung Ryu[1], Sangjun Lee[3], Chany Lee[4], Chang-Hwan Im[5], Elisha P Merriam[2], Sang-Hun Lee[1]; [1]Seoul National University, Seoul, Republic of Korea, [2]National Institute of Mental Health, [3]University of Minnesota, [4]Korea Brain Research Institute, Daegu, Republic of Korea, [5]Hanyang University, Seoul, Republic of Korea

Receptive field properties in the human visual cortex are fundamental to sensory processing and perception. Recent advances in brain stimulation techniques have enabled non-invasive modulation of neural activity, offering new opportunities to explore how baseline neural activity influences receptive field properties. In this study, we examined the effects of experimentally modulated baseline activity using anodal transcranial direct current stimulation (tDCS) on population receptive fields (pRFs) in the visual cortex. Using fMRI responses to ring-shaped traveling-wave stimuli, we modeled pRFs for individual voxels and quantified key properties, including baseline activity, pRF width, and surround suppression strength. We previously found that anodal tDCS increased the baseline fMRI activity for a prolonged period. Further analyses revealed that the increment in baseline activity was positively correlated with a narrower width of pRF ($z = -5.19$, FDR-adjusted $p < 0.05$) and surround suppression strength ($z = 4.47$, FDR-adjusted $p < 0.05$). These findings suggest a possibility that baseline neural activity modulation contributes to spatial tuning by strengthening inhibitory mechanisms, such as surround suppression, within the visual cortex. This likely involves potential changes in the balance of lateral interactions among neurons, emphasizing the role of inhibitory-excitatory dynamics in shaping receptive field properties. Our work provides insights into the functional organization of the visual cortex and highlights how receptive field properties can adapt under varying neural activity states, contributing to a deeper understanding of the cortical computations that underlie visual perception.

S.-H. L was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(grant #: RS-2024-00349515) and by the Korea Basic Science Institute(National research Facilities and Equipment Center) grant funded by the Ministry of Education (grant # RS-2024-00435727).

56.310 HOW DOES THE EARLY VISUAL CORTEX COPE WITH MESOPIC CONDITIONS? EVIDENCE FOR COMPENSATORY MECHANISMS IN CORTICAL SPATIAL SUMMATION AND SURROUND SUPPRESSION

Michaela Klimova[1], MiYoung Kwon[1]; [1]Northeastern University

Despite the prevalence of mesopic (dim light) environments, their effects on human visual processing remain poorly understood, as most vision research and assessments are conducted under idealized photopic (bright) conditions. Animal studies suggest that mesopic conditions enlarge retinal ganglion cell receptive fields and reduce surround antagonism, thereby enhancing light sensitivity but compromising spatial resolution. However, how mesopic conditions affect human cortical spatial integration remains unclear. To address this, we investigated cortical population receptive fields (pRF) and surround suppression, alongside behavioral measures of surround suppression, under mesopic and photopic conditions. We began by acquiring fMRI BOLD responses from visual areas V1-V3 in 11 normally sighted observers. In Exp. 1, we estimated pRF properties with standard stimuli and procedures, while Exp. 2 measured cortical surround suppression using high-contrast center and surround grating stimuli. We also obtained full contrast sensitivity functions (CSFs) with and without surround suppression from a subset of observers. CSFs were fit using a difference of Gaussians model, enabling a psychophysical comparison to the fMRI-derived pRF size and shape estimates. To further link cortical measures with perceptual spatial resolution, we assessed mesopic and photopic visual acuity in all observers. Our results reveal that, mesopic conditions, contrary to findings in animal retinal studies, were associated with smaller pRF sizes across early visual cortical areas ($p < 0.01$). Surround suppression remained robust, with no significant differences in suppression strength between mesopic and photopic conditions at the group level. Interestingly, observers with greater reductions in mesopic visual acuity exhibited larger decreases in V1 surround suppression, highlighting the critical role of surround suppression in contrast coding and spatial resolution. Diverging from retinal electrophysiology studies, our findings suggest that the human early visual cortex may employ compensatory mechanisms to maintain spatial resolution in response to degraded retinal input under mesopic conditions.

56.311 NEURAL AND BEHAVIOURAL EFFECTS OF VISUOMOTOR STIMULATION DURING DUAL RESPONSE TASK

Raffaele Costanzo[1] (raffaelecostanzo.phd@gmail.com), Bianca Maria Di Bello[1], Luca Boccacci[1], Sveva Montesano[3], Merve Aydin[1], Andrea Casella[1], Stefania Lucia[1], Francesca Strappini[1], Camilla Panacci[2], Francesco Di russo[1], Sabrina Pitzalis[1]; [1]"Foro Italico" University of Rome, [2]IRCCS Santa Lucia Foundation, [3]University of Bologna, Italy

Some visual brain areas are prominently involved in the computation of the continuous visual changes induced by locomotion (flow-fields). Moreover, event-related potential (ERP) can be used to study brain functions such as motor readiness, associated with Bereinshaftpotential (BP), and cognitive control, associated with prefrontal negativity (pN). The study aims to verify whether cognitive and motor preparation can be influenced by the presence or absence of visual stimulation (flow-field) or concomitant locomotion. A

visuomotor discriminative-response task (DRT) during electroencephalographic recordings was administered to 40 participants in two counterbalanced conditions: a Still and a Walking. The participants were divided into two equal groups. One group performed the conditions with a flow-field (FF) on the DRT stimuli background. The other group performed the conditions with a black background. A 2x2 mixed design was employed for statistical analyses with FF and Walking as factors. At the behavioral level, a significant positive main effect of Walking on both RT and accuracy was found. The main effect of FF was observed, improving the RT but reducing the accuracy. At the brain level locomotion significantly increased the BP and the pN, but the FF condition significantly increased the BP and reduced the pN. No interaction between conditions was observed. The study corroborates that locomotion may improve cognitive performance. The flow-field presence, likely interfering with the visual tasks, produces a sort of response-speed trade-off favoring speed at the cost of accuracy. At the brain level, a similar trade-off was detected in the condition with the flow-field facilitating motor readiness and suppressing cognitive/inhibitory control. The study showed that cognitive processing is modulated by locomotion. During not realistic walking, probably more neurocognitive resources are invested in the task for better performance. However, in realistic walking with the flow-field presentation, cognitive resources were less available still favoring speed but reducing accuracy.

This work was supported by the Italian Ministry of University and Research grant PRIN-2022_2022BK2NPS and grant PNRR MNESYS-J33C22002970002, both to Sabrina Pitzalis

56.312 SITUATING REDUNDANCY MASKING IN THE VISUAL PATHWAYS

Ângela Gomes Tomaz[1], YuYi Lu[1], Dennis M. Levi[1], Bilge Sayim[2]; [1]Herbert Wertheim School of Optometry & Vision Science, University of California Berkeley, [2]Cognitive and Affective Sciences Lab (SCALab), UMR CNRS 9193, University of Lille, Lille, France

In redundancy masking (RM), observers perceive fewer items than presented in repeating patterns (e.g., perceiving two lines when three are presented). Previous research has found RM for both simple (e.g., lines) and complex stimuli (e.g., faces) indicating that it might occur at multiple stages of visual processing. Here, we investigated whether RM occurs from retinal level neuronal interactions or higher up in the visual pathways. Twenty observers were presented with 3, 4 or 5 lines (0.04° width, 1° height), both monocularly and binocularly, in a mirror stereoscope. In the binocular presentation, lines were presented ipsilaterally (i.e., all to the same eye) and contralaterally (i.e., some lines to the right eye and others to the left eye). The center of the line array was set randomly at +10° or -10° of eccentricity along the horizontal meridian. Lines were radially aligned with respect to central fixation and presented for 150ms. Center-to-center line spacing was 0.78°, and confirmed with a Landolt C orientation discrimination task to be above the observers' minimum angle of resolution at the experimental eccentricity. Observers were instructed to report the number of lines perceived. Results showed that observers reported fewer lines than presented for all numbers of lines revealing RM. The (absolute) magnitude of RM increased with the number of lines presented. Importantly, RM was found both for ipsilateral, as well as contralateral presentations suggesting that its locus is at or beyond the site of binocular integration. However, bigger deviations than expected for contralateral presentations suggest that other factors (e.g., contralateral suppression, binocular rivalry) may be playing a role in the findings. Future psychophysical and neurophysiological research on RM will reveal the precise cortical locus – or loci – of RM.

Supported by a grant from the National Eye Institute awarded to Dennis Levi (R21EY030609) and an NIH institutional training grant (T35EY007139).

56.313 FOVEAL VISUAL ACUITY AND THE JITTER AFTEREFFECT

Lauren Sigda[1], William Tuten[1]; [1]University of California, Berkeley

Illusory motion can be perceived in a stationary stimulus following adaptation to circumjacent dynamic noise (Murakami and Cavanagh, 1998). This phenomenon, known as the jitter aftereffect (JAE), indicates image-based mechanisms play a role in generating a stable percept from the moving retinal image. Previous computational work proposed such perceptual stabilization mechanisms also contribute to resolving fine spatial details (Burak et al, 2010). We examined whether visual acuity is influenced by viewing conditions known to induce the JAE. Using a 60-Hz adaptive optics ophthalmoscope, we measure foveal visual acuity in 8 subjects. Diffraction-limited, tumbling-E optotypes (λ=680 nm) were superimposed on a circular gray background (diameter: 1.25°) provided by an external display. The central uniform disk was surrounded by an inner annulus (diameter: 4°) containing a static array of 2D Gaussian elements (σ: 8') and an outer annulus (diameter: 17°) composed of a binary noise texture with square elements subtending 8'. The boundary between each background zone was spatially blurred. Prior to testing, subjects adapted to dynamic (60 Hz) or static (0 Hz; control) noise in the outer annulus for 30 seconds. Stimuli were delivered for either 100 or 750 ms in a 1-second trial interval during which the background was static. A 6-second top-up adaptation period was inserted between each trial. As expected, the minimum angle of resolution improved with increasing exposure duration (mean±STD: 0.90±0.15 arcmin at 100 ms; 0.66±0.11 arcmin at 750 ms). At 750 ms, no systematic difference between test and control acuity was observed (signed-rank statistic: 16; p=0.844). At 100 ms, visual acuity exhibited a modest improvement in 7 of 8 subjects following exposure to dynamic random noise, though this effect was not significant (signed-rank statistic: 4; p=0.055). Our findings suggest, under the conditions studied, that the mechanisms perturbed during the JAE are not required for high-resolution vision.

This work was supported by the Air Force Office of Scientific Research (FA9550-20-1-0195, FA9550-21-1-0230), the National Institutes of Health (R01EY02359, T32EY007043), the Hellman Fellows Fund, and the Alcon Research Institute.

56.314 SSVEP-BASED ESTIMATION OF CONTRAST SENSITIVITY, VISUAL ACUITY AND ORIENTATION SENSITIVITY FURTHER SUPPORTS THE ABSENCE OF A COMMON FACTOR IN VISION

Martina Morea[1] (martina.morea@epfl.ch), Simona Garobbio[1], Marina Kunchulia[2], Michael Herzog[1]; [1]Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, [2]Free University of Tbilisi and Ivane Beritashvili Center of Experimental Biomedicine, Tbilisi, Georgia

Steady-state visually evoked potentials (ssVEP), elicited by stimuli flickering at a specific frequency, are widely used in vision research and clinical settings as an alternative to behavioral tests. Despite their widespread use, the relationship between behavioral and ssVEP thresholds remains unclear. To address this, thirty-five participants viewed a grating stimulus flickering at a specific frequency with its contrast, spatial frequency, or orientation gradually adapted, to extract the EEG thresholds (no behavioral responses were given). The same participants also completed behavioral tests to assess contrast sensitivity, visual acuity, and orientation sensitivity. We first correlated the two measures systematically in all three conditions. Second, visual tests often correlated only weakly with each other in behavioral studies. We sought to determine whether reducing the influence of confounding factors, such as attention, motor execution, and response strategy, could uncover potential correlations between different visual tests. Despite good test-retest reliability for both methods, no significant correlations were observed between EEG and behavioral measures, suggesting that these two methods may assess distinct aspects of visual functioning. Additionally, correlations between different tests were small and non-significant, providing further evidence that there may not be a single factor underlying visual abilities. These findings suggest that low correlations between visual tests may not lie in the poor specificity of the tests but rather in the inherent independence of the mechanisms they measure.

56.315 INDIVIDUAL DIFFERENCES IN VISION: VARIABILITY RATHER THAN NOISE

Michael Herzog[1] (michael.herzog@epfl.ch), Melissa Faggella[1], Simona Garrobio[1]; [1]EPFL, Switzerland

Classically, vision scientists treat individual differences as noise and eliminate them by aggregating individuals (grand average). For example, individual-level EEG waveforms often appear highly heterogeneous. We recorded EEG during a backward masking task and found between 0 and 3 peaks in the waveforms, depending on the participant. When averaging the waveforms across participants, we obtained one peak (the classic N1), which we interpreted as the “true” waveform, while we assumed that the additional peaks “disappeared” because they were smoothed out as noise. However, when the same participants were tested again 5 and 10 years later their individual waveforms had not changed - a participant with 3 peaks still exhibited 3 peaks a decade later. Hence, individual differences in waveforms are stable traits, not noise. Importantly, many participants did not show an N1 at all. Similar results were observed in behavioural tests. Participants performed a battery of illusions. We found almost no correlations between the magnitudes of different illusions. For example, an observer may experience a strong Ebbinghaus and a weak Müller-Lyer illusion. Hence, there is no common factor for illusions. Nevertheless, the magnitude of each illusion was stable within individuals over a year. We will discuss the implications of heterogeneity for vision research in general, propose methods to measure it properly, and explore how inter-individual differences may pave new avenues in vision research.

SNF

56.316 ORIENTATION COLUMNS IN V1 CANNOT BE DETECTED WITH FMRI AT 0.6 MM RESOLUTION

Cheryl Olman[1]; [1]University of Minnesota

We collected GE fMRI data with 0.6 mm resolution with the hope of being able to characterize the orientation preference of individual voxels in primary visual cortex (V1) as a function of cortical depth (columns) and location across the cortical surface (pinwheels). Stimuli were large 2 cycle per degree gratings, briefly presented in a 2-interval forced-choice paradigm as observers performed a foveal orientation discrimination task on the portion of the grating that was in a 0.5 degree diameter circle at the center of the grating. The experiment used a block design, presenting 4 12-sec blocks of gratings at each of 8 different (average) orientations in each of 4 task scans during a scanning session. High-quality datasets were acquired from 13 of the 16 individuals scanned. Regression analysis estimated the amplitude of response for each voxel in V1 for each of the 8 different orientations. Visual responses in parafoveal regions of interest (ROIs) defined by an independent localizer were robust: 76% +/- 10% s.d. of the voxels in ROIs defined on the surface and then propagated through the cortical depth were modulated by visual stimuli at the $p < 0.001$ (uncorrected) level. However, orientation preference of individual voxels was not reliable (e.g., not the same in the first and second half of the scanning session in a split-halves analysis). Average voxel orientation selectivity, computed as 1 minus the circular variance of responses to the 8 orientations, was not significantly different from the null hypothesis. Interestingly, computed orientation selectivity did vary through depth, but only because the inherent contrast-to-noise ratio of fMRI data varies through depth. Also, in spite of the lack of reliable orientation preference or selectivity in individual voxels, a reliable radial bias was evident in the data.

NIH R01 NS123482, NIH R01 MH111447

56.317 STRONGER VISUAL SURROUND SUPPRESSION UNDER PSILOCYBIN: A PSYCHOPHYSICAL AND EEG PILOT STUDY

Michael-Paul Schallmo[1] (schal110@umn.edu), Sophia Jungers[1], Ranji Varghese[1], Kathryn Cullen[1], Michael Evans[1], Jessica Nielson[1], Link Swanson[1]; [1]University of Minnesota

Perception of visual contrast depends on the surrounding spatial context. Typically, the salience of a central target is reduced by a high contrast surrounding stimulus, an effect known as surround suppression. Although this phenomenon is well-studied, the role of specific neurotransmitter systems during surround suppression in human vision remains unclear. Psilocybin is a serotonin (5-HT2A) receptor agonist known to affect visual perception (e.g., psychedelic visual phenomena). We asked whether surround suppression may be altered by psilocybin. In a double-blind crossover pilot study, healthy adults completed psychophysical ($n = 6$) and electroencephalography (EEG; $n = 5$) measures of surround suppression after taking either 25 mg of psilocybin or placebo (100 mg niacin), in separate sessions. We found that psilocybin increased the strength of surround suppression, as measured in both our psychophysical contrast matching task, and in the strength of the visual N1 component from EEG. Accuracy on catch trials was not significantly impaired under psilocybin. Our results, although preliminary and limited by a small sample size, suggest that serotonergic neuromodulation plays a role in regulating the strength of surround suppression. Our findings may also be relevant for understanding differences in visual perception in psychiatric conditions, such as the weaker surround suppression reported in individuals undergoing major depressive episodes.

University of Minnesota Foundation, the Heffter Research Institute, NIH Center for Advancing Translational Sciences (UL1 TR002494)

56.318 CONTRAST IMPACTS POPULATION HETEROGENEITY OF ORIENTATION TUNING IN V1

Joseph Emerson[1], Ryan Holland[2,3], Gordon Smith[1,2,3], Audrey Sederberg[4,5], Cheryl Olman[1,6]; [1]Graduate Program in Neuroscience, University of Minnesota, Twin Cities, [2]Department of Neuroscience, University of Minnesota, Twin Cities, [3]Optical Imaging and Brain Sciences Medical Discovery Team, University of Minnesota, Twin Cities, [4]School of Physics, Georgia Institute of Technology, Atlanta, GA, [5]School of Psychology, Georgia Institute of Technology, Atlanta, GA, [6]Department of Psychology, University of Minnesota, Twin Cities

The tuning properties of primary visual cortex (V1) display a large degree of heterogeneity across neurons which help support efficient coding of stimulus features. However, the mechanisms supporting heterogeneity in V1 are not well understood. We ask how orientation tuning heterogeneity is impacted by recurrent processing, which is known to shape orientation preference of cells in V1. To this end, we examine 2-photon calcium-imaging data from two anesthetized ferrets that viewed drifting grating stimuli at varying contrasts. We find that in both animals there is substantial variability in orientation selectivity across neurons. Interestingly, the variance of orientation selectivity across the population decreased with increasing contrast, suggesting a relationship between the strength of feedforward inputs and the population heterogeneity of orientation tuning. This trend is not due to contrast-dependent changes in measurement signal-to-noise ratios as a control analysis shows no significant trend in population heterogeneity across contrast when orientation labels are scrambled. We hypothesize that the change in tuning heterogeneity is brought on by a network transition characterized by a shift from feedforward-dominated inputs to a recurrence-dominated regime. To better understand the mechanistic underpinnings of this transition, we are using a stabilized supralinear network (SSN) model to investigate the contributions of recurrent connections in V1 to the population heterogeneity in the tuning properties of neurons. As part of ongoing work, we are testing how spatial and functional tuning properties of recurrent connectivity within the network affect orientation tuning variability across the population in a manner consistent with trends in ferret calcium-imaging data.

NIH R01 NS123482

56.319 ORIENTATION DECODING FROM NEURONAL POPULATIONS IN MACAQUE V1: AN EXTERNAL NOISE INVESTIGATION

Cong Yu[1], Xin Wang[1], Cai-Xia Chen[1], Shi-Ming Tang[1], Zhong-Lin Lu[2]; [1]Peking University, [2]New York University

The external noise paradigm (Barlow, 1957) has been widely applied in psychophysics (Pelli, 1981) and neurophysiology (Anderson-et-al, 2000) to investigate properties of human observers and neuronal tuning, as well as changes associated with attention and perceptual learning (Lu&Dosher, 2013). In this study, we employed this paradigm with two-photon calcium imaging and a transformer model to study orientation decoding from neuronal populations in macaque V1. Specifically, we simultaneously recorded the responses of more than 1,000 neurons within each FOV to a Gabor stimulus (12 orientations, contrasts = 0.03–0.50) embedded in white external noise (RMS: 0.0–0.29) in two awake, fixating macaques. We found that external noise suppressed the population orientation tuning functions, reducing the amplitude and widening the bandwidth with increasing external noise contrast. At low Gabor contrasts and high external noise, the bandwidth became unmeasurable. A gain control model provided a good fit for the observed external noise suppression effects. To decode the population responses from the recorded neurons, we applied a transformer model to reconstruct the trial-by-trial Gabor images from the neuronal response vectors. Orientation information was then extracted from the reconstructed images to evaluate orientation decoding accuracy. In a wide range of stimulus conditions, where the Gabor contrasts were not too low and the external noise levels were not too high, orientation decoding accuracy was unaffected by the contrast of external noise, despite the suppression of the population orientation tuning function. This stability was achieved by recruiting more hub neurons and establishing more effective neuronal connections. However, decoding accuracy was severely compromised under the highest external noise conditions (RMS: 0.16 & 0.29). The constant decoding accuracy across many experimental conditions, achieved through the recruitment of hub neurons and more effective connections, offers important new insights into how the visual system extracts information from noisy neuronal responses

STI2030-Major Projects grant (2022ZD0204600)

Attention: Visual search

Tuesday, May 20, 2:45 – 6:45 pm, Banyan Breezeway

56.320 3D DEPTH-INVARIANT LEARNED SPATIAL SUPPRESSION

Seah Chang[1] (seahchang@ajou.ac.kr), Julie D. Golomb[2]; [1]Ajou University, [2]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. Previous studies have primarily focused on spatial suppression in two-dimensions (2D), however, we live in a three-dimensional (3D) world where visual stimuli appear across multiple depth planes. Is suppression learned in a representation of space that includes depth information? In the current study, participants searched six items for a shape oddball target (e.g., a diamond among circles) while ignoring a salient color singleton distractor that appeared on some trials (singleton-present trials). The six search items appeared either in front of (near disparity) or behind (far disparity) the fixation plane (a fixation square presented at the

center of the screen at the middle position in depth). In Experiment 1, the salient color distractor appeared more frequently in one specific high-probability location, defined in 3D. I.e., one of the six locations in either the front or back depth plane contained the salient distractor on 54.17% of trials (depth-matched high-probability location). The salient distractor appeared less frequently (4.17%) at the other locations, including the same 2D location in the other depth plane (depth-mismatched high-probability location), other locations in the same depth plane (depth-matched low-probability locations), and other locations in the other depth plane (depth-mismatched low-probability locations). In Experiment 2, each depth plane had its own, different high-probability location, resulting in two high-probability locations in 2D. The results from both experiments consistently showed evidence for depth-invariant suppression: the suppression effect (reduced attentional capture by a salient distractor in high-probability locations compared to low-probability locations) was observed for both the depth-matched and depth-mismatched locations. These results suggest that learned spatial suppression operates on 2D representations, despite the 3D nature of our visual environment.

NIH R01-EY025648 (JG), NSF 1848939 (JG), Ajou University (SC)

56.321 DO IMPLICIT COLOR-EMOTION ASSOCIATIONS MODULATE DISTRACTOR SUPPRESSION?

William Secrest[1], Brad Stilwell[1], Brian Anderson[1]; [1]Texas A&M University

Distracting stimuli can be suppressed based on their simple features, such as color or shape. For example, individuals can learn to suppress a salient distractor, such as a uniquely colored item, when searching for a specific shape. However, in the real world we often encounter distractors that are associated with more complex properties such as emotional valence. Individuals can form implicit associations between arbitrary stimuli and emotional valence, but it remains unclear whether these valence associations can modulate distractor suppression. To test this, participants associated colors with emotionally valent words and we tested whether these color-valence associations modulated distractor suppression when those same colors served as distractors. To establish color-valence associations, we adapted the Implicit Association Task in which participants classified centrally presented words as either “good” or “bad.” Critically, the words were presented in one of three colors: one associated with positive words, one with negative words, and one with both positive and negative words. To assess whether these color-valence associations modulated distractor suppression, participants performed a visual search for a target shape (i.e., a unique shape in Experiment 1 or a specific shape, such as a circle, in Experiment 2). A salient, uniquely colored singleton distractor was either present or absent and participants were instructed to ignore it. The color singleton distractor could be one of the three valence-associated colors. In the association phase, participants were faster to respond to words in the positive and negative color, in addition to being faster to make a positive judgment. However, suppression of these valence-associated colors was not modulated. These results suggest that if there is any effect on attentional priority via implicit associations with emotion, it is either small, short-lived, or difficult to observe.

56.322 FROM CAPTURE TO CONTROL: INITIAL CAPTURE LEADS TO LEARNED SUPPRESSION

Yue Zhang[1] (yuezhang@missouri.edu), Nicholas Gaspelin[1]; [1]University of Missouri

Much research has shown that distraction can be reduced by suppressing salient stimuli to prevent attentional capture. This ability to suppress seems to result, by and large, from implicit learning and this learning develops within a few exposures to a salient stimulus. However, the exact mechanisms underlying learned suppression remain unclear. The current study will explore one potential explanation: Salient stimuli initially capture attention which jumpstarts a suppressive process to prevent future distraction. To study this, we introduce a new color-blocking technique. Participants searched for a target and attempted to ignore a salient distractor. Importantly, we changed the display colors every 5 trials. This allowed us to examine learned suppression that develops quickly (i.e., within a few trials). The results show that salient distractors initially captured attention and were rapidly suppressed within a few trials of exposure. Other experiments explored the causal relationship between initial capture and later suppression. These findings indicate that feature-based suppression is rapidly learned, perhaps taking only a single trial. They also demonstrate that initial capture may be necessary for suppression to develop.

56.323 IGNORING SALIENT DISTRACTORS INSIDE OF THE ATTENTIONAL WINDOW

Xiaojin Ma[1], Steven J. Luck[2], Nicholas Gaspelin[1]; [1]University of Missouri, [2]University of California, Davis

Salient stimuli are often assumed to have an inherent power to attract attention. However, formal research has shown that attentional capture by salient distractors can often be attenuated. This ability to ignore salient distractors is typically thought to reflect top-down control of attention. However, an alternative theory has been proposed. According to the attentional window account, attention can be narrowly focused to prevent salient distractors from capturing attention. Importantly, it has been suggested that most prior evidence of top-down control could be a result of narrow attentional focusing. The present study examined attentional capture by salient distractors under different breadths of attentional focus, using ERP indices of attentional selection. Participants completed a shape discrimination task. Importantly, the shapes were arranged so that a color singleton appeared either inside or outside of attentional focus. Across several experiments, we found that the color singleton did not elicit evidence of attentional capture, as measured by the N2pc component and behavioral indices. In addition, control conditions that required the color singleton to be attended did produce an N2pc, showing the paradigm was sensitive to detect attentional selection of the salient stimulus. Altogether, these findings suggest that attentional capture by salient stimuli can be prevented even when attention is broadly focused across an entire display. This is inconsistent with the attentional window account and instead supports models of attention that allow for top-control control.

This project was supported by the National Science Foundation Grant BCS-2345898 to N.G. and NIH Grant R01MH065034 to S.J.L.

56.324 INVESTIGATING THE TEMPORAL DYNAMICS OF LTM-TO-VWM REINSTATEMENT DURING VISUAL SEARCH

Jessica Kespe[1], Marielle Noack[1], Naseem Al-Aidroos[1]; [1]University of Guelph

When observers are shown a visual target to later search for, they form an attentional template—¬a mental representation of the target's features. This template acts as a guide, allowing only stimuli that match the target's features to capture attention. Recent research has shown that templates stored in visual working memory (VWM) do not only reflect the target object that was shown to the observer, but also features from long-term memory (LTM) that were previously associated with that object. For example, if a shape was previously associated with a specific colour, that association can be retrieved and represented in VWM, even in the absence of direct visual input: a process termed LTM reinstatement. Interestingly, reinstated features in VWM can prefentially capture attention even when they are task irrelevant. How long do you need to see an object for task-irrelevant features from LTM to be reinstated, and how long do you need to actively maintain that feature in VWM to see the attentional effects? To investigate both questions, participants memorized a set of objects with specific colours and then completed a search task where they were instructed to search for one of the memorized objects' shapes in any colour. We manipulated how long participants saw the search target object (Exp 1) and how long they had to hold the target in VWM before the search task started (Exp 2). We found that participants needed to see the object for a minimum of about 50ms, and then maintain it in VWM for another 50ms, for LTM reinstatement to later occur. These results highlight the rapid interaction between LTM and VWM, suggesting that even brief perceptual exposure and VWM maintenance can trigger the reinstatement of task-irrelevant features.

NSERC Grant to Naseem Al-Aidroos and Jessica Kespe

56.325 SELECTION HISTORY UPWEIGHTS AND DOWNWEIGHTS ATTENTIONAL PRIORITY WITHOUT MODULATING PERCEIVED SALIENCE

Molly R. McKinney[1], Brad T. Stilwell[1], Brian A. Anderson[1]; [1]Texas A&M University

In visual search, attentional priority can be upweighted to enhance target-related features and downweighed to suppress distractor-related features based on experiential learning, or selection history. However, it is unknown whether these selection history effects modulate information processing at the level of perception. To test this question, individuals performed a visual search task with predictable target and distractor features. Specifically, in Experiment 1, in a between-subjects design, we presented a color singleton among three nonsingleton items that served as a distractor for one group, the target for another group, or was equally likely to be the target or distractor in a third group; the color of the singleton was fixed across trials in each group. In Experiment 2 we used two-color, eight-item displays. One color was always used for distractors when presented and a second color was always used for the target when presented; a pair of other colors unpredictably included the target in their subset. To test whether selection history would change perception of the predictive features, we used a psychophysical task whereby participants attempted to detect either the presence of a briefly presented color singleton (Experiment 1) or identity of the color of a briefly presented, uniformly colored stimulus array (Experiment 2). Across both experiments, participants showed more efficient search for the learned colors, enhancing the consistent target color and suppressing the consistent distractor color. However, we did not observe corresponding changes in perception of these colors, with no change in accuracy to detect or identify the briefly presented stimuli. These results suggest that attentional priority for previously learned features can boost visual search performance without influencing the perceived salience of briefly presented stimuli.

56.326 WHAT IS FEATURE-GENERAL SUPPRESSION AND DO PEOPLE ACTUALLY USE IT?

Isaac Savelson[1], Andrew B. Leber[1]; [1]The Ohio State University

Everyday life demands that we minimize distraction by irrelevant visual information. Fortunately, humans have a robust ability to ignore distracting stimuli through repeated experience with the distractor's properties. Originally, distractor ignoring was thought to rely on suppression of specific features carried by the distractor. However, recent work has revealed that suppression sometimes generalizes to never before seen features. At present, little work has been done to investigate how feature-general suppression compares to its feature-specific counterpart. Here, we set out to investigate how feature-general suppression is achieved. Is it a complete suppression of all color-space or merely a special case of feature-specific suppression? To answer this, we first needed to establish a more robust method of testing feature-specific and general suppression strategies. We accomplished this using a probe procedure where we presented select distractor colors during a visual search task while, in a probe task, we compared attentional prioritization of these distractor colors to that of test colors that were never presented as distractors. Prioritization of test colors was greater than distractor colors when we presented a single fixed distractor color in search, but not when we randomly varied the distractor among multiple colors. With this new design, we tested our original question about feature-general suppression by presenting a range of test colors varying in color-space distance from the distractor color(s). Overall, distractor colors were prioritized less than test colors but a greater difference was found when a single fixed distractor was presented compared to multiple colors. This pattern suggests that when variable distractors are presented in search, participants do not automatically suppress all salient color singletons; instead, they accumulate suppression in a feature-specific manner. However, this suppression can partly generalize across color space when variable distractor features appear during search.

Funding: NSF BCS-2021038 to ABL

56.327 PROACTIVE CONTROL PREVENTS SALIENCE-DRIVEN ATTENTION CAPTURE IN A CUED GO/NO-GO SEARCH TASK

John McDonald[1], Daniel Tay[1], Rebecca Carson[1]; [1]Simon Fraser University

According to salience-driven selection theory, observers automatically orient attention to the most salient stimulus in the visual field unless that stimulus falls outside of the currently monitored region (i.e., the attentional window). Recently, we introduced a go/no-go search paradigm to test this theory in a novel way (Tay et al., 2022; doi:

10.1037/xhp0000972) Participants viewed arrays of 16 blue items or 16 yellow items and indicated whether arrays of one colour (Go trials) contained an orientation singleton. No response was required for displays containing other-colour items (No-Go trials). The singleton elicited an event-related potential (ERP) component associated with attentional selection—the N2pc—on Go trials but not on No-Go trials, which is inconsistent with salience-driven selection theory. Here, we designed a cued Go/No-Go search task to determine whether distractor-suppression processes can be engaged proactively on a trial-by-trial basis. Each search display contained 16 blue bars that were all oriented horizontally or vertically on randomly intermixed trials. On 50% of trials, one bar was rotated 90° to produce an orientation singleton. Approximately one second before the search display, an array-wide square or circle was presented for 100 ms to indicate whether a response (singleton present or absent) was required for the upcoming array. Thus, participants had ample time to engage inhibitory control processes prior to display onset on No-Go trials. Unsurprisingly, the singleton elicited an N2pc over the posterior scalp on Go trials. On No-Go trials, the same singleton elicited a long-lasting distractor positivity (PD) that began 150–200 ms after array onset. Anterior-scalp ERPs indicated that participants did not passively ignore the search array on No-Go trials. This suggests that observers can decide in advance, and on a trial-by-trial basis, to search for a singleton or to prevent such search from occurring.

56.328 SEARCHED BUT NEVER FOUND: ATTENTIONAL CONTROL SETTINGS ARE PRUNED BASED ON INTERACTIONS WITH THE EXTERNAL ENVIRONMENT

Samantha Joubran[1], Rebecca McCabe[2], Kelyn Young[1], Naseem Al-Aidroos[1]; [1]University of Guelph, [2]Brock University

When visually searching for a large set of everyday objects, only those objects capture our attention, indicating we have adopted attentional control settings (ACS) based on long-term memory representations of those objects. Beyond our internal attentional goals, are long-term memory ACSs also influenced by the external environment? Across three experiments, we had participants memorize and then search for sets of 12-24 objects during a spatial blink task. In this task, a stream of 20 non-studied objects were presented in the center of the screen and participants reported the target (one of the memorized objects) that appeared within the stream. Either 2, 5, or 8 positions before the target, two distractors appeared above and below the central stream. One distractor was always a non-studied object while the other was sometimes a critical distractor (a memorized object). When the critical distractor was a memorized object, attention was captured away from the stream to the distractor's spatial location which caused participants to miss the target; this did not occur if the distractors were both non-studied objects. The difference in accuracy is called a spatial blink and it showed attentional capture by the distractor. Across our experiments, we manipulated which memorized objects were presented as search targets allowing an investigation into how our interactions with the external environment can impact our ACS. While all memorized objects initially captured attention, over time, any objects that never appeared as search targets were pruned from the ACS and stopped capturing attention. This suggests that objects that are searched for, but never encountered as a target, will eventually be pruned from the ACSs. These results seem to demonstrate that while we can have internal goals that contribute to an ACS, our interactions with the environment can impact the maintenance and updating of our ACS.

Funded through NSERC

56.329 TRACKING THE CAPACITY BOTTLENECK IN MULTIPLE-COLOUR SEARCH

Ziyi Wang[1], Anna Grubert[1]; [1]Durham University

Visual search for known objects is guided by target templates stored in visual working memory. Previous (two-forced choice) search tasks found that reaction times (RT) were substantially delayed when multiple attentional templates had to be activated simultaneously, e.g., during search for two as compared to one target colour(s). Interestingly, such load costs were much smaller when measured directly at the neuronal level of attentional target selection (as indexed by the N2pc component of the event-related potential; ERP). This suggests that most of the behavioural load effects originate at processing stages that follow initial target selection. To track the capacity bottleneck during multiple-colour search, we analysed distinct ERPs associated with attentional target selection (N2pc), target identification (SPCN), and response selection and execution (stimulus-locked and response-locked lateralised readiness potentials; sLRP/rLRP) under low- and high-load conditions of three search tasks with increasing complexity: simple RT task (key press if the target colour is present), Go/Nogo task (key press only if the target-colour object has a specific identity), and two-forced choice task (2FC; different key presses for different target identities). Results revealed slower RTs in high- than low-load trials of all tasks, but these load costs were significantly increased in the 2FC than the simple RT and Go/Nogo tasks. ERPs mirrored these behavioural findings perfectly. Load costs on target selection and identification stages (N2pc/SPCN) were small and did not differ between the three tasks. However, load costs on response selection stages were substantial in the 2FC task, but were entirely absent in the simple RT and Go/Nogo tasks. rLRP components were entirely unaffected by load. These results demonstrate that multiple attentional templates can guide target selection and identification in parallel with minimal costs and that the behavioural costs observed in previous 2FC search tasks were likely caused by response-based capacity limitations.

This work was funded by research grants of the Leverhulme Trust (RPG-2020-319) awarded to AG.

Attention: Capture

Tuesday, May 20, 2:45 – 6:45 pm, Banyan Breezeway

56.330 ALL EYES ON THE ANIMALS: ANIMACY GUIDES VISUAL ATTENTION

Ece Yucer[1] (ece.yucer@mail.utoronto.ca), Kristina Knox[1], Tyler Puhlick[1], Jay Pratt[1]; [1]University of Toronto

The ability to detect moving entities within our visual field is essential for human survival, as these entities often signal potential threats or opportunities. Animate beings, such as animals, exhibit dynamic and unpredictable movement patterns, making them highly salient to our visual attention system. This study seeks to explore the extent to which animacy influences saccadic eye movements using a forced-choice saccade task. In Experiment 1, an animal image and a real-world size-matched inanimate object image both from the THINGSPlus Database, grey-scaled and backgrounds removed, were presented simultaneously on a monitor. The participants completed a series of trials organized into four blocks, where they were instructed to focus their gaze, as quickly as possible, on the animate image during two blocks and on the inanimate object image during the other two blocks. Experiment 2 presented a single image at a time, either an animal or an object, with the same block trial structure. This allowed for a more precise comparison of reaction times between the two categories of objects without the potential decision-making implicated in simultaneous stimuli presentation. In both experiments, participants' saccadic eye movements were recorded with an eye tracker. The results show that participants exhibited shorter saccadic reaction times when looking toward animate objects than their inanimate counterparts. Eye movement errors and amplitudes did not differ between the two types of stimuli. Overall, this indicates that our visual system is tuned to prioritize and quickly respond to potential living threats or opportunities in our environment, underscoring the importance of animacy in our attentional processes.

56.331 ATTENTIONAL CAPTURE VS. ATTENTIONAL BIAS: NOVEL EVIDENCE FOR THE PRIORITY ACCUMULATION FRAMEWORK AND DISSOCIATION BETWEEN EXOGENOUS AND ENDOGENOUS CONTROL OF OVERT ATTENTION.

Mor Sasi[1], Daniel Toledano[1], Shlomit Yuval-Greenberg[1], Dominique Lamy[1,2]; [1]Tel-Aviv University, [2]Sagol School of Neuroscience, Tel Aviv University

Most visual search theories assume that attention is automatically allocated to the location with the highest priority at any given moment. The Priority Accumulation Framework (PAF) challenges this assumption by suggesting that priority weights accumulate across time until action-relevant information signals the appropriate moment for deploying attention to the location with the highest accumulated priority. This account suggests a distinction between attentional capture - by which attention is deployed at the wrong moment, and attentional-priority bias - by which attention is deployed at the right time but biased towards the locations of previous irrelevant events. In previous work, we recorded eye movements in a free-viewing spatial-cueing task. In line with PAF's predictions, we found that only a minority of first saccades occurred early, in response to the onset cue (attentional capture), whereas most first saccades occurred later, after processing of the search display had already started, but were nevertheless biased towards the cue location. This bias was stronger the more difficult the search (attentional-priority bias). Here, we replicate these findings when the cue-target interval is longer, thereby precluding the possibility that the impact of the cue on late saccades resulted from delayed capture of overt attention. In addition, we show a dissociation between early and late saccades, which we take to uncover a qualitative difference between exogenously and endogenously triggered saccades. Specifically, we show that unlike for endogenously triggered saccades, the identity of the object (target or distractor) on which an exogenously triggered saccade lands, does not determine how long this object remains fixated or whether or not an additional saccade is initiated. These findings, which were replicated in three experiments, suggest that processing is shallower at the locus of a stimulus-driven saccade than at the locus of endogenously triggered saccades.

Support was provided by the Israel Science Foundation (ISF)

56.332 DOES INHIBITION OF RETURN CARE ABOUT SPATIAL FREQUENCY ?

Daniel Lougen[1], Jay Pratt[1]; [1]University of Toronto

Inhibition of Return (IOR) is a phenomenon where responses to targets appearing at previously attended locations are slower after a certain temporal interval, enhancing visual search efficiency by discouraging attention from revisiting examined areas. This study investigates whether the spatial frequency of cues and targets modulates IOR by selectively engaging either the magnocellular (sensitive to low spatial frequencies; LSFs) or the parvocellular (sensitive to high spatial frequencies; HSFs) visual pathways. We conducted a typical visual cueing task (two horizontally aligned cue/target locations with placeholders) with participants completing one of four cueing conditions: LSF cues, HSF cues, LSF targets, or HSF targets. These cues and targets were circles filled with black and white bars at 3 cpd (LSF) or 12 cpd (HSF). The control cues and targets were a white outline with a solid black circle. The inter-stimulus interval (ISI) between cue and target ranged from 100 ms to 1400 ms in order to examine the time course of IOR in each condition. In the spatial frequency cue conditions, we found that LSF and HSF cues generated typical – and equivalent – IOR effects and time courses. In the spatial frequency target conditions, we observed a surprising result; the lack of IOR with both spatial frequencies across all ISIs . Across all conditions, we additionally found the expected decrease in RTs synonymous with longer ISIs, indicating that participants attended to both cues and targets. Overall, the evidence from this study is indicative of a stronger effect centered around relationship between the physical features of cues and targets rather than what pathway is primed to process the stimuli.

56.333 ELECTROPHYSIOLOGICAL MARKERS OF DISTRACTOR SUPPRESSION ARE INTERACTIVELY SHAPED BY SEARCH MODE AND DISTRACTOR SALIENCE

Dock Duncan[1,2] (dockyd@gmail.com), Dirk van Moorselaar[1,2], Jan Theeuwes[1,2,3]; [1]Vrije Universiteit Amsterdam, [2]Institute for Brain and Behaviour, Amsterdam, [3]William James Center for Research, ISPA-Instituto Universitario, Lisbon, Portugal

Recent advancements in understanding visual search dynamics have illuminated the complex interplay between attentional selection and distractor processing. A key finding is the divergent mechanisms by which the brain handles salient distractors across different search modes, with interference patterns dramatically shifting between parallel and serial visual search conditions. While previous research has established distinct patterns of distractor processing across search modes—with high salient distractors interfering more in parallel search and being more efficiently ignored in serial search—the underlying neural mechanisms remain unclear. Specifically, whether the observed discrepancy should be attributed to proactive or reactive mechanisms remains unresolved. To investigate this, we relied on lateralized event-related potentials, the N2pc and the Pd that index attentional selection and distractor processing respectively. In a series of EEG experiments, using different variations of the additional singleton paradigm we systematically varied search dynamics and distractor salience across three levels (low, medium, and high) to unpack these critical issues. Participants performed search tasks with distractors ranging from low to high salience across both parallel (Experiment 1) and serial (Experiment 2) search conditions. At the behavioral level, we replicated the established pattern: in parallel search, distractor interference scaled linearly with its salience, while in serial search, the scaling pattern was reversed, with high salient distractors leading to faster search times compared to low salient distractors. Critically, our encephalographic analysis demonstrated that both early and late phases of the Pd component were systematically modulated by distractor salience, with distinct patterns across search modes. These neural-behavioral findings converge to challenge the predominant paradigm that distractors are proactively suppressed in serial search tasks, motivating a reevaluation of how we study distractor suppression. Furthermore, these results provide additional insight into the use of variable latency of the widely used distractor positivity component to dissociate between proactive and reactive mechanisms.

ERC grant (833029 - LEARNATTEND) and NWO grant (406.21.GO.034)

56.334 **HOW STATISTICAL IS "STATISTICALLY-LEARNED" DISTRACTOR SUPPRESSION?**

Brad T. Stilwell[1], Darrell A. Worthy[1], Brian A. Anderson[1]; [1]Texas A&M University

Salient items, such as uniquely colored stimuli, can be suppressed. Suppression is greater for a salient distractor that is frequently encountered. Such suppression has been argued to occur via statistical learning such that the priority of a distractor is down-regulated based on the predictiveness of its features. However, there are more generalized mechanisms that could lead to learned distractor suppression, such that each color becomes less distracting after each encounter via either feature-specific or feature-nonspecific decay. To test between these plausible mechanisms, participants searched for a unique shape target in the presence or absence of a salient color singleton distractor. Across all experiments, the frequency of the color of the singleton was manipulated to create one high- and several low-frequency singleton colors. Each experiment varied the manner in which these frequencies were realized: The high-frequency and each low-frequency color were presented either intermixed within blocks of trials or individually across blocks. When the colors varied across blocks, the ordering of blocks varied either unpredictably, in a fixed order, or were front-loaded so that all the high-frequency singleton blocks occurred before each low-frequency singleton block. We observed greater suppression for the high-frequency than the low-frequency colors across experiments. Fitting the RT data with computational models revealed that this learned distractor suppression was best explained by both a feature-specific and a feature-nonspecific exponential decay function; both models fit the data better than a model in which attentional priority reflected the frequency with which the color of the distractor was encountered, although some individual participants were best fit by this model. These results suggest that learned distractor suppression may be better explained as a skill that develops from experience suppressing stimuli that can be both feature-specific and feature-nonspecific, as opposed to a product of the learned predictiveness of distractor features.

56.335 OPTIMIZING MEMORY TASK RATIO AND DISENTANGLING PRECISION AND GUIDANCE

Mihrican Yaren Kaynar[1], Eren Günseli[1]; [1]Sabanci University

Working memory (WM) plays a critical role in guiding attention to stimuli that match its contents. Such memory-driven attentional guidance is often studied using search tasks. Reaction time costs when the memory item is a distractor are taken as evidence for involuntary attentional guidance. However, variations in the ratio of memory to search trials across studies may confound comparisons. To address this, we manipulated the memory/(memory + search) trial ratio (25%, 50%, and 75%) to evaluate its effects on guidance. Participants were shown a color to remember and either reported the color in a memory task or the location of a diamond among disks in a search task. The distractor color either matched or differed from the memory item. Memory-guided attention was observed for all ratios. However, when the number of search trials was equated to match the minimum across conditions, the guidance effect became insignificant in the 25% memory block. The 50% ratio emerged as the most trial-efficient condition, requiring only 68 trials to detect a significant effect. Additionally, in the memory task, the mean error was smaller in 75% memory blocks than in 25% and 50% memory blocks, indicating that more frequent memory tests led to better performance. The difference in memory performance with the lack of variation in guidance between the 50% and 75% memory blocks, indirectly supports the notion that memory strength alone cannot predict attentional guidance. Together, our findings provide practical recommendations for optimizing experimental designs and offer theoretical insights into the relationship between memory strength and attentional guidance by WM.

56.336 SALIENCE DOES NOT ALWAYS EQUAL DISTRACTION

A. Kane York[1], Han Zhang[1], Jacob Sellers[1], Taraz Lee[1], John Jonides[1]; [1]University of Michigan

Visual attention is highly susceptible to interference, particularly from two types of distractors: uniquely colored items (color singletons) and suddenly appearing items (abrupt onsets). How do individuals process and ultimately disregard such distractions? To explore this, we used a forced-response paradigm to examine the temporal dynamics of attentional shifts during a feature-search task. Three types of distractors were investigated: color singletons, abrupt onsets, and a combination of the two (color singleton abrupt onsets). In this paradigm, participants were trained to make a saccade in response to a fixed "go" signal, while the duration of the stimulus display was systematically varied. This method revealed the temporal landscape of visual processing, showing that abrupt onset distractors elicited a stronger and more sustained attentional bias compared to the other two distractors. However, this bias diminished as saccade initiation was delayed. Separately, a novel salience measurement method was used with a different group of participants to evaluate distractor salience. Interestingly, this measure indicated that abrupt onsets were the least salient among the three distractor types. In a follow-up experiment, participants completed both the salience measure and the forced-response task with the same stimuli. This allowed for direct comparison within the same group, and the results replicated the findings from the separate populations: abrupt onsets produced a distinct temporal attentional bias in the forced-response task, while being less salient. These findings, combined with insights from a computational model that detailed differences in distractor suppression across conditions, highlight the dynamic and evolving mechanisms of attentional control in the presence of distractors.

National Science Foundation

56.337 THE ROLE OF PREDICTIVE PROCESSING AND PERCEPTUAL LOAD IN SELECTIVE VISUAL ATTENTION: AN EXAMINATION WITH SEMANTICALLY SALIENT AND LESS SALIENT DISTRACTORS

Burcu A. Urgen[1], Ataol B. Ozsu[2]; [1]Bilkent University

Theories explaining how selective visual attention is mediated by top-down regulation highlight various factors, including the presence of a task and its perceptual load (Lavie, 1995), the salience of visual elements (Eltiti et al., 2005), and expectations about the perceived visual input (Summerfield & Egner, 2009; Rauss et al., 2011). Although these accounts are not theoretically conflicting, empirical studies integrating these factors to explore their interplay remain scarce. In this study, we investigate how expectations about task demands influence the processing of distractors with varying degrees of semantic salience. We adapted the letter search task (Lavie & Cox, 1997) and introduced predictive cues ("EASY" or "HARD") about upcoming task demands (low-load vs. high-load). These cues correctly predicted the task demands in 75% of trials. Two experiments were conducted: in the first (n = 20), we used Gabor patches as distractors with low semantic salience, while in the second (n = 20), we used faces as distractors with high semantic salience. Our analyses revealed that when the cues about task demands were incongruent with the actual sensory input, participants exhibited slower response times. Furthermore, the effect of cue-task congruency was amplified in the presence of semantically salient distractors, as participants showed significantly greater differences between congruent and incongruent trials when faces were used as distractors. Lastly, the interaction between cue congruency and the distractor presence was only present in the second experiment as a larger difference in reaction time between congruent and incongruent trials was observed when faces were presented compared to no distractor condition. In conclusion, we suggest that expectations about task demands influence attentional processing for both semantically salient and less salient stimuli. However, when visual elements compete more intensely for attention, utilizing prior information about task demands becomes increasingly important for effective attention regulation.

56.338 USING CONTINGENT CAPTURE TO IDENTIFY THE MECHANISM OF LEARNED RELEVANCE ON ATTENTION

Nancy Carlisle[1], Greta Manini[2]; [1]Lehigh University, [2]Universita di Verona

Attention researchers often split the world into targets and distractors, but in the real world we may learn that certain information probabilistically cues a target through our experience. It is unclear how selection history effects interact with attentional top-down control settings. We manipulated the relevance of salient items by changing how frequently a salient item was a target vs. a distractor across groups (0%-100%; 50%-50%, or 100%-0%). Participants completed a singleton detection task where attention control should already be tuned to attend to salient items and a feature-search task where attention should be directed to a target feature and salient items are ignored. We found more attention is directed to the salient item as relevance increases, with larger costs for salient distractors and larger benefits for salient target compared to search trials where no salient item appears. Importantly, in this study we included contingent capture stimuli that could either match the color of the salient item or another color to determine if the salient item color is becoming a part of a top-down task set, or whether participants are shifting how much attention they are directing to salience. In singleton detection mode, increasing relevance led to increasing contingent capture for both salient item color match and non-match stimuli, suggesting a graded increase in attention towards salience. In feature detection mode, we found a contingent capture effect only for the highest level of relevance, but still found similar capture for both salient item color match and non-match stimuli, once again suggesting this was a shift in attentional control towards salience instead of incorporating the salient item color into a top-down control set. These findings highlight that attentional control settings can be flexibly altered to direct more attention towards salience, as the relevance of salience increases in a task.

1R15EY030247

Undergraduate Just-In-Time 2

Tuesday, May 20, 2:45 – 6:45 pm, Banyan Breezeway

56.339 THE ODDBALL EFFECT EXTENDS TO DYNAMIC STIMULI

Ian Triplett[1], Cathleen Moore[1]; [1]University of Iowa

The oddball effect (OE) is a temporal illusion, according to which, rare stimuli (oddballs) are perceived as having a greater duration than more frequent events (standards) even when their physical durations are identical. Aside from frequency, oddballs are typically defined by their having a physical feature which varies from that of the standards (e.g. size, color, shape). In general, the greater an oddball's "oddness", the greater the magnitude of the OE. However, this relationship is not strictly linear nor is oddness the only relevant factor. As demonstrated by Tse et al. (2004), the magnitude of the OE can be modulated by dynamics. Specifically, an expanding oddball among stationary standards induced a greater OE than a stationary oddball among expanding standards. The same degree of oddness (expanding vs. stationary) resulted in OEs with different magnitudes, suggesting that dynamics may play a special role in the OE. To test this, we replicated the serial-streaming paradigm from Tse et al. (2004) with all dynamic stimuli — i.e. disks that continuously changed color. The oddball feature was a dynamic increase in size that either began immediately after stimulus onset (Experiment 1) or after a variable, short delay (Experiment 2). Robust OEs were found in both experiments, confirming that (1) the OE occurs even when comparing two dynamic stimuli — i.e. stimuli which have at least one feature which is constantly changing; and (2) that the oddball does not need to be recognized as such from the outset — i.e. a stimulus can be identified as an oddball after the onset of stimulus presentation via the introduction of an unexpected feature change.

56.340 THE EFFECTS OF PRIORITIZATION ON THE ALLOCATION OF ATTENTION

Mimi Juffe, Stephen Emrich[1]; [1]Brock University

How is spatial attention affected by the allocation of visual working memory (VWM) resources? To address this question, we used the capture-probe paradigm intermixed with a two-alternative forced choice (2AFC) VWM task. We manipulated VWM resource allocation by changing the amount of priority given to a specific item in the 2AFC task. In Experiment 1, the capture-probe letters were presented at the locations of the memory stimuli after a delay of 700 ms. As in previous studies, VWM performance increased as the likelihood of an item being probed increased, consistent with flexible resource allocation. In the letter-probe trials, correct letter recall for the letter in the prioritized location increased in conditions with higher levels of prioritization. These results reveal that the distribution of spatial attention follows the allocation of VWM resources. In Experiment 2, we addressed the question: since spatial attention is affected by the allocation of VWM resources, when is this allocation of spatial attention happening? To test if spatial attention allocation happens during encoding or maintenance, we reduced the delay between the offset of the memory stimuli and onset of the letters to 100 ms. In the 2AFC task, VWM performance was not significantly different from performance in the first Experiment, indicating that participants were distributing VWM resources just as efficiently as in Experiment 1. In the letter-probe trials, correct letter recall in the prioritized location was lower for conditions with low and no priority when compared to Experiment 1. However, correct letter recall for the prioritized location in the highest priority condition was not different from the first Experiment. These results indicate that spatial attention allocation happens between 100 – 700 ms for items of low importance but occurs within 100 ms for items of high importance.

56.341 PREDICTING INDIVIDUAL DIFFERENCES IN VISUAL SEARCH USING MEASURES OF ATTENTIONAL BREADTH AND SACCADIC INHIBITION

Bradley F Stewart[1] (bradleystewart49@gmail.com), Zachariah Weir[1], Chloe Alvarado[1], Kelly Karagias[1], Ryan V Ringer[1], Carly J Leonard[1]; [1]University of Colorado

In a world of abundant sensory stimulation, people must move their eyes to sample the visual environment. Previous studies have shown those with slower saccadic latency during visual search demonstrate higher saccadic accuracy and complete the task with fewer eye movements. In the current experiment, we use well-established measures to understand how attentional and inhibitory functioning relate to these individual differences in oculomotor behavior. A useful field of view (UFOV) task estimated attentional breadth by finding peripheral contrast-sensitivity thresholds, followed by a dual-task block requiring both central and peripheral responses. A saccadic stop signal task (SST) tested inhibitory abilities by asking participants to cancel planned eye movements when a stop signal appears. It was hypothesized that individuals who perform better on the UFOV task would have longer first saccade latencies as they spend more time accumulating peripheral information before initiating their eye movements. Additionally, poor inhibition, as measured by the SST, was predicted to contribute to faster saccade latency, limiting time to accumulate peripheral information before an eye movement decision. The results replicate previous findings showing that individual differences in first eye movement latency were significantly correlated with fixation count and saccadic accuracy during visual search. Regression analyses were conducted to determine which attentional and inhibitory measures predicted oculomotor behavior in visual search. For models of first saccade latency, there was a complex interaction of attentional and inhibitory factors. This was further influenced by a measure indicating how sensitive individuals were to target eccentricity during search. Overall, the results demonstrate that individual differences in eye movements during visual search are at least partially related to variability in attentional and inhibitory functioning as measured by these independent tasks. Nevertheless, further research is needed to investigate what other factors play meaningful roles in determining individual differences in saccadic timing during visual search.

Funded by National Eye Institute & Office of Behavioral and Social Science Research (R15EY035056 to CJL).

56.342 ATTENTION TO COLOR AND ORIENTATION SPREAD GLOBALLY TO A SIMILAR EXTENT

Nora Indenberga[1] (noraindenberga@gmail.com), Seeun Oh[1], Sydney T. George[1], Melisa Menceloglu[1], Taosheng Liu[1]; [1]Department of Psychology, Michigan State University

Feature-based attention operates globally, modulating neural responses to stimuli inside and outside of the focus of spatial attention, which is assumed to enable guided visual search. Because feature dimensions differ in their guiding power (e.g., color > orientation), this may be also reflected in their ability to spread. Here, we compared the global spread of attention to color and orientation. We used a dual-task paradigm where participants maintained central fixation and completed a primary task on one side and a secondary task on the other side of fixation. In the color version, participants were cued to attend to one of two overlapping red and blue dot fields to detect occasional dimming in the primary task, while discriminating which of two separate groups of dots in 8 colors contained an overrepresentation of red or blue in the secondary task, regardless of the cue. In the orientation version, participants were cued to attend to vertical or horizontal lines to detect occasional dimming in the primary task while discriminating which of the fields of lines with 8 orientations contained an overrepresentation of vertical or horizontal lines in the secondary task. For both, the feature cue was only predictive for the primary task. We compared accuracy on the secondary task as a function of match vs. mismatch between the cued and overrepresented feature to quantify the global spread. We observed a spread effect for color and only marginal effect for orientation with no difference between those effects. In a follow-up experiment, we strengthened the attentional selection of the feature for the primary task and observed greater and significant spread effects both for color and orientation, which were comparable. Our results may suggest that the differences in guiding power between features may be due to reasons other than their ability to be selected globally.

Funding Information: NIH (R01EY032071) and NSF (2019995)

56.343 CUEING COLOR BUT NOT ORIENTATION EFFECTIVELY GUIDES VISUAL SEARCH WITH HETEROGENEOUS DISTRACTORS

Seeun Oh[1] (ohseeun@msu.edu), Nora Indenberga[1], Sydney T. George[1], Melisa Menceloglu[1], Taosheng Liu[1]; [1]Department of Psychology, Michigan State University

Attention to visual features can guide visual search. While research has shown that color is a more powerful guiding attribute than orientation, this may partly be explained by stimuli parameters and distractor heterogeneity favoring color selection. Here, we tested this with closely-matched visual search tasks that isolated feature-based top-down guidance. Participants searched for a target T among distractor Ls (1T and 7Ls or all 8Ls). In the color task, the target could be either red or blue while distractors were in six other colors and red and blue. In the orientation task, the target could be either vertical (0° rotated upright letter) or horizontal (270° rotation) while distractors were in six oblique orientations or 0° and 270°. If a target was present, distractors did not share its feature. In both tasks, participants experienced cued and uncued blocks. In the cued blocks, the color (red/blue) or orientation (vertical/horizontal) of the target was cued with 100% validity and participants were instructed to use them to search for T. No cues were provided in the uncued block. We found that while color cueing significantly speeded responses, orientation cueing actually slowed responses. We reasoned that the letter stimuli might have rendered the orientation cueing detrimental as the stimuli consisted of orthogonal lines. Thus, we conducted a follow-up experiment with the same paradigm but using elongated shapes (rectangles with rounded or pointed ends) instead of letters. We found that color cueing speeded responses while orientation cueing had no effect on response times. Overall, our findings support previous reports and indicate that color more effectively guides visual search than orientation, in a well-matched design with heterogeneous distractors. The dominance of color may be due to its greater invariance and saliency. Future studies may explore the stimulus parameters that can enable effective orientation cueing.

Funding Information: NIH (R01EY032071) and NSF (2019995)

56.344 THE CAPACITY OF ACTIVE ATTENTIONAL TEMPLATES: TESTING THE SINGLE-TEMPLATE VS MULTIPLE-TEMPLATE HYPOTHESES

Sydney George[1] (georgesy@msu.edu), Xiaoli Zhang[1], Melisa Menceloglu[1], Taosheng Liu[1]; [1]Michigan State University

Many studies have shown that multiple features can be used to guide attention, but it is still debated precisely how multiple templates are maintained in visual working memory. In the current study, we aimed to differentiate two hypotheses: a multiple-template model where more than one template is maintained simultaneously; or a single-template model, where only one template can be active at a given time with the possibility of switching. We used a brief stimulus duration to prevent template switching in order to examine these competing hypotheses. Participants were shown a dot array (33 msec) with multiple colors and were asked to decide whether there was an over-represented color. In separate blocks, participants were either given no cue, one cue, or two cues at the beginning of the trial. The cue(s) were 100% predictive of the over-represented color when presented. Analysis of detection sensitivity suggests that participants can use either one or two features to guide attention. We then examined accuracy for target present and absent trials separately and found that compared to the no-cue baseline, the hit rate for the one-cue condition was the highest, followed by the two-cue condition, which is predicted by both hypotheses. Given our design, the multiple-template model predicts two weaker templates in the two-cue condition as compared to one strong template in the one-cue condition, which would result in a higher false alarm rate (FA) in the two-cue condition. On the other hand, the single-template model would predict similar FA rates if switching is not possible. The results showed that FA rates were significantly higher in the two-cue condition than the one-cue condition. This indicates that two weak templates were maintained in the two-cue condition, which supports the multiple-template hypothesis. We propose that attentional guidance of multiple features may be enabled by simultaneously active templates.

This work was supported in part by grants from NIH (R01EY032071) and NSF (2019995).

56.345 IRRELEVANT FEATURES OF DELAY-PERIOD MEMORY TARGETS CAUSE VISUAL WORKING MEMORY DISTORTIONS.

Ryan Miller[1], Blaire Dube[1]; [1]Memorial University of Newfoundland

Representations in visual working memory (VWM) can be distorted by irrelevant perceptual information when the items share a feature dimension. When tasked with remembering the orientation of a rotated bar across a delay, for instance, exposure to an irrelevant, differently oriented bar during the delay causes a subtle shift in the remembered orientation towards the distractor orientation. This attractive bias is robust, but can be attenuated when delay-period perceptual input is memory-relevant (e.g., when the orientation of the second bar must also be encoded). Such attenuation of attractive bias has been demonstrated when single-feature items are encoded sequentially, and therefore share a single relevant feature dimension. Are VWM representations also distorted by a shared but irrelevant feature of a delay-period item when another of its features (e.g., its colour or location) must be encoded? In Experiment 1, two differently oriented bars were presented sequentially and participants reported the orientation of both; results replicated previous work showing no attractive bias in the reports of the initially encoded item. In Experiment 2, two differently oriented bars were again presented sequentially, but the second oriented bar appeared in a random location on the screen and participants were instructed to remember only its spatial position and ignore its orientation. Reports of the first orientation exhibited attractive bias toward the irrelevant orientation of the second memory item. Remembering a feature other than the orientation of the second bar fails to attenuate attractive bias, suggesting that robust distractor interference can be driven by an irrelevant dimension of an otherwise memory-relevant stimulus. A single delay-period item can serve as a memory target while simultaneously exerting distractor-specific biasing effects on an existing item in VWM.

56.346 USING GENERAL RECOGNITION THEORY TO CHARACTERIZE FEATURE BINDING IN VISUAL WORKING MEMORY

Madilyn Dungan[1] (madilyn.p.dungan-1@ou.edu), Nadia Iranpour[1], Michael Wenger[2], Sarah Newbolds[2]; [1]Undergraduate, [2]University of Oklahoma

This study investigates the binding of sound and color in working memory, examining how these features interact within the context of attention-driven memory processes. Previous research has explored this question from various theoretical and methodological angles, but we apply general recognition theory (GRT) to mathematically assess the relationship between sound and color in working memory. GRT characterizes binding in three ways: violations of informational independence, informational separability, and decisional separability. These constructs are defined mathematically, allowing for rigorous tests for how sound and color features are or are not bound together. Stimuli consisted of a single visual form (triangle) whose color (purple) varied in saturation (low, medium, and high), that was paired with a pure sine-wave tone that varied in frequency/pitch (low, medium, and high). Three working memory conditions were tested using an n-back memory task at three levels of demand: 0-, 1-, and 2-back. In the 0-back condition, participants identified the saturation and pitch of the current stimulus pairing. In the 1-back and 2-back conditions, participants compared the current stimulus with the one presented 1 or 2 items before. Data were analyzed at the level of individual participants. Although there were individual differences, the GRT analyses generally pointed to independence and separability in both immediate perception (0-back) and working memory (1- and 2-back), with there being very few failures of marginal response invariance. Results are considered in the context of current accounts of feature binding in working memory and suggest a variety of steps to be pursued in this line of work.

56.347 THE PERCEPTION OF COUNTABILITY: A CASE STUDY OF 'MENTAL AFFORDANCES'

Lana Milman[1], Ian Phillips[1], Chaz Firestone[1]; [1]Johns Hopkins University

In addition to physical actions (e.g., climbing a staircase, or grasping an object), we also perform mental actions (e.g., counting objects in our head, or shifting attention). Recent work in philosophy of mind proposes that, just as we can appreciate whether and how easily we can execute various physical actions (physical affordances), we can also do the same for mental actions — appreciating in advance how effectively we will be able to execute a certain cognitive operation before actually carrying it out (the "mental affordance hypothesis"; McClelland, 2020). Here, we explore this hypothesis for counting and its corresponding mental affordance "countability" — i.e., how quickly and accurately an array of objects can be precisely counted. Subjects briefly (500ms or 2500ms) saw two "cookies" (circles) containing within them a number of "M&Ms" (dots). The M&Ms varied in size (large, small, or mixed), color (single or mixed), and opacity (full or partial). Subjects selected whichever cookie seemed easier to count (causing it to reappear onscreen), and then went ahead and counted that cookie's M&Ms. Results showed that even a 500ms preview was sufficient for subjects to accurately predict many aspects of their own counting performance on a given display, including that larger M&Ms would be easier to count than smaller M&Ms, that opaque M&Ms would be easier to count than semitransparent M&Ms, and so on. However, they also "misperceived" countability: Subjects preferred to count cookies with mixed-size M&Ms over smaller-sized M&Ms, even though counting performance was better on the latter. Our results suggest that naive observers can rapidly form impressions of a mental affordance and use it to guide behavior. Moreover, like physical affordances, we may be imperfectly calibrated to our actual capabilities.

Hopkins Office of Undergraduate Research Bloomberg Distinguished Professor Summer Program Award

56.348 SELECTIVE ATTENTION BY CONTEXTUAL SPATIOTEMPORAL REGULARITIES

Kaiki Chiu[1] (kaikichiu@gmail.com), William Narhi-Martinez[1], Irene Echeverria-Altuna[1], Pranava Dhar[1], Anna (Kia) Nobre[1]; [1]Yale University

Our current behaviour is often guided by long-term memories triggered by present contexts. Extensive research has established how learned spatial regularities proactively guide attention, yet the impact of spatiotemporal regularities in long-term memory remains less understood. Building on the work of Cravo et al. (2017), who demonstrated through behavioural and EEG results that contextual temporal associations can guide attention, we investigated the dynamics of learning spatiotemporal regularities and their enduring influence on attention. To probe how spatiotemporal regularities are acquired and maintained over time, we designed a paradigm in which participants completed a discrimination task within repeated contexts (across eight real-world scenes) during a one-hour session. During each trial, participants differentiated between a target object and multiple foil objects (50/50 probability), which appeared at varying locations within each scene and at distinct time points (early or late) relative to scene onset. Four scenes contained contextual spatiotemporal regularities (predictable; targets appeared at consistent locations and times within each scene), while the other four contained no spatiotemporal regularities (unpredictable). The assignment of target/foil objects and predictable/unpredictable scenes was counterbalanced across participants, and all scenes were presented in random order within each of the five completed blocks. Behavioural results revealed that over time, participants became progressively faster at identifying targets in predictable compared to unpredictable scenes, showing evidence that context-specific spatiotemporal associations were gradually learned and used to speed performance. An explicit post-task memory test further indicated that these performance enhancements did not depend on participants' awareness of the spatiotemporal regularities within each context, as no significant interactions emerged between explicit knowledge of spatiotemporal regularities and behavioural benefits. These findings provide evidence that contextually cued spatiotemporal expectations can be learned to guide attention and flexibly deployed to optimise behaviour.

56.349 CAN ARTIFICIAL INTELLIGENCE ELIMINATE THE NEED FOR EYE TRACKER CALIBRATION IN BOTH SCREEN-BASED AND WALKING TASKS?

Elizabeth Aje[1] (eaje@iu.edu), Stephanie Shields, Kathryn Bonnen; [1]Indiana University

Eye tracking has been an important method in vision science since its introduction in the 19th century, but for most of that time, eye trackers had to be attached to laboratory setups that kept the participant's head in a fixed position. Only relatively recently has the development of head-mounted eye tracking allowed for the recording of eye movements while participants move both their head and body. Unfortunately, the freedom head-mounted eye trackers allow also makes accurate calibration more difficult. One potential strategy for improving calibration is the use of artificial intelligence (AI). Baumann & Dierkes (2023) report that the Pupil Labs Neon eye tracker successfully uses a deep learning approach to accurately track eye movements and estimate gaze direction in a wide variety of use cases, without requiring any calibration. We sought to validate their findings by testing the Neon's gaze estimation accuracy using two screen-based tasks and two walking tasks. To test for an impact of viewing angle, participants performed screen-based tasks in one of three configurations: (1) while sitting, looking straight ahead; (2) standing, looking slightly downward; and (3) standing, looking at the ground. To test for an impact of slippage, participants performed the screen-based tasks both before and after the walking tasks. Overall, our results suggest that the Neon estimates gaze accurately, even without applying participant-specific corrections. Accuracy sometimes varied across viewing angles and across repeats, but the differences we observed were relatively small and were not consistent across participants. We did find differences in the overall level of accuracy across participants, and our next steps will be to explore the cause of those differences. Additionally, future work will test how much of an improvement the Neon provides over mobile eye trackers that require calibration.

56.350 THE INFLUENCE OF WORKING MEMORY LOAD ON VISUAL SEARCH AND QUITTING BEHAVIOR

Laurielle Kelly Feuzeu Mekoue[1], Vrishak Duriseti[1], Libby Kotei-Fearon[1], Natalie Chace[1], Jeffrey Moher[1]; [1]Connecticut College

We examined the relationship between visual search and working memory, particularly focusing on quitting behavior when performing visual search while simultaneously holding information in working memory. Study participants completed an online task with two parts. First, participants were shown one (low-load), two (medium-load), or four (high-load) colored squares. Then, they completed a visual search task where they searched for a vertical blue line target among slanted lines. Targets were present on a randomly selected half of all trials. Finally, they were asked to indicate whether a single square matched the color and location of one of the squares that was previously presented. We hypothesized that when participants have higher working memory loads, they will reach their decision threshold quickly during search, resulting in shorter response times and an increased amount of miss errors. As expected, we observed that as working memory load increased from one to two to four squares, memory accuracy significantly dropped from approximately 90% with 1 square to 70% with 4 squares. Miss rates in search also increased with higher working memory loads, which would be consistent with early quitting. However, on target-absent trials, reaction time increased from 2.44 seconds (s) with 1 square to 2.53 s with 2 squares, then spiking to 2.98 s with 4 squares. This result is not consistent with early quitting with high working memory loads, but may support the idea that increasing memory load generally disrupts search, as more time was taken to search and more errors were made. This is consistent with shared cognitive resources between the two tasks but not consistent with our initial hypothesis that increasing working memory load would decrease search time. This data provides an understanding of how multitasking impairs decision-making by influencing the relationship between working memory, distractor suppression, and visual search.

This work was supported by the National Science Foundation Grant BCS-2218384 to JM.

56.351 EARLY QUITTING IN VISUAL SEARCH: HOW DIFFERENT CUEING METHODS INFLUENCE PERFORMANCE

Idah Mbithi[1] (imbithi@conncoll.edu), Alexa Warren[2], Aleena Seyedzadeh[3], Melisa Sierra[4], Jeff Moher[5]; [1]Connecticut College, [2]Connecticut College, [3]Connecticut College, [4]Connecticut College, [5]Connecticut College

The Early Quitting Effect refers to the tendency of individuals to prematurely abandon a visual search task when encountering distractions. When distractors or inaccurate cues cause participants to quit early, miss errors are higher on target present trials, and response times are shorter on target absent trials. In this study, we explored how different cue types influence early quitting. Participants completed a series of visual search trials indicating whether a target "T" was present among nontarget "L's". There were three conditions, randomly assigned between participants: a control group with no cues, a group where one letter was highlighted with a circle cue, and a pre-trial cue group where a red fixation cross cue appeared at the location of one letter for 500 milliseconds before search display onset. The cues sometimes highlighted the target, and sometimes a non-target. We first focused on accuracy on trials in which the cue highlighted a non-target, and a target was present elsewhere in the display. The pre-trial fixation cross increased accuracy (56%)compared to the red circle cues (50%), but was still lower than the control condition (66%). This suggests that the pre-trial cue decreases one aspect of the early quitting effect relative to circle cues. However, when the cue highlighted the target, accuracy dropped with the pre-trial cue (72%) compared to the red circle condition (92%). Additionally, response times (RTs) on target absent trials were similar across both cueing conditions. These data suggest that cue types impact visual search and influence the early quitting effect. While the pre-trial fixation cross cue decreased miss rates for inaccurate cues, it also proved to be less effective in guiding attention towards targets. In future studies, we will explore alternative cuing approaches that may effectively guide attention without triggering early quitting when cues are inaccurate.

This work was supported by the National Science Foundation grant BCS-2218384 to JM

56.352 HOW DOES EFFORT AVOIDANCE RELATE TO THE STRATEGIC USE OF ATTENTIONAL CONTROL?

Saachi Kuthe[1], Tianyu Zhang[2], Andrew Leber[3]; [1]The Ohio State University

As humans, we frequently use visual search, but the strategies people use to engage in search tend to be inefficient. Why would people choose suboptimal strategies? Previous research has suggested that people avoid effort related to specific cognitive processes required to implement the optimal strategy. However, evidence for such effort avoidance has only been demonstrated across groups of participants, and we have yet to find a task that is sensitive enough to link individuals' effort avoidance to their strategy choices (see Zhang & Leber, 2024). To better quantify individuals' effort avoidance, we designed a new paradigm, the Voluntary Engagement Task, which provides participants the option of completing a trial of a task or doing nothing at all. We expected this new paradigm could reveal subtle individual differences across varying demand conditions by measuring the number of trials completed in a fixed period. In this study, we used the Adaptive Choice Visual Search task (ACVS; Irons & Leber, 2016) to measure individuals' search strategies. In addition, we created two modified conditions. The first required a numerosity judgment, the assumed key component required by the optimal strategy in ACVS, to find the target. The second, a control condition, did not include a numerosity judgment. We calculated the difference in voluntarily completed trials between the numerosity and control conditions to assess the degree to which participants selectively avoided performing the numerosity judgment. Results showed that this avoidance metric was significantly correlated with optimality in the ACVS task, indicating that the more participants avoided the numerosity judgment, the less optimal their search strategies were. Overall, this study demonstrates how individuals' effort avoidance of specific cognitive components predicts their search behaviors and furthers our understanding of the important role of cognitive effort in driving individuals' choice of attentional strategies.

56.353 MEASURING AFFECTIVE PROCESSING THROUGH PUPIL DYNAMICS DURING CONTEXT-BASED EMOTION PERCEPTION

Jasmine Lopez[1] (jaslopezm@berkeley.edu), Jefferson Ortega[1], David Whitney[1]; [1]UC Berkeley

The ability to perceive the emotions of others is influenced by how we physiologically respond to social cues. Pupil size, in particular, reflects fluctuations in emotional arousal (Oliva & Anikin, 2018) and may serve as a dynamic marker of real-time emotion processing. While previous work has linked pupil size to emotional salience, it remains unclear whether moment-to-moment pupil dynamics can predict individual differences in emotion perception performance. In this study, participants (N = 83) completed the Inferential Emotion Tracking (IET) task, in which they watched eight emotionally evocative video clips and continuously rated the affect of a target character using a two-dimensional valence and arousal grid. The character's facial expressions were blurred, encouraging reliance on contextual cues to infer emotion. Pupil size was recorded throughout using the Eyelink 1000 eye-tracking system. To assess whether pupil dynamics tracked emotion processing, we computed intersubject-pupil correlations for each participant—measuring how closely their pupil size time series aligned with the group average (excluding themselves) for each video. We then tested whether these pupil synchrony scores predicted IET task performance, defined as the correlation between each participant's ratings and the leave-one-out group average. Results revealed a significant positive relationship between pupil synchrony and valence rating accuracy (Pearson $r = 0.25$, $p = 0.026$), but not arousal accuracy (Pearson $r = 0.12$, $p = 0.29$). This suggests that individuals whose pupil fluctuations more closely mirrored the group were better at tracking emotional valence, but not necessarily emotional arousal. These findings support the idea that pupil dynamics reflect meaningful aspects of real-time emotion processing, particularly for valence. As a non-invasive and accessible physiological signal, pupil size may serve as a valuable marker for studying emotion perception in naturalistic contexts and populations with limited behavioral responses.

Supported in part by the National Institute of Health (grant no. R01CA236793) to D.W and (grant no. 1F99NS141343) to J.O.

56.354 FREQUENCY OVER TIME MATTERS: TEMPORAL EXPOSURE SHAPES IMPLICIT FAMILIARITY

Thiti Chainiyom[1], Chaipat Chunharas[1]; [1]Cognitive Clinical & Computational Neuroscience Lab, Faculty of Medicine, Chulalongkorn University

Are we more likely to notice something that comes back after being gone, or something that slowly slips away? Kind of like a familiar song that briefly disappears from the radio and then returns. This study explores how the timing and frequency of exposure—rather than meaning—shape implicit learning and familiarity. We used basic visual shapes (colored squares or circles) with no emotional or semantic content and showed them in two patterns: first the shape would appear frequently, rarely occur, and then back to frequent again. Then, it would turn into a B shape (reverse pattern). All stimuli were shown the same number of times and were counterbalanced across participants. We created a visual target detection test consisting of 480 trials with 16 participants (n=16). We started with a baseline of 160 trials (equal exposure to A and B), followed by 10 blocks of 32 trials (16 A and 16 B). Over time, we gradually changed how often the A and B stimuli appeared. We looked at performance across two phases (Phase 1: blocks 1–5, Phase 2: blocks 6–10), analyzing both accuracy and response time (RT). Participants responded faster to B stimuli as the task went on, while accuracy was higher for A stimuli, especially during Phase 2. Participants showed significantly different responses between conditions, as indicated by a paired-samples t-test ($p = .026$). Although accuracy was high overall, only the A condition showed a clear improvement during Phase 1. Our results suggest that people can pick up on patterns just from when and how often something appears. The A condition felt more familiar when it came back, possibly reflecting memory reactivation.

56.355 FLICKER AND READING: DOES THE PHANTOM ARRAY IMPAIR READING?

Sidney K Hulsey[1], Alayna M McGuire[1], Sarah M Haigh[1]; [1]University of Nevada, Reno

Flickering lights are common in modern environments, partly due to the efficiency and flexibility of light emitting diodes (LEDs). While certain flicker frequencies are known to produce a variety of visual symptoms including discomfort, illusions, migraines, eye fatigue, and seizures, little is known about its effects on cognition. LEDs can flicker at much faster rates than the standard 120Hz from older fluorescent lights. Originally, faster flicker was believed to be innocuous; however, the human visual system can resolve flicker at much higher rates during eye movements. The flicker is briefly apparent as a stripped pattern known as the phantom array. In this study, focused on how detection of the phantom array affects cognition, specifically errors during reading aloud. We compared phantom array detection accuracy under 600Hz flicker (where visibility of the phantom array is maximal), to number of errors when reading stripy and less stripy text under two flicker conditions: 600Hz and 60,000Hz (perceived as a steady light). Across nine participants, we found that participants made more errors when reading the stripy compared to non-stripy text (p=.034), particularly for the 60,000Hz flicker condition (p=.056). We also found that the number of errors made under the 600Hz flicker correlated with the accuracy in detecting the phantom array at 600Hz (r=.63). Together, this suggests that flickering light impairs reading and that the impairment is related to the sensitivity to flicker – in this case, the ability to detect the phantom array. Specific effects of flicker will support calls to regulate LED electronics to limit the effects of flicker on cognition in daily life.

Funding: ENDURE NINDS NS119709

56.356 CONVOLUTIONAL NEURAL NETWORKS LACK HUMAN-LEVEL ROBUSTNESS AT RECOGNIZING REVERSED-CONTRAST AND TWO-TONE MOONEY FACES

Zelin (Linda) Zhao[1] (zz3130@nyu.edu), Connor Parde[1], Frank Tong[1]; [1]Vanderbilt University

Although convolutional neural networks (CNNs) have attained human-level accuracy at identifying faces (e.g., O'Toole & Castillo, 2021), it remains unclear whether their visual strategies match those of human observers. Here, we compared human and CNN performance using image manipulations known to challenge human perception. Mooney faces are two-tone black and white images that can be perceived as faces while lacking clear facial features (Mooney, 1957). Reversing contrast polarity (i.e., photographic negation) further impairs face recognition by disrupting the interpretation of shading and shadows. We tested humans and CNNs with four conditions: blurry grayscale images (Gaussian blur, sigma 2 pixels), blurry grayscale images with reversed contrast polarity, thresholded Mooney images, and Mooney images with reversed contrast polarity. Participants viewed face images of 10 different celebrities and were asked to report the face identity. Each face image was presented with one of the four image manipulations. Likewise, the entire set of images was presented to two different CNNs (AlexNet and VGG19), both trained on the FaceScrub database. Humans outperformed CNNs overall with the performance gap becoming more pronounced under more challenging conditions. Average human accuracy was 97.1, 69.8, 76.2, and 53.1% for blurry faces, reversed-contrast blur, Mooney, and reversed-contrast Mooney, respectively, whereas CNN accuracy was 96.7, 24.2, 37.5, and 18.0%. CNN face identification performance was disproportionately impaired by reversing contrast polarity, with accuracy dropping sharply in the reversed-contrast and Mooney conditions. By comparison, humans showed greater robustness to these manipulations, highlighting fundamental differences in face-processing strategies. Our findings reveal the vulnerability of CNNs when tasked with identifying faces in ambiguous contexts and generalizing to novel visual conditions.

Supported by NEI grant R01EY035157 to F.T.

56.357 METACOGNITIVE ABILITY LIMITS VISUAL ATTENTIONAL CONTROL

Sydney Trowbridge[1], Owen Vegas[2], Michael Grubb[2], Silvia Lopez-Guzman[1]; [1]Unit on Computational Decision Neuroscience, National Institute of Mental Health, National Institutes of Health, Bethesda, MD., [2]Psychology Department, Trinity College, Hartford, CT.

Attentional Blink (AB) is the decreased ability to detect a second target when it follows an initial target by ~200ms and has been widely studied in perceptual neuroscience. AB appears to improve after engaging in metacognition through a meditation intervention. We hypothesized that these training effects could be limited by differences in individuals' ability to introspect. We addressed this question by measuring AB before and after completing a perceptual decision-making task where participants had to introspect and report their decision confidence. 72 participants completed the experiment online. In the AB task they had to report on the presence of a second target (T2) after identifying an initial target (T1), in a rapid stream of visual stimuli. In the metacognition task, participants were presented with four visual gratings, with three oriented vertically at 0° and one tilted to the left or right by a range of 0.5°-30°. Participants indicated with a keystroke whether the orientation of the tilted Gabor was clockwise or counterclockwise, then reported their confidence that they were correct on a 1-4 scale. AB data was fit to a signal detection theory (SDT) based linear model through hierarchical Bayesian estimation, to derive individual-level decision criterion and decision sensitivity and their change pre and post confidence task. Confidence data was fit to the CASANDRE model (also rooted in SDT) to derive individual confidence sensitivity. SDT modeling revealed individuals with greater metacognitive ability exhibited higher decision sensitivity in the baseline AB task. However, we found no significant relation between hit rate in the AB task and confidence sensitivity either at baseline or pre vs. post, displaying the confounds associated with standard hit-rate analyses commonly used in AB literature. Given that better metacognition is related to enhanced attentional control, future studies could address whether introspection and visual attention share common mechanisms.

This research was supported (in part) by the Intramural Research Program of the NIMH - ZIAMH002988

56.358 METACOGNITIVE PRACTICE INDUCES CONSERVATIVE RESPONSE BIASES

Owen Vegas[1], Sydney Trowbridge[2], Silvia Lopez-Guzman[2], Michael Grubb[1]; [1]Trinity College, Hartford CT, [2]Unit on Computational Decision Neuroscience, National Institute of Mental Health, National Institutes of Health

Cognitive training interventions like meditation practice and video game playing have been shown to increase attentional control, but the impact of metacognitive training has yet to be fully examined. We designed an experiment to test the hypothesis that making metacognitive judgements about perceptual decisions induces state changes in attentional control. 186 participants completed an attentional blink task, which required reporting the identity of target 1 and the presence/absence of target 2 at varying temporal lags. Next, participants completed a simple orientation discrimination task: one of four Gabors was rotated clockwise or counterclockwise from vertical. Depending on group assignment, participants then reported their confidence in their orientation decision (1-4 scale), reported the location of the non-vertical stimulus (1 of 4 locations), or made no secondary report. All participants then redid the attentional blink task. We analyzed the data with a signal detection theoretic, linear mixed effects model fit with Hierarchical Bayesian estimation, designed to quantify the change in sensitivity and response criterion from pretest to posttest in each group. Sensitivity quantifies participants' ability to correctly identify the presence of target 2, reflecting attentional control. Criterion quantifies participants' biases to report present or absent. We found that sensitivity significantly increased in the confidence group from pretest to posttest, as well as in the single judgment group. We conclude that these improvements were likely due to practice effects. Notably, the localization group did not experience these improvements. Interestingly, metacognitive practice impacted response criterion: the confidence group became more conservative in the posttest, while the single judgment and localization groups became more liberal. In short, repeated self-assessment of one's confidence in an orientation discrimination task induced conservative response biases in a subsequent detection task, highlighting the impact of metacognitive practice on visual decision making.

This research was supported (in part) by the Intramural Research Program of the NIMH - ZIAMH002988, and by NSF-2141860 CAREER Award to MAG.

56.359 LUMINANCE-MATCHING IN PUPILLOMETRY IS NOT ENOUGH: THE CURIOUS CASE OF ORIENTATION.

Matthew Parrella[1], Isshori Gurung[1], Michael Grubb[1]; [1]Trinity College

Abrupt onsets reflexively shift covert spatial attention. Our lab recently demonstrated that delivering trial-to-trial information about the probability of a peripheral onset modulated the magnitude of the attentional cueing effect (low-probability > high-probability). Although onsets were physically identical, pupil responses could have been modulated by the onset-probability signal, which provided information about the probability of the onset's appearance. Specifically, anticipatory constrictions may have preceded high-probability onsets. Here, we tested this possibility in a new experiment using centrally-presented, luminance-matched onset-probability signals. For half the participants, vertical lines at fixation signaled high probability (0.8) of onset appearance (a small, white, peripheral circle), while horizontal lines signaled low probability (0.2). These contingencies were reversed for the other half of participants. Importantly, both onset-probability signals had equal luminance. Participants fixated the onset-probability signal for 2,000 milliseconds before the onset was presented for 67 milliseconds, or omitted altogether, in line with the signaled probability. To maintain engagement, participants completed a simple localization task. We measured pupil area through the experiment using an EyeLink 1000. Preliminary evidence for an expectation-induced anticipatory constriction was obtained in Experiment 1 (n=15). However, this effect disappeared in Experiment 2 with a larger replication sample (n=39). Exploratory analyses of the two datasets uncovered a violation of a fundamental methodological assumption: despite being task-irrelevant, luminance-matched, and presented at fixation, a vertical onset-probability signal consistently generated larger pupillary constrictions, relative to the presentation of a horizontal signal. Statistical evidence for this orientation-dependent modulation was present in both experiments independently, and a combined analysis confirmed that this effect was not confounded by expectation. Our assumption—that one could eliminate a differential pupil response by using luminance-matched stimuli presented at the same location—was wrong. This seemingly benign manipulation demonstrated the real effects of stimulus design, no matter how inconspicuous, on pupillometric results.

Supported by NSF-2141860 CAREER Award to MAG

56.360 PORTABLE PSYCHOPHYSICS: MEASURING VISUAL CROWDING WITH THE APPLE VISION PRO

Kaila Dowd[1], Kamilla Volkova[1], Michael Grubb[1]; [1]Trinity College

Visual crowding is a perceptual phenomenon in which a peripheral stimulus is identified more easily when presented in isolation than when surrounded by flankers. To ensure that stimuli appear in the periphery, eye-tracking is essential. Here we leveraged the eye-tracking capabilities of the Apple Vision Pro, a relatively cheaper and immensely more portable alternative to commercial eye-trackers, to demonstrate a feasible use case of the Apple Vision Pro for visual psychophysics. In this project, we measure crowding and replicate the finding that covert exogenous spatial attention reduces the critical distance. On each trial, we presented a target T at one of four possible orientations, which was flanked above and below by Hs or Is. Then, participants reported the orientation of the T, with their eyes and a finger tap. We used eight distinct spacings between the target and flankers. Before the stimulus array, a brief onset appeared in either the same (valid condition) or opposite (invalid condition) hemifield, eliciting the reflexive allocation of covert spatial attention. Fourteen participants completed the study. For each participant and each attention condition, we calculated proportion correct at each spacing and fitted an exponential model to the data to obtain the parameters needed to calculate the critical distance for each cue type. The Wilcoxon signed rank test verified that valid cues resulted in a lower critical distance than invalid cues. We also ran a 2x8 repeated measures ANOVA where we found significant main effects of cue type and stimuli spacing on accuracy. In short, we were able to successfully measure visual crowding and replicate previous work on the impact of exogenous cueing using the Apple Vision Pro. This is an exciting proof-of-principle demonstration, attesting to the possibility of using cheap, portable, and immersive VR systems for visual psychophysics in a diversity of settings.

Supported by NSF-2141860 CAREER to MAG

TUESDAY AFTERNOON POSTERS IN PAVILION

Visual Memory: Working memory and attention

TUESDAY, MAY 20, 2:45 – 6:45 PM, PAVILION

56.401 ATTENTIONAL PRIORITIZATION AND DEPRIORITIZATION IN PERCEPTION AND VISUAL WORKING MEMORY

Issam Tafech[1], Karla Matic[1,2,4], Polina Iamshchinina[3], Thomas B. Christophel[1,2]; [1]Humboldt University of Berlin, Germany, [2]Bernstein Center for Computational Neuroscience Berlin, Germany, [3]Princeton Neuroscience Institute, USA, [4]Max Planck School of Cognition, Germany

Attentional prioritization has consistently been shown to improve information maintenance and recall. However, it has remained an open question whether attentional prioritization affects visual working memory and perceptual encoding in similar ways. Prior work has used pre-stimulus and post-stimulus cues to prioritize perceptual input and working memory content respectively, allowing for their comparison. However, experimental variations such as cue predictability and the sensitivity of the report task have led to mixed findings on the relative strength of these effects. Here, we aimed to quantify and compare the effects of prioritization in perception and in working memory on recall error using pre-cues and post-cues, respectively. In our design, participants had to do a continuous recall task, reporting one of two oriented gabor patches. Two separate experiments were run: in Experiment 1, the cues were probabilistic (75% report cued item, 25% report uncued item); in Experiment 2, the cues were deterministic (100% report cued item) but participants were cued twice per trial, possibly requiring an previously uncued item to be reported later. We showed that in both experiments, both prioritization in perception and in working memory have significant effects on performance. By comparing performance for cued and uncued items to a neutral condition with no pre-response prioritization, we found improved recall accuracy for cued items (prioritization effect) and reduced recall accuracy for initially uncued items (deprioritization effects). In both experiments, prioritization and deprioritization effects were stronger in perception than in working memory. The initial prioritization effect on cued items was similar in both tasks and participants reported similar levels of subjective prioritization. Furthermore, subsequent effects of prioritization and deprioritization on the second reported items were weaker than initial effects. These results show that attention can modulate perceptual input more strongly than mnemonic representations and suggest that prioritization effects are subject to decay.

This work was supported by DFG Emmy Noether Research Group Grant CH 1674/2-1

56.402 ATTENTIONAL REFRESHING MULTIPLE WORKING MEMORY OBJECTS VIA EXOGENOUS RETRO-CUES

Juyeon Joe[1], Min-Shik Kim[1]; [1]Yonsei University

Working memory (WM) is an important cognitive function that enables the temporary storage and manipulation of information necessary for daily life. Attention enhances WM performance and is often studied through internal attention manipulations such as retro-cues. Recent studies suggest that non-predictive retro-cues can selectively refresh WM contents (Souza et al., 2015; van Ede et al., 2020). However, it remains unclear whether nonpredictive exogenous retro-cues can facilitate the maintenance of multiple objects in WM. To address this, we investigated the effect of attentional refreshing on multiple objects. Participants were instructed to memorize four different colored shapes, and after the memory array disappeared, two of the four placeholders flickered to direct internal attention to specific objects in WM. Participants then determined whether a test probe presented at the end of each trial matched any of the memorized objects. For non-matching probes, feature combinations varied across conditions: containing both shape and color, only one feature, or no features from the WM items. Sensitivity was higher for cued compared to uncued objects, demonstrating that non-predictive exogenous retro-cues effectively refresh multiple WM objects. However, the presence of cued features in nonmatching probes did not improve rejection accuracy. Additionally, accuracy decreased as the number of shared features between the test probe and WM objects increased. These findings suggest that while attentional refreshing in WM operates on bound object representations, WM performance is constrained by feature-based processes. Thus, non-predictive exogenous cues can enhance attention to multiple objects, but only when the objects' features are integrated.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2024S1A5B5A19043380)

56.403 BIND 2, GET 1 FREE?: WILL THE VISUAL SYSTEM ENCODE A THIRD FEATURE EVEN WHEN IT'S IRRELEVANT?

Moussa Kousa[1], Brad Wyble[1]; [1]The Pennsylvania State University

Attribute amnesia (AA) revises the conventional understanding that attention is the key determinant in how information is retained in memory. It demonstrates that participants can fail to remember an attended attribute, such as the color identifying a target, immediately after using that same attribute to detect the target. The AA effect is believed to reflect a cognitive efficiency mechanism. We theorize that reporting a single feature in typical AA tasks does not require linking it to a specific object token, leading to the other features of the target stimulus being discarded from memory. This project modified the traditional AA task by requiring participants to report two target attributes, aiming to promote their integration into a unified object token and potentially bypass the mechanisms underlying AA. The study included two experiments: a control and an experimental condition. Both used bar stimuli defined by color, texture, and orientation. In the control condition, participants reported one attribute in pre-surprise trials and were tested on another in a surprise trial, following standard AA procedures. The experimental condition differed by requiring participants to report two attributes in pre-surprise trials. We hypothesized that the attribute amnesia effect, marked by improved accuracy from the surprise to the first control trial, would occur in the control experiment but not in the experimental condition. However, the results indicate a statistically significant improvement in accuracy from the surprise trial to the first control trial, demonstrating an attribute amnesia effect in both conditions—though less pronounced than what is typically observed in standard AA tasks. Closer analysis of sub-conditions revealed that when color served as the key attribute, the AA effect was absent. This finding suggests a potentially unique property of color for targets that are large textured shapes, which contrasts with previous research, which found an AA effect with color.

56.404 BOTH ACTIVE AND PASSIVE STATES IN VISUAL WORKING MEMORY RECRUIT SENSORY STORAGE

Xinran Chen[1] (12439002@zju.edu.cn), Huixin Song[1], Mowei Shen[1], Hui Chen[1], Yingtao Fu[1]; [1]Zhejiang University

Visual working memory (VWM) has been thought to be supported by sensory storage and has a close interaction with visual perception. However, a series of behavioral and neural evidence emerging in recent years demonstrates that only the active, but not the passive state in VWM relies on sensory storage, raising the debate regarding whether VWM selectively recruits sensory representation. The current study aims to delve into this debate further by testing whether increasing the load of active/passive states in VWM affects detection sensitivity to an incoming visual stimulus, a paradigm which has been verified to specifically uncover the sensory nature of working memory storage. In Experiments 1-3, we consistently found that loading either active or passive VWM state impaired the visual detection to a similar degree, suggesting that both states involved sensory storage. Experiment 4 validated the state manipulation by observing dissociative memory-driven attentional bias effect of different states. Experiment 5 showed that information released from VWM no longer impaired visual detection, further confirmed the specific role of VWM storage (in either active or passive state) in interfering with sensory processing. Together, the current findings demonstrate that VWM relies on sensory storage regardless of representational states.

56.405 THE IMPACT OF RETRO-CUE VALIDITY ON WORKING MEMORY AND ATTENTIONAL TEMPLATE EFFICIENCY

Alexandre Fortuna[1] (alexandre.fortuna@unige.ch), Martin Constant[1], Dirk Kerzel[1]; [1]University of Geneva, FPSE

An attentional template refers to features in visual working memory that guide attention in search tasks. Across three pre-registered experiments, we compared predictions from the state-based model (Olivers et al., 2011) and the resource model (Huynh Cong & Kerzel, 2021) concerning the effects of retro-cues in dual target search. In our experiments, participants memorized two colors at the start of a trial and one of the two colors was cued retrospectively. After the retention interval, participants performed either a search task where they indicated the orientation of the stimulus in one of the two potential colors, or they performed a memory task where they indicated one of the two colors on a color wheel. The two tasks occurred randomly and with equal probability. The retro-cue promoted the status of attentional template because search and memory performance for the retro-cued color improved. In addition, effects of retro-cueing increased when the validity of the cue increased from 50% to 70%, suggesting that participants allocated working memory resources continuously according to the validity of the retro-cue. The continuous performance adjustments are at odds with the state-based model, which postulates discrete switches between attentional template and accessory status but support the resource model.

Swiss National Science Foundation (SNSF 10001E_219517)

56.406 THE STABILITY OF INDIVIDUAL WORKING MEMORY CAPACITY MEASURES

Courtney Turner[1] (courtney.l.turner@durham.ac.uk), Ella Williams[2], Roger Remington[3], Anna Grubert[1]; [1]Durham University, [2]Oxford University, [3]University of Minnesota

The ability to temporarily memorise visual information is an essential cognitive ability used in everyday life. Such information is held in visual working memory (vWM) and the capacity of this short-term storage differs individually. Importantly, higher levels of vWM capacity have been linked to higher levels of fluid intelligence and enhanced cognitive abilities such as logical reasoning, problem solving and comprehension. However, such correlations are only meaningful if individual vWM capacity is stable and the measures to assess vWM capacity are robust. We tested this assumption by means of five change detection tasks in which we manipulated visual/perceptual stimulus parameters between conditions and measured the effects of these manipulations on individual vWM capacity as measured with Cowan's K. In each trial, participants were presented with a memory display (100ms) with differently coloured squares. After a retention period (900ms), they were shown a test display (100ms) and had to decide whether it was identical to the memory display or contained a colour change. In different tasks, we manipulated the number of memory items (Experiment 1), the eccentricity of the memory items from fixation (Experiment 2), the figural organisation of the memory items (Experiment 3), the featural identity of the memory items (colour, shape, alphanumerical category; Experiment 4), or the predictability of the stimulus locations in the test display (Experiment 5). Individual K values changed significantly between task conditions in all experiments (apart from Experiment 3), suggesting that individual vWM capacity measures are not robust against changes in visual/perceptual task parameters. However, in each experiment, the different K values produced in the different task conditions were positively correlated. Taken together, these findings suggest that K values may lack validity as absolute measures of vWM capacity, but that they seem to be reliable measures of relative individual vWM abilities.

This work was funded by research grants of the Leverhulme Trust (RPG-2020-319) awarded to AG.

56.407 VISUAL WORKING MEMORY INTRUSIONS CAUSED BY SALIENT ONSETS, WITH AND WITHOUT SPATIAL DISTRACTION

Derrek T. Montalvo[1] (montalvo.46@osu.edu), Blaire Dube[2], Julie D. Golomb[1]; [1]Department of Psychology, The Ohio State University, [2]Department of Psychology, Memorial University of Newfoundland

The visually complex environments around us make it difficult for our brains to select relevant information while ignoring irrelevant items. Thus, our brain uses various mechanisms to encode relevant information into our visual working memory (VWM), while filters prevent irrelevant information from being encoded. Dube and Golomb (2021) recently showed that spatial distraction not only captures external spatial attention, but also disrupts non-spatial attentional filters responsible for gating access to VWM, resulting in the encoding of irrelevant features at the time of capture (WM intrusions). However, is spatial distraction required to disrupt these attentional filters and cause WM intrusions, or is the presence of a salient stimulus alone enough? Participants performed two consecutive search tasks on each trial. In search 1 (S1), participants located a target 'T' among non-target 'L's presented within task-irrelevant colored squares. In 40% of the trials, a salient cue (white border) abruptly flashed around the target square (onset at 50ms, lasting 100ms), serving as a salient visual event while keeping spatial attention focused on the target. In search 2 (S2), participants located a uniquely oriented Landolt-C. The stimuli in S2 were all white, except one color singleton that either matched the color of the S1 target, color of a S1 non-target, or novel color. We tested if the S1 salient cue would cause its associated, task-irrelevant color to intrude into VWM and drive memory-driven capture in S2 (slower RTs when the singleton matched). Despite no disruption to spatial attention in S1 (S1 RTs were faster on cue-present vs. cue-absent trials), memory intrusions were triggered by the salient cue: on cue-present trials, RTs were longer when the S2 singleton color matched the S1 target square, compared to non-target matches and cue-absent matches. This suggests that perceptually salient cues alone can disrupt attentional filters and cause WM intrusions.

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

56.408 VISUAL WORKING MEMORY LOAD MODULATES DISTRACTION IN VISUAL SEARCH

Jennifer Trujillo[1], Bo Yeong Won[2]; [1]California State University, Chico

Visual working memory load (WML) influences distraction in visual search, yet it remains unclear how different types of WML—working memory capacity (WM capacity) and working memory precision (WM precision)—interact with different types of distraction, such as distraction from a similar distractor versus multiple dissimilar distractors. This study investigated two key questions: (1) whether memory load modulates distraction, and (2) whether WM capacity and WM precision loads interact differently with distraction. In Experiment 1A (WM capacity), participants completed a dual task involving memory and visual search. They memorized either a low-load (2 colors) or a high-load (4 colors) set of colors while searching for a target color among colored distractors (one similar color or three dissimilar colors). After the search, they judged whether a test color was part of the memorized set. In Experiment 1B (WM precision), a separate group performed a similar task but memorized a single color following a precision cue indicating low or high precision. After the search, they reported the memorized color on a color wheel. Results revealed that high WML reduced distraction regardless of the type of WML or distraction. However, WM capacity and WM precision interacted differently with distraction. Specifically, participants were less distracted by a similar color distractor compared to multiple dissimilar distractors during the WM precision task relative to the WM capacity task. These findings were replicated in Experiment 2, where participants completed a "T among Ls" search task instead of a color search, and WM capacity and WM precision trials were randomly intermixed within subjects. Overall, this study demonstrates that WML reduces distraction in visual search. Notably, remembering precise information in WM reduces distraction from similar distractors more effectively than from multiple dissimilar distractors, suggesting that the mechanisms underlying detailed information processing may govern both WM and attentional control.

56.409 WORKING MEMORY CONSOLIDATION INTERRUPTION BY FORCED CHOICE DECISION IS LOCATION SPECIFIC

Benjamin J. Tamber-Rosenau[1], Paulina A. Kulesz[1], Brandon J. Carlos[1,2]; [1]University of Houston, [2]Ball State University

Preserving visual items in working memory entails a consolidation process that results in a distraction-resistant representation. An initial, rapid stage of consolidation, sometimes called encoding, is terminated by a mask, while a second, slow stage continues after masking. However, the second stage of consolidation can be interrupted by a decision task (T2) that follows sample presentation by up to 1 s, with memory performance reduction attenuating over sample-T2 delay. Because of the cross-representational-format nature of this interference (Nieuwenstein & Wyble, 2014, DOI: 10.1037/a0035257; Carlos et al., in revision), consolidation is thought to depend on central executive processes, and it has been argued that interference with consolidation by T2 reflects a structural limit (Carlos et al., 2023, DOI: 10.3758/s13414-023-02757-7). Such a central structural limit is not expected to be location specific—that is, the location in which the visual sample is presented should not relate to the degree of interruption of WM consolidation by T2. To our surprise and in contrast to numerous past studies, in two cohorts of participants (N=23, N=30), we observed no interruption of consolidation by the presentation of a number parity decision T2. The lack of consolidation interruption was observed across multiple kinds of visual memoranda and set sizes. It does not reflect a failure of sensitivity, as we still observed the typical visual masking (first stage of consolidation) effect. What can explain the missing interference? In past studies, both the sample and T2 were presented centrally. Here, T2 was presented centrally, but the sample was presented at 10° eccentricity, in peripheral vision. Thus, we suggest that the interruption of working memory consolidation from vision by a decision task is surprisingly location specific.

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

Visual Memory: Working memory and visual functions

Tuesday, May 20, 2:45 – 6:45 pm, Pavilion

56.411 VISUAL WORKING MEMORY AND THE PRIMARY VISUAL CORTEX: BEHAVIORAL ASYMMETRIES AND STRUCTURAL CORRELATES

Julia Papiernik-Kłodzińska[1,2], Simon Hviid del Pin[3], Michał Wierzchoń[1,4], Marisa Carrasco[5,6], Renate Rutiku[1]; [1]Consciousness lab, Institute of Psychology, Jagiellonian University, Krakow, Poland, [2]Doctoral School in the Social Sciences, Jagiellonian University, Krakow, Poland, [3]Norwegian University of Science and Technology, Gjøvik, Norway, [4]Centre for Brain Research, Jagiellonian University, Krakow, Poland, [5]Department of Psychology, New York University, [6] Center for Neural Sciences, New York University

Whether the primary visual cortex (V1) plays a role in working memory performance remains unclear. We explored this question by testing whether polar angle asymmetries, namely the vertical meridian anisotropy and horizontal-vertical asymmetry, along with visual field asymmetries (upper-lower and left-right) are reflected in individual working memory performance variation across the visual field and compared them against the individual brain structure. 262 participants performed an object recognition working memory task, adjusted to measure the visual asymmetries. Additionally, all participants underwent a structural MRI scan during a separate session. We used a multiparametric mapping sequence to acquire four different brain maps per participant: MT, PD, R1, and R2*, each contrast denoting different aspects of the brain microstructure. The regions of interest were V1, the intraparietal sulcus (IPS), and the frontal eye field (FEF), with the first one known to reflect the polar angle and visual field asymmetries, and the latter two known to be linked to working memory processing. The behavioral results show, that the working memory performance is impacted both by the visual field and polar angle asymmetries, with the vertical meridian anisotropy appearing reversed compared to the previous literature. The MRI results showed that the individual asymmetries in visual working memory performance do not explain the differences in the volume of FEF and IPS. On the other hand, the differences in volume of V1 were impacted by the vertical meridian anisotropy, with significant results present in three out of four MPM maps, and R2* results being the most systematic. This suggests that V1 plays a role in working memory performance, as well as that there may be a link between the iron content in V1 tissue, and the elevated polar angle asymmetry in visual working memory.

This work was supported by the National Science Centre, Poland (projects 2021/42/E/HS6/00425 to Renate Rutiku and 2017/27/B/HS6/00937 to Michał Wierzchoń for data collection) and by the European Cooperation in Science and Technology (COST, Action CA18106).

56.412 A DENSE SAMPLING STUDY ON VISUAL WORKING MEMORY ACROSS THE HUMAN MENSTRUAL CYCLE

A.Y li[1] (alisonlee1115@gmail.com), H Grotzinger[1], E Murata[1], E.G Jacobs[1], T.C Sprague[1]; [1]University of California, Santa Barbara

Working memory (WM) is the ability to temporarily store and manipulate information to guide later behavior (Baddeley, 2010; Ma et al., 2014), and recent studies have shown that WM information is stored by distributed activity across early visual cortex and higher areas including parietal and prefrontal cortex (PFC) (Christophel et al., 2017; Curtis & Sprague, 2021). While most studies average data across participants, there is important evidence that these areas (especially PFC) show variations in neural function across the menstrual cycle, and measures of cognitive function (e.g., n-back) reflect these hormone-related fluctuations (Jacobs & D'Esposito, 2011). However, this previous work does not address how these hormone fluctuations impact aspects of visual WM performance, including capacity, precision, and inter-trial serial dependence (Fischer & Whitney, 2014; Bliss et al., 2017). This is important, because visual WM is a tractable system for modeling relationships between neural function and behavior in humans (Li et al., 2021). Here, we employed dense sampling methods (Pritschet et al., 2021) to assay WM performance across n = 6 participants' natural menstrual cycles. At each of ~15 sessions per participant (approximately every other day), we measured WM capacity (Change Localization Task; Zhao et al., 2023) and WM precision (memory guided saccade [MGS] task, Funhashi et al., 1989; Li & Sprague, 2023), along with salivary measures of ovarian hormones (estradiol/progesterone) and survey measures of state anxiety, sleep quality, and caffeine intake. We observed fluctuations in visual WM performance as measured by MGS response time, precision, and magnitude of serial dependence throughout the menstrual cycle, while WM capacity was remarkably stable. Importantly, these differences could not be explained by variations in measures of sleep quality or anxiety. This implies that the rapid fluctuation of ovarian hormones may be responsible for the change in visual WM performance measured with the MGS task.

UCSB Undergraduate Research and Creative Activities Awards; Sloan Research Fellowship; Ann S. Bowers Women's Brain Health Initiative

56.413 ASSOCIATING MEMORIZED COLOR WITH SPATIAL LOCATION IMPROVES RECALL AND BIASES GAZE

Amit Rawal[1,2] (amitrawal1643@gmail.com), Michael J. Wolff[1], Rosanne L. Rademaker[1]; [1]Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with the Max Planck Society, Frankfurt, Germany, [2]Faculty of Behavioural and Human Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands

Visual working memory (VWM) allows brief maintenance of information to serve behavioral goals. Recently, Henderson and colleagues (2022) highlighted this goal-oriented nature, demonstrating that a predictable response allows people to flexibly change a visuo-spatial memory representation into a motor-based one. Can such flexibility also be observed between two visual features, color and space? In this eye-tracking study, participants performed two blocks of trials during which they remembered a centrally presented (250ms) color for 3 seconds, subsequently reporting it on a color-wheel via method-of-adjustment. Critically, this color-wheel was either randomly rotated on every trial (random block) or was identical on every trial (fixed block). In a third block, participants remembered a spatial location. When people remember a color in the fixed color-wheel blocks, specific colors are tied to fixed locations. Do people take advantage of this inherent spatial information to aid behavioral performance? And is this reflected in gaze behavior? Confirming that people can utilize color-space associations, we find higher accuracy and quicker responses in fixed compared to random color-wheel blocks. This replicates comparable recent work (Bae & Chen, 2024). Spatial biases in gaze can reflect the location of a prioritized item (van Ede et al., 2019), which we replicate during the delay of the single-item spatial memory block. Importantly, there was also a bias when the color-wheel was fixed, but none when it was random (as color was not predictive of space). During recall, a further gaze bias towards the correct color was observed, which was larger in the random compared to fixed color-wheel blocks. This may reflect a lack of integration of spatial knowledge in random color-wheel blocks. Together, these results show that behavioral performance can be improved by incorporating spatial information to aid color memory, and that systematic changes in gaze can index this flexible utilization of VWM.

56.414 BIDIRECTIONAL INTERFERENCE BETWEEN WORKING MEMORY AND PERCEPTION FOR FACES AND MOTION OVER TIME

Sanika Paranjape[1], Simon Kaplan[1], Sarah Shomstein[1], Dwight J. Kravitz[1,2]; [1]The George Washington University, [2]US National Science Foundation (SBE/BCS)

The sensorimotor recruitment theory of visual working memory (VWM) posits that the same cortical areas are recruited for both encoding and maintenance, suggesting that memory influences ongoing perception, and ongoing perception influences memory. Previous studies show that this bidirectional relationship exists for low-level features (e.g. orientation and color). However, it is unclear whether this bidirectional relationship is unique to low-level features or extends to high-level complex stimuli. We hypothesized that, for motion direction and face identity, stimuli held in VWM influence ongoing perception, consistent with the common recruitment of motion area MT and face-selective FFA/OFA for VWM and perception. In a series of experiments, two discrimination conditions were used: 1) the middle condition, in which the two discrimination stimuli are equidistant from the maintained stimulus (in a continuous stimulus space); and 2) the side condition, in which both discrimination stimuli are off to one side. The maintained information predictably interfered with discrimination, both decreasing and increasing discriminability depending on the relationship of the maintained and discriminated information. Moreover, the discrimination stimuli introduced systematic biases in the continuous report of the maintained information. We further tested the effect of delay, hypothesizing a theta rhythm in the strength of interference effects. For motion, the strength of this interference was modulated by the time between encoding and discrimination, evidencing sustained interference and systematic theta oscillations. This oscillatory pattern and decay over time was not found for faces, highlighting the robust and complex nature of faces and their resistance to decay. Overall, the results suggest that sensory recruitment is a general VWM mechanism, not limited to maintenance of low-level features. The inconsistency between faces and motion in the temporal pattern of the interference is consistent with the recruitment of distinct neural circuits for different visual information, resulting in different specific dynamics.

NSF BCS-2022572 to SSS and DJK; W911NF-23-2-0097 to DJK and SRM

56.415 CAUSAL EVIDENCE FOR THE ROLE OF VISUAL AREAS IN WORKING MEMORY MAINTENANCE

Wing Kwan Hannah Chu[1] (wkc267@nyu.edu), Bas Rokers[1,2,3], Kartik K Sreenivasan[1,2]; [1]Psychology, New York University Abu Dhabi, [2]Center for Brain and Health, New York University Abu Dhabi, [3]Aspire Precision Medicine Institute Abu Dhabi, New York University Abu Dhabi

Items held in working memory (WM) can be decoded from fMRI activity in visual cortex, motivating the hypothesis that visual regions store information to support WM. Causal evidence for this hypothesis, however, is lacking; transcranial magnetic stimulation (TMS) studies have yielded decidedly mixed results, and cannot rule out unwanted downstream perturbation of higher order regions. We used artificial scotomas as a novel means of modulating visual cortical activity to test its role in WM. An artificial scotoma is a simulated blind spot, which can be induced by presenting texture or motion around the scotoma location. The scotoma location is filled in by information interpolated from the surround presumably via feedback from higher-order regions. Forty participants completed a WM task that required them to memorize the motion directions of two drifting gratings in opposite quadrants, hold these in WM over a memory delay, and report the motion direction of the grating cued at the end of the trial. During the delay, we induced an artificial scotoma by presenting a gray patch superimposed on a drifting plaid background. The scotoma could occur in one of three locations on each trial: the location of the to-be-probed (i.e., cued) memory item, the location of the uncued memory item, or a non-memory item location. Induction of the scotoma resulted in increased memory error, decreased memory precision, and a greater propensity to misreport memorized motion direction by 180°. Critically, performance was disrupted only when the scotoma was induced in the location of the to-be-probed memory item and only on trials when subjects confirmed that a scotoma was successfully induced. By using spatially selective artificial scotomas that target early visual processing, we overcome the limitations of previous fMRI and TMS studies to provide causal evidence that visual areas play an active role in WM storage.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYUAD Research Institute grant CG012

56.416 LEARNED AFFORDANCES AND ACTION SIMILARITY MODULATE REPULSION BIASES IN VISUAL WORKING MEMORY

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

Recent research suggests that action plans can influence sensory representations in visual working memory (VWM). For example, we found that two similarly oriented bars concurrently held in VWM were perceived as more dissimilar (i.e. repelled each other) when associated with different actions than when linked to the same action. In that study, the pairings between orientations and action plans were random, varying from trial to trial, indicating that the action-based repulsion effect resulted from rapid, short-term sensorimotor associations. In the current study, we examined whether the action-based repulsion effect is modulated by longer term associations. For this purpose, we implemented consistent mappings between specific to-be-remembered orientations and specific actions, and compared this to random couplings. Participants performed a task in which, in each trial, they memorized two bar orientations and then reproduced these using cued actions (either a grip or a slide action). Crucially, for some pairs, the bars were consistently associated with the same action (e.g., always grip) (fixed coupling condition), while for other pairs, they were randomly linked to either action (random coupling condition). Participants reported no explicit awareness of the fixed action-orientation associations, but when asked to guess did indicate the correct action at above-chance levels, thus providing evidence of implicit learning of action-orientation associations. We replicate our previous finding: orientations linked to different actions repelled each other more than those associated with the same action, across both conditions. Preliminary results further indicate that, in the fixed coupling condition, memory repulsion is larger than in the random coupling condition when the two bars are assigned to two different versus the same action. These preliminary findings suggest that action learning may enhance differences between VWM representations that belong to different action groupings, supporting the notion that affordances play an important role in structuring our memory landscape.

56.417 REMEMBERING MULTIPLE OBJECTS IN THE SAME CATEGORY LEADS TO BETTER PERFORMANCE AT A SACRIFICE OF HIGHER FALSE ALARM

Payachana Chareunsouk[1,2], Anantaporn Sena[1,3], Chaipat Chunharas[1,3]; [1]Cognitive Clinical & Computational Neuroscience Lab, Faculty of Medicine, Chulalongkorn University, [2]Medical Science, Faculty of Medicine, Chulalongkorn University, [3]Chulalongkorn University Chula Neuroscience Center, King Chulalongkorn Memorial Hospital

When remembering multiple objects, people often group them into chunks to improve memory. Previous studies showed that people group objects by their basic features, such as location, orientation, or color. In the real world, objects often possess multiple basic features and belong to semantic categories. Whether and how semantic categories influence visual working memory remains debated. An answer might not simply be a yes-or-no as individuals may use categorical labels as a gist, resulting in more stable memory but a higher probability of within-category false alarms. This study investigates the role of semantic categories in real-world object visual working memory. Eleven participants performed a delay-match-to-sample task involving five to-be-remembered items, with a 5-second viewing of an object array, followed by a 2-second delay. The number of objects within a single category was manipulated across conditions: 5 (all same category), 4, 3, or 1 (all different categories). Participants were then shown a single object and asked to indicate whether it appeared in the original array. Test objects included items from the initial array (50%), new objects within the same category (25%), and new objects from different categories (25%). Participants had more accuracy when all objects belonged to the same category than from different categories (79±8%, 69±4%, 65±7% for 5 , 4 , 3 objects in the same category respectively). However, this improved accuracy was accompanied by a slightly higher false alarm rate, suggesting participants relied on semantic labels for memory grouping (11±6%, 8±6%, 9±6% false alarm rate of 5, 4, 3 objects in the same category respectively). Our findings indicate that semantic categories enhance visual working memory performance, albeit at the cost of increased within-category false alarms. This aligns with long-term memory research and highlights the potential interplay between semantic organization and working memory, offering insights into real-world memory processes.

56.418 REWARDS GUIDE VOLUNTARY UTILIZATION OF VISUAL WORKING MEMORY

Xutao Zheng[1] (zhengxutaowk@foxmail.com), Chenxiao Guan[1], Mowei Shen[1], Jifan Zhou[1]; [1]Zhejiang University

Processing vast amounts of information with limited resources is a key challenge to human cognitive system. Visual working memory (VWM) is capable to retain only 3-4 objects. This number is even smaller in natural working memory tasks, where individuals need to repeatedly sample objects, simulating real-world scenarios. Research findings suggest that humans tend to minimize cognitive costs: when sampling cost is low, they sample more frequently but with fewer objects per sample to reduce VWM cost. In addition to minimizing costs, the voluntary utilization of VWM is supposed to maximize rewards for better cognitive outcomes in complex environments. This study adopts a copying task to examine how rewards influence the voluntary utilization of VWM. Participants copied six bars from a model to a workspace, autonomously deciding how many bars to copy each sample. The distance (long/short) between the model and the workspace determined sampling cost (high/low). In Exp.1, high or low rewards were assigned to the bars. We found high-reward bars were copied prior to low-reward ones and took longer time to copy, although accuracy reached a ceiling for both high and low-reward bars. Exp.2 replicated Exp.1's findings, with the bars changed to the same color to prevent potential semantic encoding. Exp.3 introduced a time limit, then accuracy for high-reward bars was significantly higher than that for low-reward ones. Also, we found copying group for high-reward bars increased slightly as sampling costs increased, which should beneficial for maintaining accuracy. Exp.4 raised reward for copying multiple bars at once, leading to a significant increase in the VWM load despite high cognitive cost. In conclusion, rewards drive the adjustment of voluntary VWM operations to increase gains, even when it requires greater costs. Our research from a cost-benefit perspective could deepen the understanding of VWM's mechanisms and its real-world applications.

the National Natural Science Foundation of China (Grants 32371088, and 62337001), Science and Technology Innovation 2030—"Brain Science and Brain-like Research" Major Project (2022ZD0210800), and the Fundamental Research Funds for the Central Universities (226-2024-00118).

56.419 DECISION-MAKING MODULATES PERCEPTUAL INTERFERENCE BEYOND SENSORY INTERACTIONS

Joseph M. Saito[1] (josaito@ucsd.edu), Philip Li[1], Timothy F. Brady[1]; [1]University of California San Diego

Human observers are known to exhibit biases in their reports of visual memories following interference by perceptual distractors, such that memories are reported as being more alike distractors than they actually were. Leading sensory recruitment models posit that these biases can be fully explained by interactions between sensory representations of memory and perception within the local circuitry of visual cortex. Here, we challenge these models by showing that observers' decision history modulates report biases beyond low-level sensory interactions. Across multiple experiments, we asked observers to explicitly compare a target stimulus held in visual working memory to a perceptual probe stimulus to determine their similarity before submitting a memory report. Target-probe pairs were sampled systematically across consecutive trials to form short runs of highly-similar or dissimilar combinations that were interleaved by critical pairs whose similarity was more ambiguous. Despite their fixed visual features, observers perceived the same ambiguous pairs to be similar more often when presented in runs of similar pairs than dissimilar pairs. More importantly, these history-related changes in perceived similarity were associated with corresponding changes in the magnitude of report biases, suggesting bona fide modulations in target-probe integration rather than mere priming of the similarity judgment. In a series of follow-up experiments, we extended this trial history effect by testing its generalizability across abrupt changes in stimulus type and by showing that changes in report bias are not explained by changes in memory strength. Finally, using simulations, we grounded out the predictions of sensory recruitment models to show that sensory interactions alone are unable to reproduce these behavioral patterns without the inclusion of separate decision-making processes that implement additional top-down weighting of mnemonic and perceptual information. Together, these findings highlight the necessary collaboration between early- and post-perceptual processes to determine the readout of memory contents following perceptual interference.

This work was supported by NSF BCS-2146988 awarded to TFB

56.420 **SCAN IT LIKE IT'S HOT: WORKING MEMORY AND** VISUAL SCANNING

Micalee Segers[1] (msegers@vols.utk.edu), McKenzie Sheets[1], Tori Jones[1], Shannon Ross-Sheehy[1]; [1]University of Tennessee Knoxville

Change detection tasks are commonly used as a measure of visual working memory (Luck & Vogel, 1997). In these tasks, adults are typically given instructions to encode and remember array items, then report if they detect a change from sample array to test. However, additional work has shown that memories are stronger when the changed item was overtly fixated during the sample array (Eschman & Ross-Sheehy, 2023). This is important, as participants who naturally scan more might be more likely to have fixated the to-be-changed item, resulting in especially strong memory. If overt fixation is advantageous, then we would expect adult viewers to fixate more array items when they are instructed to detect a color change, than when they are passively viewing arrays with no response requirement. To test this, 46 adults were tested in a change detection task (sample=1500ms, retention=500ms, test=3000ms) consisting of a passive block, during which participants were given no instructions, immediately followed by an active block in which participants were instructed to indicate if any of the array items changed color from sample to test. Eye-tracking was used to sample gaze at 500Hz. Preliminary results indicate that explicit instructions did influence array scanning $F(1,21981)=34.22$, $p<.001$; participants fixated more array items in the active block (3.37) than the passive block (3.19). Interestingly, this did not influence preference for the changed item, as change preferences did not differ between groups, $t(389.95)=-0.33$, $p=0.742$). This suggests that having an explicit task can influence array scanning, but this does not influence relative preference for the changed item. These results demonstrate that passive behaviors like change preference may reveal implicit change detection, and that these automatic processes are less influenced by task goal. Additional analyses will use passive change preference to explore the relationship between scanning strategy and subsequent working memory representations.

56.421 EXPLORING THE ROLE OF SACCADES IN VISUAL WORKING MEMORY UTILIZATION

Rania Qais[1] (raniaqais@mail.tau.ac.il), Shlomit Yuval-Greenberg[1,2]; [1]Israel

Visual working memory (VWM), the ability to maintain visual information, is typically studied using a change detection task where participants observe a set of items that disappear and then reappear either identical or altered and are asked to identify changes. Studies using this task have shown that observers can store around 3–4 items in their VWM. However, how this lab-based capacity relates to VWM utilization in natural environments remains unclear. Specifically, in real-world settings, items do not typically disappear and reappear as they do on a computer screen. Instead, the primary mechanism by which items are brought into or removed from the visual field is through saccadic eye-movements. Correspondingly, the central role of VWM is to store information across eye-movements. We present a novel task to explore the link between saccades and VWM. In two experiments, participants (N = 46) compared two sets of items to identify the item common to both. They either moved their eyes freely between the two sets (eye-switch) or switched between them by pressing a button (manual-switch). We measured the number of saccades performed within each set before switching to the second set, and used this as an index to assess the number of items stored in VWM, reflecting VWM utilization. Results showed fewer saccades within a set—indicating reduced VWM utilization—in the eye-switch condition compared to the manual-switch condition. Furthermore, increasing cognitive load (i.e. increasing the number of items) reduced VWM utilization in the eye-switch condition, while the opposite effect was observed with the manual-switch. These findings demonstrate that VWM remains underutilized in naturalistic settings, where saccades play a critical role in determining reliance on memory. They also highlight the importance of considering eye-movements when studying VWM and emphasize that task design influences the balance between working memory and perceptual strategies.

The study was funded by ISF grant 1960/19 to S-Y.G

56.422 RESPONSE MODALITY MODULATES SPATIAL AND TEMPORAL BIASES IN VISUAL WORKING MEMORY

Sihan Yang[1], Yueying Dong[1], Anastasia Kiyonaga[1]; [1]University of California, San Diego

Visual working memory (VWM) temporarily stores visual information to guide behavior but is also modulated by expected actions and contextual factors. Prior studies show that the same VWM information can evince different neural representations or be perceived differently depending on the expected response format. VWM recall can also be biased by the features of other items concurrently perceived (i.e., surrounding bias) or from the recent past (i.e., serial bias), and these contextual biases may be modulated by motor and response demands. Here, we show how the VWM motor response modality influences the manifestation of spatial and temporal contextual biases. Participants viewed two Gabor patches and reported their orientation after a short delay using a digital pen. In half of the blocks, they responded in the conventional way by continuously adjusting a probe wheel ('clicking'). In the other blocks, they reported by drawing a line that matched the remembered orientation ('drawing'). In general, we observed typical surrounding and serial biases in that recall was repelled from the concurrently-presented stimulus but attracted toward the relevant stimulus from the previous trial. We also found that the magnitude of both biases changed with response modality and predictability, even when perceptual inputs and memory content were matched. For instance, the serial bias was magnified for drawing responses (vs. clicking) when participants were able to prepare their motor plan during the delay. In a follow-up study using eye-tracking, we observed that gaze patterns showed progressively more serial bias across the delay for drawing conditions (but not clicking), revealing that peripheral oculomotor signatures can flexibly track the development of contextual biases. These results highlight how the integration of contextual information and visuomotor transformation interweave in VWM, emphasizing that VWM biases may be expressed differently depending on upcoming action demands.

56.423 TASK-RELEVANT ACTIONS TRIGGER AUTOMATIC WORKING MEMORY UPDATING

Sahcan Özdemir[1], Eren Günseli[2], Daniel Schneider[1]; [1]Leibniz Research Centre for Working Environment and Human Factors, [2]Sabanci University

It has been proposed that action control mechanisms can influence gating of visual working memory (WM) from task-irrelevant sensory input. Further, we hypothesized that a relevant action pattern could trigger WM updating through this mechanism, facilitating the interference of irrelevant sensory input on WM. To examine this idea, we conducted a preregistered (OSF) EEG experiment using a delayed-match-to-sample task. Participants memorized the color of a target stimulus and, after a retention period, identified its color on a color wheel. The target's shape indicated which hand they should use to adjust the wheel. During the retention phase, participants performed a secondary task requiring a response with either the right or left hand, which either matched or mismatched the hand used in the primary task. This secondary task introduced two types of interference: motor interference or visuomotor interference, with the latter involving a task-irrelevant color of the cue. First, EEG analysis revealed mu/beta suppression during WM encoding, indicating simultaneous preparation of action plans with visual encoding (H1). Second, visuomotor interference caused greater disruption to target representation compared to motor interference, as reflected in WM task accuracy (H2). Finally, interference involving the same hand as the main task triggered WM updating, evidenced by a stronger attraction bias of the target color toward the interfering color (H3). Further exploratory analyses showed increased frontal theta activity in the matched hand condition. This activity was negatively correlated with the attraction bias on a single-trial level. We suggest that WM updating with irrelevant sensory information is triggered automatically given the action-based relevance of interference; however, a control mechanism, reflected in frontal theta, actively responds to mitigate the impact of interference.

Visual Memory: Neural mechanism of working memory

Tuesday, May 20, 2:45 – 6:45 pm, Pavilion

56.424 DISTRACTED BUT NOT DEFEATED: VOXEL POPULATION CODES FLEXIBLY TRANSFORM TO SUPPORT HIGH-FIDELITY WORKING MEMORY REPRESENTATIONS

Janna Wennberg[1], Kirsten Adam[2], John Serences[1]; [1]University of California, San Diego, [2]Rice University

Working memory (WM) is surprisingly robust to sensory distraction, but there is debate about how neural codes protect WM against sensory interference (Kamitani & Tong, 2005; LaRocque et al., 2012; Rademaker et al., 2019). Some studies propose that memory representations transform within early visual areas to minimize interference (Libby & Buschman, 2021), while others suggest that these memory representations are offloaded to frontoparietal regions, reducing early visual regions' role in WM storage (Xu, 2021). We used fMRI to test whether early visual regions "multi-task" to concurrently support WM and the processing of incoming sensory information. Eight participants performed a spatial WM task, remembering the angular spatial position of a flickering checkerboard circle. On half of the trials, the screen was blank during the delay; on the other half, identical checkerboard circles flickered in other locations. We used circular ridge regression to decode spatial position, testing within-condition and cross-condition classification accuracy. The frontoparietal account predicts that sensory distractors degrade WM representations, leading to poor classification accuracy in early visual regions when training and testing on distractor-present trials. The flexible multi-tasking account predicts a cross-over interaction because mnemonic information is transformed but not lost. Decoding accuracy was similar in the distractor-present and distractor-absent conditions, consistent with the lack of a behavioral effect on recall error. However, we observed a crossover interaction as early as V1: Training on distractor-absent trials yielded better decoding for distractor-absent trials than distractor-present trials, and vice-versa. Cross-time generalization analyses showed a particularly robust cross-over interaction when training on late-delay period data. Notably, this interaction was absent in areas such as IPS. While our findings accommodate a prominent role for frontoparietal regions in storage and maintenance, they also suggest that voxel population codes in early visual regions flexibly transform to support storage of mnemonic information.

This work was supported by RO1-EY025872 awarded to John T. Serences

56.425 EEG DECODING REVEALS A LINK BETWEEN VISUAL WORKING MEMORY FIDELITY AND THE MAGNITUDE OF SIMILARITY-INDUCED MEMORY BIAS

Nursima Ünver[1,2], Rosanne Rademaker[3], Keisuke Fukuda[1,4]; [1]University of Toronto, [2]The Max Planck-University of Toronto Centre (MPUTC) for Neural Science and Technology, [3]Ernst Strungmann Institute for Neuroscience, [4]University of Toronto Mississauga

Visual working memory reports can be biased toward novel visual inputs presented during the delay. This bias is particularly strong when people report that these inputs look similar to the remembered target. In a prior study, we showed that the magnitude of this similarity-induced memory bias (SIMB) increased when certainty about the target decreased. This implies that the SIMB may be related to the strength of target encoding. To relate the magnitude of the behavioral SIMB effect to the neural fidelity of a target representation, we used EEG recordings together with multivariate decoding. Specifically, sixteen participants remembered the direction of a briefly (0.5s) presented dot motion stimulus (target) across a 2.5-second delay, after which they reported its direction. In one half of the trials, a second 0.5s dot motion stimulus (probe) appeared midway through the delay, and participants had to judge its similarity to the target. First, we replicated that when people judged the probe as "similar" to the target, target responses were biased toward the probe (i.e., the SIMB). Next, we successfully decoded the direction of target motion from the EEG for up to 0.4s after the target offset. We then hypothesized that, on trials in which the probe was judged to be similar to the target ("similar trials"), the magnitude of the SIMB would be large when the neural fidelity of the target representation is low. To test this, we median-split the similar trials based on the target decoding accuracy into the high- and low-fidelity trials and compared the SIMB magnitudes between them. Our results confirmed that the SIMB was reliably larger in the low-fidelity trials than in the high-fidelity trials. These findings demonstrate that neural fidelity, as reflected in decoding accuracy, is inversely related to the magnitude of SIMB.

This research was supported by the Natural Sciences and Engineering Research Council (5009170).

56.426 GOAL-SETTING MODULATES VISUAL WORKING MEMORY PERFORMANCE THROUGH ALPHA POWER SUPPRESSION AND EFFORT

Olga Kozlova[1], Kirsten Adam[1]; [1]Rice University

Because the capacity of Visual Working Memory (VWM) is strictly limited, previous research has aimed to improve VWM with feedback and monetary incentives. Often, however, the effects of feedback and incentives on VWM have been mixed. We have proposed an “optimal strategy” account to explain when feedback will be effective versus ineffective at improving VWM performance. Participants completed a behavioral whole-report VWM goal task in Experiment 1 (N=100), with EEG and pupillometry in Experiment 2 (N=22). At the beginning of each trial, participants were given performance goals (e.g., “remember 3 items”), and received feedback and monetary bonuses when goals were met. In Experiment 1, optimal VWM performance corresponded to goals that aligned to an individual’s typical VWM capacity (3 items). In contrast, supra-capacity goals harmed VWM performance by increasing lapses of performance. In Experiment 2, we investigated which underlying cognitive processes may support improvements to VWM with optimal goals. Specifically, we predicted that we should observe a “U-shaped” function for cognitive processes altered by goal-setting, with greater engagement for an “optimal goal” of 3 items compared to under- or over-ambitious goals. In contrast, we predicted that we should observe a monotonic increase for cognitive processes related to simple effort (Goal 5 > Goal 3 > Goal 1). First, we found a “U-shaped” pattern for posterior alpha power suppression ($p < .044$), suggesting that participants successfully maintained more items throughout the delay period when given an optimal goal. In contrast, we found a “monotonic increase” pattern for tonic pupil dilation ($p < .005$), suggesting that, paradoxically, participants exerted more effort, but remembered fewer items, for the suboptimal Goal 5 condition. Taken together, our results suggest that effective feedback can improve VWM performance, without concurrently increasing overall effort, by encouraging participants to allocate VWM resources to only a subset of items.

56.427 IMMEDIATE RESELECTION OF VISUAL AND MOTOR MEMORIES AFTER EXTERNAL INTERFERENCE

Daniela Gresch[1,2], Larissa Behnke[3], Anna C. Nobre[1,2], Sage E.P. Boettcher[2]; [1]Yale University, [2]University of Oxford, [3]University of Zurich

During natural behavior, we must often maintain internal representations in working memory while concurrently engaging with perceptual events in the external world. Though previous research has demonstrated working memory to be susceptible to external interference, internal representations generally withstand such challenges without complete failure. Here, we investigated the neural dynamics underlying the reselection of internal representations after engagement with an external task. Specifically, we asked which contents of internal representations are reselected after an interrupting task and when this reselection occurs. To address these questions, we developed a visual-motor working memory task in which participants were retrospectively cued about an item during the retention interval. In most trials, after a retro-cue was presented, participants were required to respond to a perceptual discrimination task. Using electroencephalography, we tracked the reselection of visual (i.e., spatial location) and motor (i.e., response hand) representations through contra- vs. ipsilateral modulations of posterior alpha (8–12 Hz) and central beta (13–30 Hz) activity, respectively. Our findings revealed the concurrent reselection of visual and motor contents immediately after the perceptual task was completed, rather than just-in-time when internal information was required for memory-guided behavior. These findings demonstrate the early reinstatement of visual working memory into a ready-to-use state and underscore the pivotal role of visual-spatial information in scaffolding internal representations after external interference.

56.428 NEURAL CORRELATES OF WORKING MEMORY PRECISION

Philipp Deutsch[1,2] (deutsch@med.uni-frankfurt.de), Cora Fischer[1,2], Jochen Kaiser[1,2], Christoph Bledowski[1,2], Benjamin Peters[3,4]; [1]Institute of Medical Psychology, Goethe University Frankfurt, [2]Cooperative Brain Imaging Center, Goethe University Frankfurt, [3]Institute for Language, Cognition and Computation, University of Edinburgh, [4]Centre for Cognitive NeuroImaging, University of Glasgow

A hallmark of working memory is its capacity limitation: memory performance decreases with more simultaneously memorized items. The neuronal origins of this limitation remain unclear. While capacity limitations are typically assessed by asking participants to report a single item, we applied a whole-report paradigm requiring participants to report all items. Specifically, participants memorized one, two or four orientations of Gabor gratings. After a delay, they sequentially reported each item at the locations indicated by external cues. For each item, participants provided two responses: 1) a confidence rating about memory quality (scale from 0=forgotten to 3=best possible remembered) and 2) a reproduction of the memorized orientation on a continuous scale. This allowed us to assess the number of items rated as forgotten and to estimate the memory precision for each item. We developed a composite measure of memory precision by integrating the confidence rating and orientation report, while accounting for the systematic impact of report order on memory. We recorded brain activity using magnetoencephalography while participants performed the task. Using beamforming, source level estimates of brain activity were obtained. We then applied multivariate pattern analysis to the activity across the human cortical surface. Preliminary results suggest that the activity patterns in posterior cortical regions predicted the aggregated precision of the reproduced orientations. This prediction was most pronounced during memory encoding but remained present until report. Moreover, activity patterns from posterior regions during similar task phases predicted whether a subject rated items as forgotten. Together, these findings suggest that working memory capacity is already limited by the initial processing of the stimuli. Future analyses will examine which factors contribute to the prediction signals including fluctuations of attention between trials or spatial positions and stimulus characteristics.

56.429 PINPOINTING THE SOURCES OF TOP-DOWN FEEDBACK IN VISUAL WORKING MEMORY

Shanshan Li[1,4] (sl9712@nyu.edu), Sarah L Master[1], Hsin-Hung Li[1,2], Thomas C Sprague[3], Ying Zhou[1,4], Kartik K Sreenivasan[4,5], Clayton E Curtis[1,6]; [1]Department of Psychology, New York University, [2]Department of Psychology, The Ohio State University, [3]Department of Psychological and Brain Sciences, University of California, Santa Barbara, [4]Division of Science and Mathematics, NYU Abu Dhabi, [5]Center for Brain and Health, NYU Abu Dhabi, [6]Center for Neural Science, New York University

How are working memory (WM) representations stored over brief delays? Neuroimaging studies consistently report persistent above-baseline activity in frontal cortex during WM maintenance but not visual cortex, while conversely, WM content can be decoded from visual cortex but less consistently in frontal cortex. Based on these findings, we hypothesized that persistent activity in frontal cortex reflects top-down feedback that enhances the fidelity of WM representations in visual cortex. To test this hypothesis, we developed a novel approach combining univariate analysis to assess trial-by-trial delay-period activation with decoding to quantify the quality of WM representations on a trial-wise basis. We applied this method to four WM datasets spanning three distinct studies, enabling a comprehensive test of our hypothesis and including testing how results generalize across tasks. Consistent with previous findings, we observed persistent activity in multiple frontal and parietal regions (but not visual cortex) during WM delays, and were able to decode memory context most robustly in V1-V3 during these same memory delays. Critically, across all datasets, trial-by-trial variations in delay activity amplitude in frontal and parietal cortices - specifically the superior precentral sulcus and intraparietal sulcus - predicted the quality of decoded WM representations in V1-V3. Thus, persistent activity in frontal and parietal cortices may reflect feedback signals targeting WM representations in visual cortex. We propose that these feedback signals may sculpt population activity in visual cortex, improving the quality memory representations.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYUAD Research Institute grant CG012 to KKS, and R01 EY-016407 and R01 EY-033925 to C.E.C.

56.430 REVEALING THE EFFECT OF REGULARITIES ON REPRESENTATIONS IN VISUAL WORKING MEMORY

Nursena Ataseven[1] (n.ataseven@rug.nl), Sahcan Özdemir[2], Daniel Schneider[2], Wouter Kruijne[1], Elkan G. Akyürek[1]; [1]University of Groningen, [2]Leibniz Research Centre for Working Environment and Human Factors

Previous studies have suggested that working memory (WM) can leverage regularities in the environment to increase its capacity. This highlights how memory resources can be used more efficiently when associations in long-term memory support WM. However, the mechanisms underlying this interaction remain unclear. Here, we ask whether such associations might alter the representation of items maintained in WM. Are long-term memory (LTM) supported representations "compressed," thereby freeing up capacity for novel information? To address these questions, we recorded EEG data (N=20) as participants performed a delayed comparison task, where they compared a target orientation grating to a probe and judged whether it was rotated clockwise or counterclockwise. Before the target presentation, a central color cue (red, green, or blue) was shown. In half of the trials, this cue informed participants about the 90° range from which the upcoming target orientation was drawn, while in the other half the cue was uninformative. Participants learned the associations between colors and orientation ranges prior to the experiment. We analyzed the neural representations in working memory by decoding orientation from posterior channels during the target orientation display. This allowed us to examine how different cue conditions influenced the encoding of orientation representations and the potential role of LTM in optimizing memory processes. The results suggest that behaviorally, the informative cue improved memory performance. However, this difference was hardly reflected in the EEG decoding. During target encoding, no differences were observed. Throughout the maintenance period, we observed a very subtle difference that did not reach significance, but became larger when probe presentation was imminent. These findings suggest that the associated information hardly affected the representation in working memory, but might factor into behavior at a later stage, for example during the comparison with the probe.

56.431 SHARED AND UNIQUE VISUAL WORKING MEMORY REPRESENTATIONS ACROSS INDIVIDUALS REVEALED BY INTRACRANIAL RECORDINGS

John Farley[1] (jpfarley@umd.edu), Ziam Khan[2], Alexander Ksendzovsky[2], Sara Inati[4], Yaoda Xu[3], Kareem Zaghloul[4], Weizhen Xie[1,2,4]; [1]Department of Psychology, University of Maryland, College Park, [2]Department of Neurosurgery, University of Maryland, Baltimore, [3]Department of Psychology, Yale University, [4]Surgical Neurology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health

The ability to maintain visual short-term/working memory (STM/WM) for goal-directed behavior is often attributed to delay-period activity in neocortical regions, such as the prefrontal cortex (PFC). However, recent evidence implicates the medial temporal lobe (MTL), traditionally associated with long-term memory, in visual WM. This raises the question of whether the PFC and MTL redundantly represent WM content or serve distinct roles. One possibility is that the PFC retains task-relevant stimulus information, including task structure and prior knowledge shared across individuals, while the MTL captures idiosyncratic, context-specific aspects of the encoded stimuli. Alternatively, both regions may represent the same WM content, communicated through coupled neuronal activity during retention, as suggested in recent studies. To test these possibilities, we recorded intracranial EEG from the PFC and MTL in seven participants performing a delay-match-to-sample task. Participants viewed a single object from one of eight categories (e.g., cats, cars) for 200 ms, followed by a 1300-ms delay, and judged whether a test object drawing from the same category matched the encoded object. Decoding analyses revealed object category information in both regions during the delay. However, representational geometry analyses uncovered striking differences: whereas PFC representations were remarkably consistent across individuals, MTL representations exhibited high variability. Using the same stimuli, prior fMRI research demonstrated that WM representations during delay aligned across sensory and posterior parietal cortices. The PFC representations revealed by intracranial EEG here aligned closely with those found in fMRI, whereas MTL representations showed weaker correspondence. These findings highlight complementary roles for the PFC and MTL in supporting visual WM: While the PFC provides stable representations that generalize across observers and regions, the MTL encodes more flexible, observer-specific representations. Combined, these mechanisms integrate shared and unique aspects of stimulus information, enabling robust WM for adaptive goal-directed behavior.

Y.X. is supported by NIH grant R01EY030854; W.X. is supported by NIH grant R00NS126492

56.432 TASK-DEPENDENT NEURAL CODING IN VISUAL WORKING MEMORY

Dennis Boakye[1] (boakyed@lawrence.edu), Khayla Santiago[2], Chunyue Teng[3]; [1]Lawrence University

Adaptive behavior often requires the flexible manipulation of task goals maintained in working memory under different task demands. For instance, the same piece of information can be used either as action-guiding content or contextual information. It remains unclear whether working memory representations could be reconfigured depending on their functionality. In this study, we investigated the temporal dynamics of the neural representations of content and context information in visual working memory, using electroencephalography (EEG) on human participants (n=15, ongoing data collection). We contrasted two hypotheses: 1) functional account that content information that necessitates direct action is maintained in a different neural code from the contextual background information; and 2) feature-specific account that the representational state of information in working memory is insensitive to functionality. Participants were asked to memorize the location and orientation of grated patches in two tasks: In the orientation response task (orientation content and location context), they recalled the orientation of the memory sample and reported either the exact memorized orientation or the orthogonal orientation, depending on whether the location of the probe matched the memory sample or not. In the location response task (location content and orientation context), they recalled the memorized location, based on the comparison between the orientation of the probe and that of the memory sample. Using inverted encoding modeling (IEM), we successfully decoded the location of the memory grating during the memory delay period for both task conditions. Cross decoding (i.e., training IEM on one task and applying it to the other task) between task conditions revealed evidence for both hypotheses. While crossing decoding on activity from posterior electrodes was successful consistently throughout the delay, it failed for activity from frontal electrodes quickly after early delay. Results suggest working memory employs both stable and task-specific representations to enable flexible information use.

56.433 THE INFLUENCE OF IDENTICAL OBJECTS ON VISUAL WORKING MEMORY CAPACITY: ELECTROPHYSIOLOGICAL EVIDENCE

Lijing Guo[1,2], Ruyi Liu[1], Dan Nie[1], Chaoxiong Ye[1,3]; [1]University of Jyvaskyla, Jyvaskyla, Finland, [2]Anyang Normal University, Anyang, China, [3]Sichuan Normal University, Chengdu, China

Identical memory items can potentially reduce the cognitive demands on visual working memory (VWM) and enhance its memory performance. While previous studies have preliminarily explored this topic, controversy remains, particularly regarding the generalizability of these benefits to complex stimuli. This study further investigates whether identical items within the memory range reduce the number of items maintained in VWM and explores the conditions under which this effect occurs. Participants performed a change detection task, memorizing the orientations of memory arrays under three conditions: (1) four identical orientations, (2) two pairs of identical orientations, and (3) four different orientations. By examining contralateral delay activity (CDA), an event-related potential component that reflects the real-time number of items stored in VWM, we observed that in the late time window, the CDA amplitude was significantly lower for the all-identical condition compared to the partial-identical and all-different conditions, with no significant difference between the latter two. However, there was no significant difference across the three conditions during the early time window. Our findings suggest that, during the early stages of VWM consolidation, individuals tend to encode as many items as possible from the visual field. However, in later consolidation stages, the VWM system processes identical information to reduce the number of maintained items. This effect, however, only applies when all items in the visual field are identical. When only some items are identical, directly storing all items is more efficient than first identifying identical ones and then processing them.

This work was supported by the Research Council of Finland (former Academy of Finland) Academy Research Fellow project (#355369 to Chaoxiong Ye) and Finnish Cultural Foundation (#00231373 to Chaoxiong Ye).

56.434 MEASURING TRUST IN ARTIFICIAL INTELLIGENCE WITH THE CONTRALATERAL DELAY ACTIVITY (CDA)

Tobias Feldmann-Wüstefeld[1], Eva Wiese[1]; [1]Technische Universität Berlin

Visual working memory is crucial for processing information in dynamic and challenging environments, making it a key factor in human-machine interaction. One increasingly significant form of such interaction is with artificial intelligence (AI). Offloading cognitive workload to AI has a great potential for enhancing human performance in complex tasks. However, a critical question in human-AI interaction is the extent to which humans trust the AI. We tested a novel approach to implicitly measure trust in AI with the contralateral delay activity (CDA), a key metric for working memory load. In solo-blocks, participants performed a change detection task by themselves. Participants had to encode items from one hemifield, indicated by a cue, and only that hemifield was probed in the end of a trial. In team-blocks, participants monitored one hemifield while a simple algorithm, framed as AI, monitored the other. If the human side was probed, participants responded immediately (change / no change). If the AI side was probed, the AI suggested selected a response (90% accuracy) and the participant confirmed or overruled the AI. In team-blocks, CDA was generally reduced compared to solo-blocks, showing that participants encoded items from both their and the AI's side, i.e., they did not offload as much memory load as they could, indicating some level of distrust. Importantly, those participants with the highest self-reported trust showed the smallest CDA amplitudes (the lowest offload) in team-blocks. In a second experiment, the AI's performance dropped after an initial phase of high accuracy. This led to a general decrease in CDA amplitude, reflecting trust dissolution. When the AI's performance later recovered, CDA amplitudes varied: some participants regained trust, while others did not. In sum our study shows that the encoding imbalance reflected in the CDA amplitude can be used as an implicit neural marker for trust in AI.

This research was supported by the Alexander von Humboldt Foundation

56.435 ANTICIPATION OF EXTRA MEMORY LOAD DOES NOT ALTER WORKING MEMORY'S RELIANCE ON LONG-TERM MEMORY

Yakup Yılmaz[1], Nursena Ataseven[2], Wouter Kruijne[2], Elkan Akyürek[2], Eren Günseli[1]; [1]Sabanci University, [2]University of Groningen

Working memory (WM) has a limited capacity, allowing storage of only three or four objects simultaneously. However, individuals often encounter more information than this limit allows. A potential strategy to handle this overload is transferring existing information to long-term memory (LTM) to free up WM capacity. In a study, we tested participants' strategies under the expectation of additional memory load. In Phase 1, participants (n=30) studied images of real-world objects, each paired with a distinct color. This would allow the recruitment of LTM for these objects. In Phase 2, their memory accuracy for these colors was assessed to ensure learning has taken place. The main experiment (Phase 3) utilized a blocked design with two anticipated load conditions (extra load expected and low load expected) and two memory conditions (studied and novel). In the extra load expected condition, 80% of the trials included two additional novel memory items, while 20% had no additional items. In the low load expected condition, these probabilities were reversed. This design allowed participants to anticipate the likelihood of additional memory load. After a retention interval, participants selected the color of one of the presented objects on a color wheel. To assess participants' strategies under load expectations, we measured contralateral delay activity (CDA) via EEG. The CDA was larger for novel vs studied items, demonstrating reliance on LTMs when available. However, this effect was unaffected by expectations of additional memory load. These findings suggest that participants prioritize relying on LTM whenever possible, and their memory strategies operate independently of expectations about future memory load.

56.436 CONTENT-INDEPENDENT POINTERS MEDIATE WORKING MEMORY STORAGE FOR BOTH VISUAL AND VERBAL STIMULI

Woohyeuk Chang[1] (woohyeukchang@uchicago.edu), Will Epstein[1], Will Ngiam[2], Henry Jones[1], Ed Awh[1]; [1]University of Chicago, [2]University of Adelaide

Visual working memory (VWM) and verbal working memory have often been treated as distinct processes. However, recent research suggests potential overlap between these two forms of memory. For instance, varying the number of letters and words elicits similar contralateral delay activity (CDA)—a load-sensitive electrophysiological signature of VWM that has typically been examined with visual stimuli (e.g., colored squares; Rajsic et al., 2019). In a re-analysis of these data, we applied multivariate load decoding and representational similarity analysis (RSA) to confirm the presence of a generalized load signal across words and colors, while also demonstrating that distinct variance in EEG activity tracked the stored content (words versus colors). To further test this finding, we replicated the study while eliminating perceptual differences by presenting colored words and asking observers to selectively store one of the features. Once again, we observed a generalized load signal across color and word features, even while other aspects of EEG activity tracked the attended feature. These results strengthen the case for a distinction between neural activity related to content-independent "pointers" and parallel neural signals that track the stored feature values.

Decision Making: Perception, memory

Tuesday, May 20, 2:45 – 6:45 pm, Pavilion

56.437 NOVELTY SEEKERS VS. FAMILIARITY SEEKERS - INDIVIDUAL DIFFERENCES AND PERSONAL PROFILING BASED ON MEMORY INFLUENCE ON VISUAL PREFERENCE

Shengjie Zheng[1], Eiko Shimojo[2], Shinsuke Shimojo[3]; [1]California Institute of Technology, [2]California Institute of Technology, [3]California Institute of Technology

Familiarity(F) and Novelty(N) preferences vary across object categories and individuals. While familiar faces are generally preferred over novel ones, the opposite is observed with natural scenes (Park et al., PNAS '10). We further investigated individual differences in F versus N preferences, towards a new kind of personality profiling based on memory-affected preference choice behavior. Fifteen adult participants (2 females, 13 males) performed a two-alternative forced choice (2-AFC) preference task. Participants were presented with photo stimuli from four categories: faces, natural scenes, cars, and geometric figures. In each trial, one image was familiar from previous trials, while the other was novel, but always within the same subcategory (e.g., the same race, gender, and age for faces). The "old" stimulus was chosen as the median rank-order attractive, from a pilot attractiveness rating result, individually. Participants rated their preference using a 7-point scale. They completed questionnaires, including Autism Spectrum Quotient (AQ), positive life attitude ("Maemuki"), and familiarity/novelty preferences in daily activities. Results indicated diverse F/N preferences across participants. Nine (out of total 15) participants replicated the trend from the previous study, preferring familiar faces but novel natural scenes. Only one participant showed consistent F preference, while another showed consistent N preference, across all the four object categories. Notably, there was a strong negative correlation between F/N preference and AQ scores (r=-0.66, P<0.01), indicating that individuals with higher autism spectrum tendencies strongly preferred novelty. A weak positive correlation was also found between "Maemuki" scores and F/N preference, indicating that individuals with a more positive and active lifestyle preferred novelty. This study introduces a new type of personality profiling, linking memory effects on visual preference with individual traits and lifestyle. It successfully combines visual psychophysics and clinical/personality psychology, offering insights into ASD and preference behavior in daily life.

The project was supported by JST/Moonshot Goal 9.

56.438 SEEING WHAT YOU WANT TO SEE: HOW BELIEF BIASES PERCEPTION AND MEMORY

Adam Malitek[1] (a.malitek@mail.utoronto.ca), Minsuk Chang[2], Cindy Xiong Bearfield[2], Keisuke Fukuda[1,3]; [1]University of Toronto Mississauga, [2]Georgia Institute of Technology, [3]University of Toronto

To behave intelligently in today's data-rich society, we need to accurately perceive, interpret, and remember the data presented to us. To this end, Xiong et al. (2022) showed that the perceived correlation of a scatterplot was biased toward participants' beliefs about the relationship between predictor and dependent variables. However, it is unclear whether the bias persists beyond the time of perception, and if so, whether the impact reflects a bias in the memory representation of the scatterplot or the interpretation of an unbiased memory of the scatterplot. To test this, participants first performed a correlation perception task. In each trial, participants saw a general statement depicting a relationship between two variables (e.g., A worker with a longer commute tends to be more stressed, X: commute time; Y: stress level) and reported their belief and predicted correlation between them. Participants then saw a scatterplot of the two variables and reported the perceived correlation. After the perception task, participants performed a memory task in which they first saw a pair of variables (e.g., X: commute time; Y: stress level) and indicated how well they remembered the corresponding scatterplot. Subsequently, they drew the remembered scatterplot from their memory and reported the correlation in a counterbalanced order. Our results (N = 70) replicated the belief-driven bias in perceived correlation. Furthermore, the belief-driven bias persisted in the scatterplots drawn from memory and the remembered correlations. These findings highlight the pervasiveness of prior beliefs in correlation estimation and suggest that belief-driven bias influences memory representation. This emphasizes the need for analysts to account for belief-driven biases in visual data processing and retention of data patterns in memory.

This research was supported by the Natural Sciences and Engineering Research Council (5009170).

56.439 ARBITRARY AND EXPLICIT PREDICTION BIASES PERCEPTUAL CATEGORIZATION

Olya Bulatova[1], Keisuke Fukuda[1,2]; [1]University of Toronto, [2]University of Toronto Mississauga

Predictions influence perception by biasing the percept in line with the expectation (Kok et al., 2012; 2018). However, this prediction-induced bias has primarily been observed by eliciting predictions through learned statistical regularities. Therefore, it is not clear whether arbitrary and explicit predictions would also bias our perception. To test this, we had participants perform a categorization task after making an arbitrary and explicit prediction on what they were about to see. More precisely, on each trial, participants first made an explicit binary prediction as to whether they would see one of the two objects (e.g., "Dog" or "Boar") by clicking either a top or bottom buttons (e.g., "Dog" and "Boar" buttons) displayed on the computer screen. Subsequently, a morph of the two objects was briefly presented (50ms), and participants categorized the object as one of the two objects by clicking either the left or right buttons (e.g., "Dog" and "Boar" buttons) displayed on the computer screen. To avoid motor priming, the response options for predictions and categorizations were orthogonalized and fully counter-balanced on a trial-by-trial basis. Here, we found that participants' categorization was indeed biased in the direction of their predictions. For instance, 50% morph objects (i.e., a morph of 50% Dog and 50% Boar) were more likely to be categorized as the predicted object than the counterpart, and we replicated this finding in two other object pairs (i.e., "Face and Tree" and "Gecko and Branch", Stöttinger et al., 2015). Our results extend the prediction-induced biases to arbitrary and explicit predictions, and the results of an ongoing EEG study examining the origin of prediction-induced biases (i.e., do predictions bias perception or categorization decisions?) will be discussed.

This research was supported by the Natural Sciences and Engineering Research Council (5009170).

56.440 CATEGORICAL REPRESENTATIONS IN SEQUENTIAL EVIDENCE ACCUMULATION

Mengting Fang[1], Alan Stocker[1]; [1]University of Pennsylvania

Perceptual decision-making often involves sequential evidence accumulation. Previous work has shown that category-level stimulus representations can play an important role in perceptual inference, even when not explicitly required. Here, we conducted a visual discrimination task to investigate how categorical representations can affect sequential evidence accumulation. Subjects discriminated the angular position (CW/CCW) of an unknown source relative to a reference based on 8 stimulus samples drawn from a Gaussian with fixed variance centered at the source position. Stimuli were presented in rapid sequence (150ms ISI). Subjects reported their categorical choice by pressing the corresponding button on a gamepad. After each trial, visual feedback displayed both the correct category and the source position. The reference was adjusted using a staircase procedure. All subjects performed the task under two conditions. In the first condition, they were asked to make a preliminary decision based on partial evidence within a 1.75s time-window, before then making their final choice after seeing all samples. The preliminary decision occurred either before the 1st sample (upfront guess) or after the 2nd, 4th, or 6th sample in the sequence. The four choice positions were randomly interleaved in each block. In the second condition (control), subjects were tested with the exact same sample sequences and reference positions as in the first condition, but simply maintained center fixation instead of making a preliminary decision. In contrast to the first condition, the reference was only shown for the final decision.

Both conditions were tested in alternating blocks. Our results show that being engaged in a preliminary decision against a reference significantly improves subjects' final decision performance compared to the control condition. This suggests that the formation of categorical stimulus representations may be crucial for accurate and robust sequential evidence accumulation.

This work was supported by the NSF CRCNS grant IIS-1912232 to A.A.S

56.441 CATEGORY-LEVEL STIMULUS REPRESENTATIONS INDUCE SIMILAR BIASES IN BOTH RETROSPECTIVE AND PROSPECTIVE PERCEPTUAL DECISION-TASKS

Marvin Maechler[1] (marvin.maechler@gmail.com), Alan Stocker[1]; [1]University of Pennsylvania

Decision-making often involves categorical choices that can bias subsequent judgments (e.g., confirmation bias). We investigated these choice-induced biases in a psychophysical task where participants first had to categorize the generative mean of a sequence of angular position stimuli (clockwise/counterclockwise relative to a reference) before ultimately estimating this generative mean by adjusting a cursor. In each trial we presented eight samples drawn from a normal distribution around different mean angular positions, displayed as small white dots (1dva diameter) on uniform gray background equidistant (5dva) from fixation. Importantly, a categorical decision was required after seeing either four samples (prospective condition) or all eight samples (retrospective condition). Unlike previous studies, detailed feedback was given in each trial to ensure that potential biases were not caused by miscalibrated inference strategies. To isolate choice-induced biases from effects of sampling noise, we presented identical stimulus sequences multiple times in both conditions. In line with previous results, we find robust estimation biases away from the reference position in both conditions. Interestingly, while estimates showed nearly identical distributions in both conditions, the correlation between categorical choice and subsequent estimate was strongly reduced in the prospective condition. This indicates that participants' categorical choices had a limited influence on their perceptual estimates in the prospective condition. Rather, we hypothesize that the mere presence of the reference induced a category-level stimulus representation that caused participants to make an implicit categorical judgment before their final estimate even in the prospective condition. Biases persisted despite monetary incentives for accuracy and trial-by-trial feedback. Therefore, these biases may arise from an intentional process that optimizes for multiple objectives beyond pure estimation accuracy. A holistic Bayesian matching model that formulates such multi-level objectives across the representational hierarchy can explain the data and demonstrates the normative nature of the observed biased stimulus estimates.

56.442 REPULSIVE BIAS IN CATEGORY FORMATION: BEYOND CONFIRMATION BIAS

Long Ni[1,2], Lucas Caceres[1], Michael S. Landy[1,2]; [1]Department of Psychology, New York University, [2]Center for Neural Science, New York University

Estimates of a stimulus feature (e.g., motion direction) after performing a categorization (e.g,. clockwise vs. counter-clockwise) are repelled from the category boundary (Jazayeri & Movshon, 2007). This phenomenon has been attributed to confirmation bias, where evidence consistent with the chosen category is overweighted (Stocker & Simoncelli, 2007). Is this true for sequential evidence accumulation for categorization? That is, in learning about two categories of stimuli from a sequence of examples, are estimates of the categories themselves biased away from the decision boundary and, if so, is this due to weighting evidence more heavily when it is consistent with the observer's categorical choice? Eight participants viewed a sequence of 40 ellipses in each block of trials. Ellipse orientation was sampled from one of two partially overlapping Gaussian orientation-category distributions with equal probability. Within a block, the two Gaussian distributions shared the same variance but had different means, equidistant from the optimal decision boundary. Blocks with 8 optimal decision boundaries (0 to 167.5 deg in 22.5 deg steps) were run in random order. Importantly, the decision boundary was never explicitly provided to participants and thus was learned while performing the task. Participants reported the stimulus category and received auditory feedback about correctness. At the end of the block, participants estimated the mean orientation of each category. Participants' estimates were substantially biased away from the decision boundary. Regression analysis showed that stimulus orientations that were correctly categorized contributed strongest to the mean estimates, consistent with confirmation bias. However, after controlling for the magnitude of orientations, we found that the repulsive bias was not due to confirmation bias but, rather, to overweighting stimulus orientations farther away from the optimal decision boundary. Our findings suggest that repulsive bias can occur in the absence of confirmation bias and thus highlight the need to consider alternative mechanisms.

NIH EY08266

56.443 SERIAL DEPENDENCE IN TARGET ESTIMATION BECOMES MORE PRONOUNCED WITH INCREASING TEMPORAL VARIATIONS IN THE CONTEXT

Jaeseob Lim[1] (jasup1883@gmail.com), Sang-Hun Lee[1]; [1]Seoul National University

Visual perception and decision-making are shaped by past experiences. One such instance is a phenomenon known as 'serial dependence,' in which current percepts assimilate to previous ones. Serial dependence has been construed as an adaptive strategy for leveraging the temporal stability found in one's surroundings. If that is the case, serial dependence is expected to manifest more strongly in a stable environment, characterized by stimuli that do not change much over time. To test this hypothesis, we asked participants to carry out an orientation estimation task, in which they viewed a white Gabor patch and reproduced its orientation after a delay. To regulate environmental stability over time, we asked participants to undertake an orientation change judgment task (e.g., clockwise or counterclockwise) for sequentially presented green Gabor patches in between the primary estimation task. We manipulated the amount of orientation change in the contextual task to affect how participants perceived the temporal stability of their environment. Our manipulation of temporal stability in the contextual task significantly modulated the degree of serial dependence in the estimation task. Surprisingly, however, the direction of modulation was opposite to our initial hypothesis: large orientation changes in the contextual task (i.e., a less stable environment) led to stronger serial dependence in the estimation task. This finding suggests that larger changes in surroundings might paradoxically stabilize perception of target over time. As one possible account of this seemingly paradoxical finding, we conjecture that the contrasting mechanism of the perceptual system in the spatial domain (e.g., tilt illusions) is also at work in the temporal domain. Substantial temporal changes in the context could downplay perceived temporal changes in the target, causing its current orientation to appear closer to past orientations. Our study highlights the complex interplay between contextual dynamics and serial dependence in visual perception.

This research was supported by the National Research Foundation of Korea funded by the Ministry of Science and ICT (RS-2024-00435727, RS-2024-00349515, RS-2023-00276729)

56.444 THE INFLUENCE OF EXPECTATIONS ON WORKING MEMORY-BASED DECISIONS AND SENSORY WORKING MEMORY REPRESENTATIONS

Ying Zhou[1], Kartik Sreenivasan[1,2], Daryl Fougnie[1]; [1]Division of Science and Mathematics, New York University Abu Dhabi, [2]Center for Brain and Health, New York University Abu Dhabi

Our memory of the world is influenced by what we expect. For example, working memory (WM)-guided decisions are biased towards highly probable stimuli (Jabar & Fougnie, 2021). However, it is unclear precisely how expectations bias WM-guided decisions: expectations may directly modulate WM representations themselves; alternatively, expectations may alter the process by which WM representations are translated to behavior. Perceptual studies provide evidence for both possibilities — some find that expectations bias perceptual representations (Kok et al., 2013), whereas others demonstrate that expectations influence decision-making instead of perception (i.e., cognitive impenetrability; Firestone & Scholl, 2016; Rungratsameetaweemana et al. 2018). To adjudicate between these alternatives, we had 15 participants perform a WM delay task for orientation. In the expectation condition, orientations were not random but instead drawn from a Gaussian distribution that was explicitly communicated to participants. Participants' reports indicated that WM-guided decisions were biased towards likely stimuli (bias = 6.62 ± 2.91°; $p < .001$). Moreover, modeling confirmed that this bias arose from the integration of memory and expectation, rather than a shift in response strategy. To investigate whether expectations were integrated with WM representations, we used a Bayesian decoder to identify the content of WM from fMRI activity measured during the delay period of each trial. We focused on WM representations in early visual areas, as these regions provide the most reliable decoding of orientation information. Despite seeing clear behavioral effects of expectation, there was no consistent evidence for bias in single-trial WM representations in early visual areas, suggesting that WM representations maintained in early visual areas may be impenetrable to expectation. How WM representations are converted into decisions is a critical, yet poorly understood topic in WM. The present work highlights that even simple WM-guided decisions may require integrating knowledge across multiple stages of processing.

This work was supported by the NYUAD Center for Brain and Health, funded by Tamkeen under NYUAD Research Institute grant CG012 to KKS

Scene Perception: Categorization, memory, clinical, intuitive physics, models

Tuesday, May 20, 2:45 – 6:45 pm, Pavilion

56.445 AVERSION TO AMBIGUITY: THE RELATIONSHIP BETWEEN CATEGORIZATION AMBIGUITY AND PLEASURE IN VIEWING REAL-WORLD IMAGES

Yikai Tang[1,2], William Cunningham[1,2], Dirk Bernhardt-Walther; [1]University of Toronto, [2]Vector Institute

Ambiguity, the state where competing perceptual interpretations or behavioral affordances coexist (Hirsh & Mar, 2011), has been shown to be aversive in various contexts. While extensively studied in decision-making, researchers suggest that aversion to ambiguity is a broader principle spanning perception and cognition. If this principle extends to visual preferences, individuals may favor images with clear, dominant interpretations and find ambiguous ones less appealing. Here, we focus on categorization as a fundamental form of interpretation. To test this hypothesis, we analyzed model-derived ambiguity estimates for natural object and scene images from the BOLD5000 dataset. For object images (N = 1,915), ambiguity was operationalized using classification models (e.g., VGG16, VGG19, AlexNet) by calculating the maximal confidence of their SoftMax outputs. Correlating these values with pre-collected pleasure ratings, we found that higher ambiguity was consistently associated with lower pleasure ratings across most models. We extended this analysis to scene images (N = 2,000), which often feature multiple objects with varying interpretability. Using the You Only Look Once (YOLO) model, we computed the average classification confidence across detected objects within each scene. Consistent with the findings for object images, higher average ambiguity in scene images correlated with lower pleasure ratings, suggesting that aversion to ambiguity influences preferences for more complex visual stimuli as well. Finally, we investigated whether the negative effect of ambiguity stems from its inherent aversiveness or the cognitive effort required to resolve it. Using BOLD signal changes in ventral visual areas (also from the BOLD5000 dataset) as a measure of processing costs, we found that these costs partially mediated ambiguity's negative effect on pleasure. This may suggest that aversion to ambiguity reflects both the aversiveness of ambiguity itself and the effort involved in resolving it. Overall, our study provides valuable insights into affective responses to categorization ambiguity in real-world images.

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.

56.446 VISUAL AND SEMANTIC SCENE INFORMATION: A STEADY-STATE VISUAL EVOKED POTENTIALS STUDY

Skylar Stadhard[1], Sage Aronson[1], Michelle Greene[1]; [1]Barnard College

It is understood that scenes can be processed rapidly. What is less understood is why some scenes are processed more quickly than others. Might there be information capacity limits to visual perception? We introduce methods for quantifying relative visual and semantic scene information, and assess the extent to which they contribute to the speed of scene recognition. We employed a 2x2 design to compare the temporal processing of images with high and low visual and semantic information. Images were selected from the Places in the Wild dataset, which contains over 67,000 RAW photographs spanning 260 scene categories. We collected descriptions from 10 participants for each image, and we used natural language processing (NLP) techniques to compute five features, using the first principal component to create a unified semantic information metric. Visual information was computed by assessing the relative file sizes of RAW and compressed PNG image versions. We reasoned that more compressible images had more redundancy, thus less visual information. Unexpectedly, we found that semantic information and visual information scores were largely uncorrelated (r=0.05). For our experiment, we used a sweeping steady-state evoked potential paradigm, with images flashing at 3 Hz. The phase coherence of each image increased over time, from 0% to 100%, in increments of 5% per second. We can assess the extent to which images were objectively recognized through examining the EEG power at 3 Hz over time. Further, by filtering data at 3 Hz, we can compute the mutual information between ERPs and images, using only stimulus-driven responses. By quantifying visual and semantic information, we have provided a method for testing two potential bottlenecks in the visual scene processing.

Supported by CAREER 2240815 to MRG

56.447 FACTORS AFFECTING BASELINE BOUNDARY EXTENSION EFFECT: A META-ANALYSIS

Jiří Lukavský[1,2], Natálie Šimčík[1,2], Filip Děchtěrenko[1]; [1]Czech Academy of Sciences, [2]Charles University

Boundary extension (BE) is the phenomenon where individuals remember a previously viewed scene as having a larger background extent than was originally presented. Over the past 30 years, numerous small-scale studies have explored this effect. This study synthesizes the existing literature to estimate the size of the BE effect and identify factors that influence its magnitude. We conducted a systematic search of the Web of Science and Scopus databases for studies citing the seminal work by Intraub & Richardson (1989). Inclusion criteria required studies to: (1) provide new experimental data, (2) focus on healthy adult participants, and (3) measure BE using a rating scale (e.g., -2 to +2). For each study, results were recorded for all close-up/wide-angle presentation combinations (CC, WW, CW, WC). When multiple experimental conditions were present, we selected the baseline condition, typically used for comparison against experimental manipulations. Data were analyzed using multi-level Bayesian meta-analysis, with results presented as credible intervals for the BE effect size and Bayes Factors for moderator analyses. The dataset included 110 outcomes extracted from 56 studies in 32 articles, encompassing data from 2,379 participants and 71,786 trials (mean: 30.2 trials per participant; 16.1 trials per outcome). On average, each study included 42.5 participants per condition (SD = 28.1). Our analysis revealed that the baseline condition produced a BE effect size of D = -0.222 (95% CI [-0.256, -0.189]). The effect was significantly influenced by presentation order of the close-up/wide-angle versions. However, there was strong evidence against the influence of other design choices (e.g., testing BE after each photograph or at the end, number of trials) or presentation factors (e.g., screen distance, viewing angle).

The research was supported by Czech Science Foundation (GF24-11506K).

56.448 THE EFFECT OF SCENE CLUTTER ON VISUAL REPRESENTATIONS

Stefania Bracci[1], Davide Cortinovis[2], Enrico Guarnuto[3]; [1]CIMeC, Trento University

In the field of object perception, the focus has traditionally been on explicit object dimensions such as object category, animacy or real-word size, often overlooking the complexity of the visual environment in which the objects are embedded. As an example, to control for possible confounds, most studies use images of objects without any background. However, our visual perception always deals with complex and extremely cluttered visual environments. This neuroimaging study explores how variations in the clutter of a scene influences object-related dimensions such as animacy and real-word size, typically represented in the ventral visual pathway. For this purpose, we created a set of images where each stimulus was presented either as a single object on a background (e.g., a butterfly) or as an object ensemble (e.g., many butterflies). In addition, animacy and real-world size were orthogonalized thus allowing to test the influence of scene ensembles on each feature space separately. Results revealed an interesting dissociation in regions encoding objects and scenes. Whereas, in object-selective areas, the animacy dimension was strongly represented in the single object condition, it did not reach significance in the object ensembles condition. On the contrary, in scene-selective areas, object size was encoded in the object ensemble condition but not in the single object condition. Together, this double dissociation suggests that the feature spaces encoded in the visual cortex are shaped by the interaction of both (1) the regional computational goal (e.g., scene processing) and (2) the visual properties of the images.

56.449 DISTINCT ROLES OF VISUAL AND SEMANTIC INFORMATION IN SCENE DETECTION AND CATEGORIZATION

Sage Aronson[1], Hooriya Aamir, Maria Adkins, Michelle Greene; [1]Barnard College, Columbia University

Scene processing is fast and effortless, yet we are often visually overwhelmed by information. Do we process such scenes more slowly? We adopt an information-theoretic framework, reasoning that visual and semantic information might limit the timecourse of visual processing. We compiled a dataset of 67,000+ RAW photographs spanning 260 different scene categories. To measure semantic information, participants provided descriptions of each photograph. We computed five metrics through natural language processing and applied principal component analysis to isolate the first component, which served as a measure of semantic information. Visual information was computed by compressing RAW images to PNG and calculating the difference in file size between the compressed and the original. This approach reasons that more compressible images contain less relative visual information. We examined both scene detection and categorization. In the detection series, participants were briefly presented with scenes or phase-randomized scenes (backwards masked), and were asked to determine whether the image was an intact scene. Images were selected from the 100 highest-information and 100 lowest-information images for both visual and semantic experiments. Results indicate that participants had higher sensitivity to images with less visual information (d' = 2.46) than images with more (d' = 2.03, p<0.001). By contrast, observers were more sensitive to images with high levels of semantic information (d'= 3.06) compared with low semantic information images (d' = 1.75, p<0.0001). In the categorization series, participants were presented with high- and low-information images and performed a 2AFC categorization task. Unlike the detection experiments, we observed no difference in categorization accuracy for high- and low-visual information. However, observers were more accurate with low semantic information images than high. Visual information limited scene detection but not categorization, while semantic information limited categorization while facilitating scene detection.

56.450 UNTYPICAL SCENE EXEMPLARS ARE EASIER TO REMEMBER, BUT HARDER TO CATEGORIZE

Charlotte Atzert[1] (charlotte.atzert@uni-muenster.de), Filip Děchtěrenko[2], Jiří Lukavský[2], Niko A. Busch[1]; [1]University of Münster, Germany, [2]Czech Academy of Sciences, Prague, Czech Republic

Why are some events remembered for a long time, while others are quickly forgotten? Specifically, which image properties make certain scene images more consistently memorable than others? While the memorability of any event can be influenced by many different factors, it has long been recognized that unusual events are remembered better than typical events. Here, we investigated whether image memorability is linked to image typicality. Specifically, we tested the hypothesis that untypical images would be harder to categorize but easier to remember. We collected a dataset of scene images from several indoor scene categories with ratings of subjective typicality. For each scene category, we also compiled two sets of feature dimensions: descriptors used by human participants to discriminate between scene exemplars (e.g., number of windows, coziness) and deep neural network features based on AlexNet layer fc7. Both sets of features were strongly predictive of scene typicality ratings, indicating that scene typicality reflects the configuration of image features of a given exemplar relative to other category exemplars. We then presented typical and untypical scene exemplars to a new sample of 90 participants. In the categorization task, participants decided as quickly as possible whether a given image was an exemplar of a specific category. In the subsequent memory task, we repeated all images from the categorization task along with new foils, and participants discriminated between old and new images. We found that typical images were categorized as exemplars of their category faster and more accurately than untypical images. Conversely, typical images were more difficult to remember than untypical images. Together, these findings highlight a trade-off between scene memorability and categorizability, suggesting that both depend on how distinct a scene's features are compared to other category exemplars.

This research was supported by the German Research Foundation (BU2400/14-1) and the Czech Science Foundation (GF24-11506K).

56.451 ENSEMBLE PHYSICS: PERCEIVING THE MASS OF GROUPS OF OBJECTS IS MORE THAN THE SUM OF ITS PARTS

Vicente Vivanco Cepeda[1], Joshua B. Tenenbaum[1], Kevin A. Smith[1], Vivian C. Paulun[1]; [1]Massachusetts Institute of Technology (MIT)

Imagine pouring a box of granola into a bowl. Are you considering hundreds of individual chunks or the motion of the group as a whole? Human perceptual limits suggest we cannot be representing the individuals, implying we simulate 'ensembles' of objects. If true, we would need to represent group physical properties beyond individual aggregates, similar to perceiving ensemble properties like color, size, or emotion. However, unlike previously discovered ensemble perception, physical properties such as mass are not directly observable. In two experiments we study whether people can perceive the mass of an ensemble, and if this representation is more than an aggregate of the properties of individual objects. Participants viewed video clips of marbles falling onto a cloth. On each trial, they had to choose which of two consecutive videos showed the larger mass. In Experiment 1, 40 participants completed two conditions: 1) comparing the masses of two individual marbles, and 2) comparing the average mass of groups of 25 marbles with different means. For the same mass ratios, we found people discriminate ensembles better than individuals ($\chi 2(9)=37$, p<.001). Does this imply ensemble processing beyond the sum of parts? In Experiment 2, 21 participants compared the average mass of a group of 25 marbles with different means against a similar set with five changed marbles. Crucially, immediately after they were asked to identify one of the marbles that had changed weight. Participants' mass judgments were significantly better (66% correct) than can be expected if they were relying on individual information (58% correct given 33% correct identification of one marble, Haberman & Whitney (2012); binomial test, p<.001). Together this supports the concept of ensemble perception in intuitive physics, extending our understanding of how people represent and simulate sets of objects.

KAS and JBT were supported by National Science Foundation Science Technology Center Award CCF-1231216, and NSF grant 2121009

56.452 INTUITIVE PHYSICAL AND DOMAIN-GENERAL REASONING ARE DISSOCIABLE IN THE HUMAN BRAIN

RT Pramod[1], Samuel Hutchinson[1], Nancy Kanwisher; [1]MIT

Recent studies implicate regions of the human bilateral fronto-parietal cortices in intuitive physical reasoning—our ability to perceive, plan in, and predict the physical world. However, it remains unclear whether this "Physics Network" (PN) constitutes a distinct and domain-specific system, or whether instead it overlaps with the domain-general multiple demand (MD) network. To answer this question, we scanned subjects (N = 20) with fMRI in a pre-registered study and measured spatial overlap and resting-state correlation between the PN and MD networks. In each participant, the Physics Network was identified with our standard localizer contrasting a physical reasoning task on block tower stimuli ("Where will the block tower fall?") with a color-based task ("Are there more blue or yellow blocks in the tower?"). The MD network was localized using a standard hard > easy spatial working memory task. We report three findings: (1) The overlap between the two networks was minimal, with an average Dice coefficient of 0.08. (2) Network-specific voxels exhibited consistent selectivity in held-out data. Voxels in the PN, identified using one half of the data, exhibited stronger physics > color task contrast in held-out data compared to the hard > easy spatial working memory contrast, and vice versa for the MD network ($p < 0.0005$ for network x localizer interaction in an ANOVA). (3) Resting-state functional correlation revealed that subregions within each network were more strongly correlated with each other than with subregions of the other network (average correlation: $r = 0.49$ within the PN, $r = 0.51$ within the MD network, and $r = 0.29$ between physics and MD networks). These findings indicate that the Physics Network is dissociable from the Multiple Demand system in both spatial location and resting functional time course, supporting its distinct role in intuitive physical reasoning.

56.453 A LARGE-SCALE VISION-LANGUAGE FMRI DATASET FOR MULTI-MODAL SEMANTIC PROCESSING

Yuanning Li[1] (yuanningli@gmail.com), Shurui Li[1], Zheyu Jin[1], Shi Gu[2], Ru-Yuan Zhang[3]; [1]ShanghaiTech University, Shanghai, China, [2]University of Electronic Science and Technology of China, Chengdu, China, [3]Shanghai Jiao Tong University, Shanghai, China

Large-scale functional MRI datasets with naturalistic stimuli provide more ecologically relevant experimental conditions and promote more reproducible research into the neural basis of sensory perception. These datasets afford the use of advanced AI models to investigate and model the neural coding and processing of language and visual information. However, existing research often focuses on isolated visual or language networks, with few studies addressing the interaction between vision and language processing at the semantic level. To facilitate the investigation of the neural mechanisms of semantic representations across different modalities, we present the Caption Scene Dataset (CSD). Specifically, we designed a paired caption-image semantic matching task and collected extensive 3T fMRI data from 8 subjects for a total of 320 hours. During the scanning, each subject viewed more than 4400 paired visual stimuli, each stimulus contained a text caption, followed by a naturalistic image. The subjects were asked to determine whether the paired text and image stimuli matched semantically. In addition to the fMRI scans, we acquired T1, T2, and diffusion MRI data, as well as eye-tracking and electrocardiogram (ECG) data to enrich the dataset's utility. We also localized the early visual cortex and category-selective areas in each participant using additional functional localizers. Preprocessing steps included slice-time correction, EPI distortion correction, and motion correction, followed by alignment to individual brain space. Single-trial neural responses to text and image stimuli were estimated using GLMSingle. To illustrate the utility of the CSD dataset, we demonstrated that deep neural encoding models could effectively predict neural responses to text and image stimuli across different cortical regions. In sum, our unique, large-scale vision-language fMRI dataset establishes a robust platform for investigating the neural basis of semantic processing across vision and language modalities, fostering cross-disciplinary advances at the intersection of cognitive neuroscience and AI.

This work is supported by the National Natural Science Foundation of China (32371154, Y.L.), Shanghai Rising-Star Program (24QA2705500, Y.L.), Shanghai Pujiang Program (22PJ1410500, Y.L.), and the Lingang Laboratory (LG-GG-202402-06, Y.L.).

56.454 SCENE UNDERSTANDING MAPS: PREDICTING MOST FREQUENTLY FIXATED OBJECT DURING FREE VIEWING WITH MULTI-MODAL LARGE LANGUAGE MODELS

Shravan Murlidaran[1], Miguel P Eckstein; [1]Department of Psychological and Brain Sciences, University of California Santa Barbara

Introduction: What objects are most frequently fixated during free viewing of scenes with images depicting complex human behaviors, actions, or interactions? We hypothesized that: 1) objects that contribute the most to the understanding of the scene would be fixated most frequently, 2) Multi-Modal Large Language Models (MLLMs) can be used to build Scene Understanding Maps (SUMs) that visualize the contribution of individual objects to accurate scene descriptions and can predict the most fixated objects by humans during free viewing. Methods: For each image (n=38), we created an MLLM-SUM by digitally removing each object from the image and using a semantic similarity measure (Gemini) of the MLLM description of the manipulated images and that of the intact images. We compared the predictions from MLLM-SUM to the impact of the same object deletion manipulation on human scene descriptions and to human fixations during a separate free-viewing task (2-sec presentation). We benchmarked the MLLM-SUM against Graph-Based Visual Saliency (GBVS), DeepGaze models, and human judgments of the meaningfulness of cropped scene patches (meaning maps). Results: The MLLM-SUM's object most critical to scene understanding (minimum of the MLLM-SUM) accurately predicted the object most disruptive to human (N=55) scene descriptions in 73.5 +- 7% of the images (chance performance = 12.5%). The MLLM-SUM also predicted the most fixated object by humans (N=50) in the free viewing task in 52.6 +- 8% of the images, significantly higher than the maximum of the GBVS (29 +- 8%, bootstrap, p=0.009) and meaning maps (24 +- 7%, bootstrap, p=0.002) but not significantly different from the maximum of Deepgaze (47.3 +- 8%, bootstrap, p=0.26). Conclusions: Multi-modal large language models and semantic similarity metrics provide a new powerful tool to identify objects critical to scene understanding and predict the most fixated object during free viewing.

56.455 FOVEATED MULTI-MODAL LARGE LANGUAGE MODEL MAPS TO PREDICT TIME TO UNDERSTAND SCENES

Ziqi Wen[1] (ziqiwen@ucsb.edu), Shravan Murlidaran[2], Jonathan Skaza[3], Miguel Eckstein[1,2,3]; [1]Department of Computer Science, UC Santa Barbara, [2]Department of Psychological and Brain Sciences, UC Santa Barbara, [3]Graduate Program in Dynamical Neuroscience, UC Santa Barbara

Introduction: Perceptual science has proposed evidence accumulation models (Ratcliff et al., 2016; Rafiei & Rahnev, 2022) to predict response times for simple perceptual tasks and clutter metrics to predict search times with complex scenes (Rosenholtz, 2005; Deza & Eckstein, 2017). There are no image-computable models that predict an observer's time to understand real-world scenes. Here, we use Multi-Modal Large Language Models (MLLMs) to generate foveated scene understanding maps of how the description (understanding) of a scene is influenced across fixation locations and propose new metrics to predict observers' time to understand a scene. Methods: For each image (n=94), we created a foveated scene understanding map (FSUM) that quantifies how an MLLM scene description for different fixation points quantitatively compares (measured using LLM sentence similarity measures, Gemini) to a gold standard description (MLLM description of non-foveated images). An FSUM metric was created by integrating the SUM map using Minkowski pooling. We assessed the ability of the FSUM metric to predict the response time and number of saccades required to accurately describe a scene by seven observers. For benchmark comparisons, we used image clutter metrics (Feature Congestion, Rosenholt, 2017 & Subband Entropy). Results: To understand the scenes, observers executed 6.9±0.18 saccades and took 1.89±0.06 sec. The FSUM metric correlation with human RTs and number of eye movements was significantly higher ($p<0.001$) than clutter metrics (RTs: FSUM=0.547; Feature Congestion=0.1; Subband Entropy=0.099; Eye movements: FSUM=0.605, Feature Congestion=0.119, Subband Entropy=0.038). As a comparison, the correlation of single observer RTs and the average of the remaining observers RTs was 0.391. Conclusion: Combining foveated architectures, multi-modal large language models, and semantic similarity metrics provides a new powerful tool to extend vision science approaches to human understanding of complex scenes.

56.456 SEEING FROM THE GROUND UP: SPONTANEOUS PERCEPTION OF 'CAUSAL HISTORY' DUE TO INTUITIVE PHYSICS

Kimberly W. Wong[1], Aalap D. Shah[1], Brian Scholl[1]; [1]Yale University

We typically think of visual perception as providing us with representations of our present local environments. But vision may also sometimes represent the causal *past*, extracting how those environments got to be that way--as when a shape with a jagged 'bite' is represented as the full (unbitten) shape to which an event (biting) occurred. Here we suggest that this perception of 'causal history' is more prevalent than previously suspected, due to intuitive physics. In a stack of blocks (or books, or dishes), for example, gravity entails that the bottom object was placed before higher objects. Here we show that such 'historical' relationships are spontaneously extracted during passive viewing, and influence perception in surprising ways. Observers viewed a table on which a stack of two blocks appeared, (1) all at once, (2) with the bottom block appearing first, or (3) with the top appearing first--and they simply reported on each trial whether the blocks appeared simultaneously or sequentially. We reasoned that possibility (3) might be less naturally perceived, since it violates the causal history mandated by the underlying intuitive physics. And indeed: when the bottom block appeared first, observers reliably perceived this sequential presentation; but when the top block appeared first, they were more likely to mistakenly perceive that blocks had appeared simultaneously--as if the actual temporal offset and the gravity-inspired prior effectively cancelled out. In fact, observers were more accurate for towers built 'from the ground-up' than for actual simultaneous presentations (which were often misperceived as sequential). And these effects seemed specific to gravity-based intuitive physics, since they disappeared when the same stimuli appeared to be lying flat on the table. These results collectively show how visual processing extracts causal history as a result of intuitive physics, and how such representations influence the perception of temporal order.

56.457 MODELING VISUAL CORTEX WITH LOCAL UNSUPERVISED LEARNING

Ananya Passi[1], Michael F Bonner[1]; [1]Johns Hopkins University

Deep neural networks (DNNs) are leading computational models of the visual cortex. However, their reliance on biologically implausible training methods, such as task supervision and backpropagation, limits their interpretability and alignment with biological principles. Additionally, the time- and resource-intensive nature of their training discourages iterative exploration, further hindering insights into high-performing models of the visual cortex. In contrast, biological visual representations are believed to emerge through unsupervised, local learning mechanisms that do not rely on backpropagation, highlighting the need for alternative computational approaches that better reflect these processes. Here, we propose a biologically inspired framework for learning hierarchical visual representations using local unsupervised learning without backpropagation. In this approach, each layer of a DNN incorporates a bottleneck mechanism that compresses and subsequently expands representations. The learning process is entirely unsupervised, with each layer optimizing only to compress its inputs. This minimalist algorithm produces representations that rival models trained using backpropagation in their ability to predict image-evoked fMRI responses in the visual cortex up to intermediate processing stages. By aligning with key principles of biological learning, our approach requires no labeled data or task-specific supervision and provides a parsimonious framework for modeling visual hierarchy formation. Our model offers neuroscientists a novel tool for conducting in-silico analyses and controlled rearing experiments with reduced computational overhead. Furthermore, its simplicity and biological plausibility provide new insights into how visual computations might be organized in the brain, advancing our understanding of the neural mechanisms underlying vision.

Scene Perception: Spatiotemporal factors

Tuesday, May 20, 2:45 – 6:45 pm, Pavilion

56.458 BIASES IN PREDICTIONS OF DYNAMIC NATURAL SCENES: CONTRIBUTIONS OF MOTION AND SCENE CONTENT ON THE ACCURACY AND PRECISION OF PREDICTION

Jiali Song[1] (jiali.song@utoronto.ca), Ginnie Wee[1], Benjamin Wolfe[1]; [1]University of Toronto Mississauga

Prediction is a fundamental part of navigating our visual world. Although there is prior evidence of prediction in memory representations of dynamic natural scenes (representational momentum), there is relatively little empirical data on an explicit prediction task in this context. Our study investigates the extent to which drivers can accurately predict non-hazardous, everyday road scenes. To this end, we created a novel stimulus set of 3D videos of real road scenes, which we plan to make publicly available, recorded using a stereoscopic dashcam setup during urban and highway driving. On each trial, we showed observers a 2s preview (video or still image) of a road scene and asked them to select the image that best represents what they think the scene will look like 2s after the end of the preview in a 5AFC task. The alternatives were frames sampled from the video at 1s intervals and always included the correct (+2s) frame. We also manipulated the presence of stereoscopic depth information using a 3D display. In a sample of 48 licensed drivers who each performed 420 trials, we found that predictions were on average 0.29s farther in time than ground truth, and such bias towards the future was larger for video compared to still image previews. Prediction proportion correct was higher for video compared to still previews and for urban roads compared to highways, suggesting an important role for motion information and environmental density in prediction. Moreover, these effects were mainly driven by increased prediction precision with relatively small changes in the magnitude of future bias. Stereoscopic depth information had negligible effects on prediction performance. Our findings suggest that drivers can make predictions about road scenes, and these predictions are subject to biases similar to those affecting memory representations.

This work was supported by funding from NSERC Discovery Grant and University of Toronto XSeed Grant awarded to BW.

56.459 DYNAMIC COLOR PRIMING IN MLDS ALTERS PERCEIVED SIMILARITY

Haley Frey[1], David Whitney[1]; [1]University of California, Berkeley

Previous work has demonstrated that changes in perception can be measured by comparing maximum likelihood difference scales (MLDS). A preceding stimulus can alter the appearance of the current stimulus, if the stimuli are similar, leading to a greater perceived similarity of those stimuli relative to others along a continuum. Here, we expand upon these results by using dynamic stimuli to prime perception and compare difference scales to a baseline (no prime) curve. The stimuli are video clips that slowly morph through a subset of colors from the continuum. We anticipated that the dynamic prime would show a similar serial dependence effect as the static prime did, reflected by greater similarity ratings of the stimuli involved in the preceding morph compared to stimuli outside of this range. Indeed, we find evidence of such changes, particularly in the center range of the continuum values. Our results demonstrate a “reverse crispening” that occurs in the MLDS curve in response to dynamic primes in this center range. This “reverse crispening” is interesting because it reflects a pattern in similarity perception that opposes the typical psychometric function. One possibility is that the perceptual effect of primes from the central region of the continuum have a greater impact on the scale. Stimuli that are already perceived as similar (usually at either end of the continuum) may not show an obvious increase in perceived similarity following a prime. However, an increase in perceived similarity among stimuli at the center of the continuum, which observers usually rate as more different, would drastically alter the resulting difference scale, leading to the seeming reverse crispening we observe here.

Institutional Training Grant T32EY007043 awarded to the Department of Vision Science, School of Optometry, University of California, Berkeley

56.460 EFFECTS OF IMAGINING PERFORMING A TASK ON THE MEMORY FOR SCENE PICTURE BOUNDARIES

Carmela Gottesman[1] (cvgottesman@sc.edu), Ashley Schirmer[1], Kristen Tehonica[1]; [1]University of South Carolina Salkehatchie

When testing scene memory, participants typically simply look at the scene or they may be instructed to study it during the presentation stage. However, most of our interactions with the scene around us in everyday situations involves doing different tasks with objects, like picking up litter, making a sandwich, etc. This study examines how thinking of, and imagining performing tasks affect the spatial mental representation of scenes. Boundary extension refers to the memory error in which people report remembering areas of the scene that were actually beyond the edges of the picture they saw. We examine if imagining doing a task will affect memory for the picture boundaries. Each picture presented in this study, was of a scene containing objects related to two tasks, one more to the left and one more to the right in the picture. Participants were asked to examine the pictures and imagine themselves doing one of the two tasks, counterbalanced across participants. Afterwards, the viewers were shown four versions of each of the scenes. Each version showed segments of the scene with different boundaries: the original image, an image where boundaries are extended in all directions, an image where the boundaries are extended in the direction more relevant to the task they were imagining, and an image where the boundaries are extended in the opposite direction. The viewers were asked to select which version best fits the picture they remember seeing before. Viewers were less likely to choose the version with overall extended boundaries, but they were equally likely to choose all other options. Notably, the choice with boundaries extended on the side of the performed task was not significantly differentiated from the one with extended boundaries on the other side. This suggests that imagining performing a task may reduce but not eliminated boundary extension.

56.461 HOW MUCH VISUAL FIELD LOSS CAN YOU TOLERATE ON THE ROAD? IMPACT OF CENTRAL AND PERIPHERAL SCOTOMAS ON ROAD HAZARD LOCALIZATION

Ido Zivli[1] (bazido@gmail.com), Ginnie Wee[1], Jiali Song[1], Benjamin Wolfe[1]; [1]University of Toronto Mississauga

Our ability to perceive the gist of a scene in a glance is well-established for static (Greene & Oliva, 2009) and dynamic scenes (Wolfe et al. 2019). Strong views of the Useful Field or Functional Visual Field argue that dynamic scene tasks require central vision whereas peripheral vision is less useful. However, little empirical evidence speaks to this claim. Here, we examined the extent to which observers could tolerate foveal and peripheral visual field loss in a road hazard localization task. Eight licensed drivers viewed 2s excerpts of 270 dashcam videos from Road Hazard Stimuli (Song et al. 2024; videos subtended 39 x 22 DVA). 66% of videos contained hazards, defined as situations which require immediate driver response to avoid a collision. Hazards occurred in the left or right of frame in equal proportion. The remaining 33% of videos contained no hazards. Participants indicated whether a hazard was on the left, on the right or absent. Using a gaze-contingent display, we simulated central and peripheral scotomas in two separate blocks. Central scotomas were simulated by removing a circular window at the gaze location, leaving the rest of the video visible. Peripheral scotomas only displayed the video inside a circular window at the gaze location. To determine the threshold scotoma size at 80% localization performance, the diameter of the circular window was controlled with a trialwise 3-up-1-down staircase. In the central scotoma condition, participants tolerated a 6.80 DVA central scotoma. In the peripheral scotoma condition, participants required an 8.34 DVA visible window. Our results suggest that detecting immediate hazards in road scenes is highly resistant to visual field loss. Moreover, central and peripheral vision both support this task, suggesting a more complex account of peripheral vision use in dynamic scene perception.

Supported by a NSERC DG, University of Toronto XSeed awards to BW

56.462 IRRESISTIBLY LOGICAL: DISJUNCTIVE INFERENCES FACILITATE VISUAL RECOGNITION OF LIKELY AND UNLIKELY EVENTS

Nathaniel Braswell[1], Chaz Firestone[2], Nicolò Cesana-Arlotti[1]; [1]Yale University, [2]Johns Hopkins University

Whereas logical inference is typically associated with symbolic notation and laborious proofs, it also arises intuitively in everyday reasoning. Previous work shows that human infants deploy basic disjunctive inferences to infer occluded objects' identities (Cesana-Arlotti et al., 2018). In two prior studies (Braswell et al., VSS 2023), we discovered that this developmentally basic logical computation arises spontaneously when adults recognize objects in visual scenes. In particular, we showed adults visual events wherein objects are hidden and then revealed in ways that either follow or violate the logically predicted outcomes; subjects responded faster when a revealed object's identity was consistent with the inference's prediction than when it violated it. Here, we explore whether such inferences are automatic and even "irresistible", arising in circumstances where the subject receives statistical evidence contrary to the inference. In Experiment 1, participants had to identify two kinds of concurrent objects: ones logically predicted by the events in the scene and logically unrelated ones. Strikingly, participants recognized objects predicted by the logical inference faster than identically looking objects that were logically unrelated to the scene, suggesting that logical inferences were facilitating and expediting visual processing. In Experiment 2, we manipulated the statistical distribution of revealed objects to create cases where the logically predicted outcome was statistically unlikely (by a ratio of 2:1). Remarkably, participants often misidentified the statistically likely object as the improbable one merely because logic compelled them to do so, despite their predictions being contradicted by previous statistical evidence. In other words, even when it would have benefited participants not to reason logically, they couldn't help but do so. This work shows how methods from vision science can illuminate the mind's logical capacities. Our findings suggest the presence of core logical inferences that automatically facilitate visual processing and are hard to resist despite prevailing counterevidence.

NSF BCS #2021053 awarded to C.F.

56.463 RELATIONAL SCENE UNDERSTANDING RELIES ON CORTICAL PROCESSING EVIDENCED BY DICHOPTIC PRESENTATION

Nayoung Kim[1], Oakyoon Cha[1]; [1]Sungshin Women's University

Relational representations can be useful in understanding a visual scene. For example, recognizing an object that provides physical support (e.g., table) to another object (e.g., book on the table) may facilitate scene understanding. Hafri and colleagues (2024) demonstrated that participants recognized a scene faster when the reference object (table in the example) was presented before the figure object (book in the example) than the reverse order of presentation. Even for species lacking a neocortex, relational information can guide their behavior in response to environmental challenges. In the present experiment, we investigated the possibility that subcortical areas contribute to the building of relational representations. To address this possibility, we used a dichoptic presentation technique. When two objects are presented to different eyes, the two objects are processed separately in the subcortical visual processing areas before information from the two eyes integrates at the primary visual cortex. If subcortical areas contribute to relational representations, dichoptic presentation should disrupt these processes. Participants first read a sentence and then viewed scenes depicting a reference and a figure object, judging whether the sentence accurately described the scene. Reference and figure objects could be presented to both eyes (dioptic condition), to the same, one eye (monoptic condition), or to different eyes (dichoptic condition). Object presentation order (reference first or figure first) was also manipulated. Across all conditions, participants responded faster when the reference object preceded the figure object , replicating previous findings. The effects of presentation order of objects were comparable across dioptic, monoptic, and dichoptic conditions (BF $<$ 1/3). These results suggest that relational representations rely predominantly on cortical processing rather than subcortical pathways, at least when it is judged based on the relational representations derived from language-based descriptions.

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).

56.464 REPRESENTATIONS OF SCENE BEAUTY IN SPACE AND TIME: AN EEG-FMRI FUSION STUDY

Philipp Flieger[1], Rico Stecher[1], Daniel Kaiser[1,2]; [1]Department of Mathematics and Computer Science, Physics, Geography, Justus Liebig University Gießen, Germany, [2]Center for Mind, Brain, and Behavior, Justus Liebig University Gießen, Philipps University Marburg and Technical University Darmstadt, Germany

How does neural activity give rise to the perception of beauty in our environment? Previous EEG work suggests neural representations of beauty emerge during early perceptual processing and are temporally sustained. Complementary fMRI work pinpoints the neural correlates of beauty to visual, frontal, and default-mode network (DMN) areas.

An integrated view of the spatiotemporal dynamics that give rise to the perception of beauty, however, is lacking. In separate EEG (N = 52) and fMRI (N = 29) studies, participants rated the beauty of 100 scene photographs. We then performed a model-based EEG-fMRI fusion to investigate the spatiotemporal dynamics of beauty-related representations. Specifically, we obtained the representational similarities between EEG responses at every time point and representational similarities between fMRI responses for 100 brain parcels. We then determined how much of the correspondence between EEG and fMRI-derived similarities is explained by beauty ratings. Our results suggest that representations underlying beauty judgments emerge early (peaking at 275ms post-onset), are long-lasting, and span visual, frontal, and DMN areas. This spatiotemporal signature remained robust when controlling for image-quality ratings (obtained separately, N = 43). When additionally controlling for visual features extracted by a VGG16 deepnet model, we found strongly reduced beauty-related activations, particularly in visual cortex. Conversely, beauty-related activations in frontal and DMN regions remained relatively unaffected. These representations still emerged remarkably fast (within 225ms). We conclude that the perception of beauty involves rapid and spatiotemporally sustained processing related to visual correlates of beauty, while the DMN concurrently forms more abstract representations of beauty.

56.465 SCENE PROCESSING IS INFORMATION CAPACITY LIMITED BY BOTH VISUAL AND SEMANTIC INFORMATION

Amy Nguyen[1], Vivian Gao, Sage Aronson, Michelle Greene;
[1]Barnard College

Scene understanding is considered rapid and effortless, yet its processing mechanisms are currently unknown. Scenes differ in content, and it may be the case that some scenes are understood more efficiently than others. However, no framework exists to predict why some images may be processed more efficiently. We adopt an information theoretic approach, measuring the relative amounts of visual and semantic information in photographs and assessing the processing efficiency via time-resolved EEG decoding. We amassed a novel dataset of over 67,000 RAW photographs from over 260 scene categories. RAW images were compressed to PNG, and visual information scores were computed as the difference in resulting file sizes. We reasoned that more compressible images had less relative visual information. To assess semantic information, participants provided descriptions of each scene. We calculated five metrics from natural language processing and, using principal component analysis, extracted the first component as a semantic information measure. Unexpectedly, visual and semantic information scores were not strongly correlated (r=0.05). For our ERP experiments, we selected the 20 images with the highest and lowest scores in visual and semantic information. Each of the 40 images was presented for 500 ms and repeated 30 times while participants performed an orthogonal task to maintain attention. We used time-resolved, whole-brain decoding with a linear support vector machine whereby image identity was predicted with five-fold cross-validation. For both experiments, decoding accuracy was higher for low information images. Moreover, the latency of maximum decoding accuracy was earlier for low information images. These results indicate that scene processing is delayed for high-information images, and this is true for both visual and semantic information. This suggests that the visual system has information capacity limitations for both visual and semantic information.

Supported by CAREER 2240815 to MRG

Scene Perception: Natural images, virtual environments

TUESDAY, MAY 20, 2:45 – 6:45 PM, PAVILION

56.466 DEPTH ESTIMATION IN REAL-WORLD SCENES

Michaela Trescakova[1] (mt1n16@soton.ac.uk), Wendy Adams[1], Matthew Anderson[2], James Elder[3], Erich Graf[1]; [1]School of Psychology, University of Southampton, [2]School of Optometry and Vision Science, University of California, Berkeley, [3]Centre for AI & Society, York University, Toronto

Human depth perception is often studied using simple shapes defined by a limited number of depth cues. In this study, we used the Southampton-York Natural Scenes (SYNS) dataset to examine the dynamics of depth estimation and the relative contributions of elevation, binocular disparity, and color on both ordinal and ratio depth judgments. Participants viewed briefly presented images (17–267 ms) from 19 outdoor scene categories in the SYNS dataset. In Experiment 1, images were viewed under monocular conditions, while Experiment 2 included both monocular and binocular conditions. Each image featured two crosshairs marking a pair of locations, and participants identified which location appeared farther away. They then used a slider to report the depth of the nearer location as a percentage of the depth to the farther location. Trials varied in the depth difference and mean depth of the target locations. Experiments 3 and 4 introduced manipulations to the color of the images (natural, color-inverted, and greyscale), transformed using the CIELuv color space. Participants displayed a strong elevation prior across the experiments - they consistently judged the higher crosshair as further. Participants demonstrated the ability to estimate local depth with extremely brief image presentations, with both elevation and binocular disparity influencing early depth estimates. Color effects were marginal relative to the impact of other cues, suggesting that color information does not significantly enhance early perception of relative depth for outdoor scenes. We found that deep networks trained on other datasets, and K-nearest neighbours regression trained on SYNS image features, were less accurate than humans. This performance gap was most pronounced when elevation cues were absent, and at longer presentation durations. Altogether, these results suggest that human depth estimation relies on cues not fully captured by local image features or current deep learning models.

The work was funded by an EPSRC grant (ROSSINI, PI: Adams), EP/S016368/1

56.467 DIMENSIONS OF TUNING FOR 3D SCENE STRUCTURE IN THE BRAIN REFLECT DIMENSIONS OF 3D STRUCTURE ENCOUNTERED DURING NATURAL BEHAVIOR.

Mark Lescroart[1]; [1]University of Nevada, Reno

Many brain regions respond selectively to particular categories of images, but why and how these response properties arise remains unclear. Here, we investigate whether selectivity for scenes (as measured by BOLD fMRI) is related to dimensions of scene structure in natural visual experience. To address this question, we used data from the Natural Scenes Dataset (NSD), which contains 7T BOLD fMRI responses to thousands of natural images, and the Visual Experience Dataset, which contains hundreds of hours of first-person video of natural behaviors. We developed an image-computable model that formalizes 3D scene structure as the distribution of surface normals and distances in each image or video frame. We used this model to compute 3D structural features in thousands of clips from VED and to every image from NSD. We then estimated selectivity for these features in the brain. We used regularized regression to estimate weights that reflect the importance of each feature in each voxel in the NSD fMRI data. To find dimensions of tuning for 3D structural features, we ran principal components analysis (PCA) on the estimated regression weights for all voxels in scene-selective areas. To find dimensions of 3D structure in visual experience, we ran PCA on the 3D structural feature values in the VED clips. Intriguingly, we found qualitative and quantitative matches among the resulting dimensions: the second dimension of tuning in the brain was reliably correlated with the first experiential dimension, and both of these dimensions reflected a distinction between open, navigable scenes and close-up scenes. Furthermore, voxels in scene-selective areas showed reliable pattens of projections onto these dimensions, with clusters of voxels responding similarly to relatively open or closed scenes. These results suggest that dimensions of visual experience may provide local organizing principles for feature representations in the cortex.

56.468 HOW VISUAL AND CONCEPTUAL FACTORS PREDICT THE COMPOSITION OF TYPICAL SCENE DRAWINGS

Gongting Wang[1,2] (generalwgt@gmail.com), Ilker Duymaz[2], Matthew Foxwell[4], Micha Engeser[2], 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

When asked to draw a typical living room, which objects would you include? The objects drawn could depend on both visual experience (i.e., the objects present in living rooms previously encountered) and conceptual knowledge (i.e., the semantic relationship between the scene and its constituent objects). Here, we tested to which extent visual and conceptual factors predict object content when participants draw rea-world scene categories. We collected data from 156 participants, which were asked to draw typical exemplars of six scene categories: bathroom, bedroom, café, kitchen, living room, and office. We annotated all objects within these drawings and then computed object occurrence frequencies for each category. Next, we modelled these occurrence frequencies using two predictors: First, to quantify visual experience, we extracted occurrence frequencies for all the objects drawn in each scene category from the ADE20K dataset of segmented scenes. Second, to quantify conceptual knowledge, we computed similarities between object and scene concepts using a word2vec language model. A generalized linear model revealed that both factors uniquely contributed to the composition of drawings, with a combined two-factor model performing better than both single-factor models. This result was relatively stable across scene categories, but the visual predictor consistently yielded stronger predictions. Further, we show that even objects drawn less frequently are still diagnostic of a scene category, but less so than objects drawn more frequently, an effect predicted by both models. Together, our results demonstrate that visual and conceptual factors jointly determine which objects are included in typical scene drawings.

56.469 STATISTICS OF EGOCENTRIC VIDEO DURING EVERYDAY TASKS

Xueyan Niu[1,2], Lili Zhang[1], Charlie Burlingham[1], Romain Bachy[1], James Hillis[1]; [1]Realty Labs, Meta Platforms Inc., [2]New York University

Power spectrum analysis on curated images provides good insight into the statistical regularities of visual inputs and explains many observations of visual processing, but those images differ strongly from the dynamic visual experience during active behavior in daily life. In this work, we revisited these frequency characterizations on a very large egocentric video dataset recorded using light-weight glasses designed for Project Aria (Engel et al., 2023), for people performing a variety of everyday activities (gardening, sports, cooking, etc.). We found that 1) While photographs of different image categories have distinct average spatial power spectra (Torralba & Oliva, 2003), our analyses of video frames showed similar power spectra (both orientation and spatial scale) across different activities. Notably, previous work analyzed photographs and identified that man-made indoor environments have a higher frequency falloff than natural outdoor ones and suggested that this may be one driver of myopia (Flitcroft, Harb, Wildsoet, 2020). However, we found much smaller differences for egocentric images indoor and outdoor, emphasizing how different data sources may lead to different conclusions; 2) The spatiotemporal power spectrum of egocentric videos showed a significant boost in mid-to-high temporal frequencies and essentially whitening of temporal dynamics for mid-high spatial frequencies, inconsistent with previous analyses on curated videos that identified an inseparable relationship between the spatial and temporal power spectra (Dong & Attick, 1995). This whitening effect was previously observed only when accounting for eye movements in participants wearing a heavy head-mounted device (DuTell et al., 2020). The differences we found between statistics from egocentric video during natural interaction compared to those from curated images and videos point to the importance of using data representative of natural human behavior and, we believe, lay a better foundation for understanding the mechanisms of visual encoding.

56.470 OBSERVERS OVERESTIMATE HOW MUCH THEY SEE ACROSS THE VISUAL FIELD

Michael Cohen[1,2], Helen Feibes[3], Mabel Shanahan[1]; [1]Amherst College, [2]MIT, [3]NIH

How much do observers believe they can see across the visual world, and how accurate are these intuitions. Here, we asked participants to make a series of predictions about their perceptual abilities across the visual field and then tested the accuracy of those predictions. First, we asked observers to predict how far out into the periphery they could tell whether 1) a face was Taylor Swift or Margot Robbie, 2) an object was big or small, 3) a scene was entirely colorful or desaturated in the periphery, and 4) a scene was entirely preserved or scrambled in the periphery. In this case, we used a custom-built mechanical device that allowed observers to move their arms into the periphery to "show" how far out they believed they could identify faces, objects, colors, and scenes. After extensively showing participants the stimuli at fixation, they predicted how far they could identify these stimuli at their 50% threshold. Next, we placed observers in a custom-built display dome that used a projection system to create a 180° field of view across the horizontal axis and used a staircase procedure to identify the point at which observers would were at chance with these tasks. In this case, the results were unambiguous: Observers drastically overestimated their perceptual abilities. For example, participants believed they could tell Taylor Swift from Margo Robbie 53° into the periphery, when they could only do it 17° into the periphery. Taken together, these results demonstrate that observers do not see nearly as much of the world across their visual field as they believe they do. The fact that these effects hold for numerous stimulus types and to such a large degree shows that it is a ubiquitous aspect of visual perception: Observers do not perceive nearly as much as they think they do.

56.471 THE INFLUENCE OF REGIONAL LANDMARKS ON SEX DIFFERENCES IN SPATIAL NAVIGATION: THE MODERATING ROLE OF SENSE OF DIRECTION

Qiliang He[1], Brandon Fross[2], John Perkins[1], Dylan Hunter[1]; [1]the University of Texas at San Antonio, [2]Trinity University

While research has extensively examined how global and local landmarks influence sex differences in spatial navigation, the effects of regional landmarks—visual features that demarcate different sections within complex environments (like color zones in hospitals)—remain unexplored. These landmarks provide general location information but offer neither the precise positional information of local landmarks nor the orientational information of global landmarks. We investigated how regional landmarks affect sex differences in navigation and subsequent decision-making, examining whether these effects are moderated by self-reported sense of direction (SOD). Participants completed wayfinding tasks of varied difficulty in virtual environments with or without regional landmarks, followed by value-based decision-making tasks that depends on the learning outcome of the previous wayfinding tasks. Results revealed that regional landmarks' effects on sex differences are moderated by SOD. In route-learning tasks with regional landmarks present, low-SOD men outperformed low-SOD women, whereas high-SOD women performed equally or better than high-SOD men. In task that require cognitive mapping, the SOD effect disappeared; women performed comparably to men with regional landmarks present but showed significantly worse performance without them. Additionally, when controlling for consistency in memory-based decision-making, high-SOD women demonstrated increased risk-taking behavior in environments with regional landmarks. These findings help reconcile mixed results in the literature regarding sex differences in landmark utilization and decision-making, suggest the importance of considering the interaction between environmental factors and spatial ability in understanding navigation-related sex differences. Our results have practical implications for designing navigational aids in modern buildings.

56.472 ASSESSING THE EFFICACY OF VISUAL AUGMENTATIONS FOR HIGH-STRESS NAVIGATION

Lily Turkstra[1], Apurv Varshney[1], Jiaxin Su[1], Scott Grafton[1], Barry Giesbrecht[1], Mary Hegarty[1], Michael Beyeler[1]; [1]University of California, Santa Barbara

Interest in integrated augmented and virtual reality (AR and VR) systems is increasing as research expands our understanding of its ability to optimize information intake and its potential to increase efficiency and safety. Visual augmentations like minimaps, in-world arrows, and compasses are widely used in AR and VR to aid in navigation tasks, but their effectiveness under high-stress scenarios remains unexplored. Stress may impair navigation by increasing cognitive load and reducing the ability to notice shortcuts and maintain necessary situational awareness. This study aims to evaluate the efficacy of these established augmentations in high-stress conditions to understand their impact on wayfinding performance and cognitive processing. We used immersive virtual reality (VR) to create a fully customizable virtual environment in order to simulate AR augmentations under controlled conditions. Participants first learned a maze by following a fixed route with visual landmarks. They then navigated the same maze under non-stress and high-stress conditions, using either a minimap, an in-world arrow, a compass, or no augmentations. High-stress scenarios included blocked paths, frightening audio cues, and a time-limited countdown. Navigation performance was assessed via direct measurements of success and efficiency. Supplementally, use of eye-tracking data to measure attention to augmentations, alongside individual differences in spatial understanding and visual attention, NASA-TLX cognitive load surveys, and qualitative feedback were gathered to assess attention and usability. Results indicate that specific augmentations, such as the compass, mitigated stress-induced navigation impairments. Others, like the minimap, increased cognitive load and hindered performance, possibly due to attentional overload. Performance improvements were often consistent with participants' preferred augmentations and perceived reductions in workload, though they did not always correlate with lower stress levels or greater precision. This work contributes to understanding how visual aids can be optimized to support navigation under pressure, offering insights for improving integrated VR and AR-based navigation tools.

Research was sponsored by the U.S. Army Research Office and accomplished under contract W911NF-19-D-0001 for the Institute for Collaborative Biotechnologies.

Scene Perception: Neural mechanisms

TUESDAY, MAY 20, 2:45 – 6:45 PM, PAVILION

56.473 THE SUPERIOR PARIETAL LOBULE SCENE REGION IS INVOLVED IN VISUALLY-GUIDED NAVIGATION

Hee Kyung Yoon[1], Yaelan Jung[1], Andrew S. Persichetti[2], Danny D. Dilks[1]; [1]Emory University, [2]National Institute of Mental Health

Previous research has suggested that a newly discovered scene-selective region in the superior parietal lobule (SPL) is involved in visually-guided navigation (i.e., our ability to move about the immediately visible environment, avoiding boundaries and obstacles). More specifically, SPL has been shown to exhibit sensitivity to first-person perspective motion information through scenes – one kind of information necessary for visually-guided navigation – and to be preferentially connected to the occipital place area (OPA) – another scene-selective region involved in visually-guided navigation. Here, we further test the hypothesis that SPL is involved in visually-guided navigation by asking whether SPL i) represents "sense" (left/right) information – another kind of information necessary for visually-guided navigation and ii) responds more while participants actually perform a visually-guided navigation task compared to a scene categorization (control) task. First, using fMRI adaptation, we found that SPL is sensitive to sense information in scenes, consistent with its hypothesized role in visually-guided navigation. Critically, such sensitivity to sense information was not exhibited for objects. Second, a reanalysis of previously published data investigating OPA's role in visually-guided navigation, we found that SPL, like OPA, responded significantly more during a visually-guided navigation task than during a scene categorization task, again consistent with its hypothesized role in visually-guided navigation. Crucially, SPL responded similarly during both the scene categorization task and a "baseline" (one-back) task, revealing its total lack of involvement in scene categorization. Taken together, these results, coupled with prior evidence, reveal that SPL is indeed a scene-selective region involved in visually-guided navigation.

National Eye Institute

56.474 THE ADDITION OF STEREOPSIS TO NATURAL SCENES ENHANCES HUMAN BRAIN ACTIVATION IN SCENE-SELECTIVE REGIONS

Sofia Varon[1], Carol Coricelli[1], Eva Deligiannis[1], Karsten Babin[1], Kevin Stubbs[1], Laurie M. Wilcox[2], Jody C. Culham[1]; [1]Western University, London, ON, Canada, [2]York University, Toronto, ON, Canada

Information about 3D layout is crucial for many tasks in real-world scenes, including reachable spaces, where binocular disparity may be particularly effective for estimating the distances of objects to guide adept reaching and grasping actions. While extensive neuroimaging research has investigated 3D vision using simple stimuli (e.g., stereograms) and scene processing using static 2D photographs, little is known about how neural scene processing utilizes realistic 3D information. We used 3-Tesla functional magnetic resonance imaging to investigate how scene processing is affected by the addition of binocular disparity and motion parallax in 24 participants with normal binocular vision. We rendered five virtual scenes of reachable spaces (e.g., desk, kitchen table) with viewing geometry and binocular disparities consistent with the real world. These scenes were presented under five viewing conditions: 2D Static (same static image presented to each eye); 3D Static (binocularly disparate images presented to each eye); 2D Translation (video of the scene shifting without relative movement between objects); 2D Parallax (video of the scene with relative movement between objects consistent with motion parallax); 3D Parallax (video of the scene with both parallax and binocular disparity). Compared to 2D Static images, 3D Static images showed enhanced activation in the right parahippocampal place area, occipitoparietal regions (V3A, V3B, IPS0, with partial overlap with the occipital place area), MT+/LO, and the supramarginal gyrus. Compared to 2D Translation, 2D Parallax led only to a modest increase in activation in MT+. Compared to 2D Parallax, 3D Parallax led only to a modest increase in activation in the left occipital place area. While we found little impact of depth from motion, the information provided by binocular vision about the distances of objects from oneself and each other may play a more important role in scene processing than previously appreciated.

Canadian Institutes of Health Research (PJT 190159), Canada First Research Excellence Fund "BrainsCAN" grant

56.475 MAINTAINING VS. UPDATING REPRESENTATIONS: ROLES OF PARAHIPPOCAMPAL, OCCIPITAL, AND RETROSPLENIAL CORTICES IN SCENE PERCEPTION

Treedom Beiyin Zhang[1] (bz1166@nyu.edu), Seoyoung Lee[2], Olivia S. Cheung[1,3]; [1]New York University Abu Dhabi, [2]University of Chicago, [3]Center for Brain and Health, NYUAD Research Institute

Scene perception requires maintaining a stable representation of the environment while updating information as we move our eyes, heads, or bodies. What are the respective contributions of various regions in the scene-selective cortical network, namely bilateral parahippocampal place area (PPA), occipital place area (OPA), and retrosplenial cortex (RSC), in balancing these demands? Using fMRI, we investigated this question by showing participants (N=20) scene segments that varied in two types of content: each image trio showed either identical or partially overlapping views of a scene, or completely different scenes. For overlapping views, each scene was divided into three segments (66% overlap between the first and second segments, and 33% overlap between the first and third segments). We hypothesized that scene-selective regions responsible for maintaining stable representations would show differential neural activity between identical/overlapping views of the same scene, compared with completely different scenes. Conversely, the regions responsible for updating representations would show differential neural activity between identical views vs. all other conditions, since both overlapping views of the same scene and completely different scenes require updates. Univariate and pairwise classification analyses revealed that the three regions showed different sensitivity to maintenance vs. updating of scene information. Specifically, PPA showed sensitivity to both maintenance and updating, with stronger activity for completely different scenes than overlapping views, and stronger activity for overlapping than identical views of a scene. These differences were confirmed using classification analysis. OPA showed some evidence for maintenance, as revealed by classification analysis, but consistent evidence for updating, as revealed by univariate and classification analysis. RSC, however, only showed consistent sensitivity to updating. A control region (lateral occipital complex) did not show evidence for either maintenance or updating. These findings highlight the complementary roles of maintenance in ventral/lateral regions and of updating in all regions in the scene-selective cortical network.

The research was supported by a faculty grant (AD174) and a Tamkeen NYU Abu Dhabi Research Institute grant (CG012) to O.S.C., and a NYU Abu Dhabi capstone grant to S.L.

56.476 VALENCE OF THE SCENE CONTEXT, NOT THE PEOPLE IN THE SCENE IS CAPTURED IN THE VISUAL CORTEX AND VISUAL ARTIFICIAL NEURAL NETWORKS

Elahe Yargholi[1], Laurent Mertens[2,3], Joost Vennekens[2,3,5], Jan Van den Stock[4], Hans Op de Beeck[1]; [1]Department of Brain and Cognition, Leuven Brain Institute, Faculty of Psychology & Educational Sciences KU Leuven, 3000 Leuven, Belgium, [2]KU Leuven, De Nayer Campus, Dept. of Computer Science J.-P. De Nayerlaan 5, 2860 Sint-Katelijne-Waver, Belgium, [3]Leuven.AI - KU Leuven Institute for AI, 3000 Leuven, Belgium, [4]Neuropsychiatry, Leuven Brain Institute KU Leuven, 3000 Leuven, Belgium, [5]Flanders Make@KU Leuven, 3000 Leuven, Belgium

Humans can evaluate the emotional meaning of complex social interactions in real-life settings, but it is unclear how this assessment is achieved. Previous evidence from the human brain and AI models pointed to visual processing as the primary seat for processing emotional valence, but this conclusion may not generalize to complex scenes involving social interactions. Here, we prepared stimuli depicting social human interactions in emotionally loaded scene contexts, e.g. funerals. Across the full set, the valence of the people in the scene was partially dissociated from the valence of the scene context, e.g. people laughing at a funeral. Neuroimaging (fMRI) responses showed that visual areas represent the emotional valence of the scene context and not the valence of people in the scene. Category-selective areas are not the main regions for coding valence; instead, they respond to properties of elements related to the category preferences of these regions. Scene-selective regions have a significantly correlated level of activation with valence of the scene context; a negative correlation consistent with negativity bias theory. However, the valence of people in the scene is not captured in face/body selective regions. Neural responses selective to the valence of people in the scene are only generalized across images in the association cortex. AI responses showed existing models for image valence processing rely mostly on the valence of the scene context while advanced multi-modal AI models that integrate text and vision partially capture the valence of the social interactions on top of the valence of the scene context. We show how basic visual processing captures the basic emotional associations of objects and scenes. Yet, higher levels of processing are needed, in AI models and distributed across the human association cortex, to capture the valence of social scenes when it is at odds with basic properties of images.

This work was funded by KU Leuven grant IDN/21/010.

56.477 PREDICTION EFFECTS IN EARLY VISUAL PERCEPTION

Anmol Kaur[1], Cheryl Olman; [1]University of Minnesota

The brain integrates different cues and information to perceive the world, but disruptions to these cues can lead to confusion about size and depth. For example, two objects with the same angular size that appear to occupy different portions of the visual field will also activate different areas of V1 (Murray, Boyaci, & Kersten, 2006). The psychophysical approach of measuring spatial frequency (SF) sensitivity allows us to determine if the size illusion ("ball in the hall") provides behavioral evidence for modification of the structure of neuron receptive fields, rather than just the amplitude of their responses, by feedback related to size perception. The present study examined the effect of feedback on SF tuning. Participants detected grating patterns inside two discs differing in perceived size for 5 different SFs (4, 6, 8, 12, 16) with two conditions (context or no-context). A mixed-design ANOVA was conducted to evaluate the effects of SF and location on contrast detection thresholds in the context condition. A significant main effect of SF was found at the Participant:SF level indicating that contrast detection thresholds varied significantly with SF. A paired t-test revealed that the average threshold (N=10) between the Top and Bottom locations was higher for spatial frequencies of 6 cycles per degree and above, and statistically significant at 6 and 8 cpd. These findings suggest that the SF response for the Top ball is shifted to the left relative to the Bottom ball. The findings demonstrate that perceived size influences SF tuning, supporting the role of feedback in modulating neural processing in V1. This result provides behavioral evidence for the modulation of neuron receptive fields by feedback related to contextual size perception, aligning with theories of predictive coding explaining how the brain integrates spatial and contextual cues to resolve discrepancies in size and depth perception.

56.478 REPRESENTATION OF GOAL-MODULATED NAVIGATIONAL AFFORDANCES IN THE HUMAN BRAIN

Jinkook Yu[1], Soojin Park[1]; [1]Department of Psychology, Yonsei University, Seoul Korea

Orientations of visible paths, often directly computable from visual input, play a fundamental role in guiding our navigation. The Occipital Place Area(OPA) has been shown to encode navigationally relevant features of scene environments, such as path orientation and distance. However, real-world navigation incorporates familiarity and goals beyond physical features. For example, two commuters standing at the same intersection may choose different routes to their workplaces based on their respective destinations and familiarity with the routes. This study investigated whether the OPA represents visually identical paths differently based on participants' biased navigational experiences. In the free exploration phase, participants explored four virtual environments to familiarize with them. Each environment included intersections with different storefronts(e.g., bookstore, theater) placed at the end of the paths. In the subsequent training phase, participants were instructed to navigate toward a specified storefront. Importantly, this instruction biased the participants' navigation goals: they were biased to direct to the left or right path within each environment more frequently (9:1 ratio), resulting in goal-driven, asymmetrical navigational experiences at the intersection. During the test phase in the fMRI scanner, participants viewed snapshots depicting the intersections of the explored environments while performing a color-dot detection task to maintain attention. Half of the snapshots depicted intersections where participants had been biased to navigate left during training, while the other half showed intersections where they had been biased to navigate right. Visually, all snapshots included both left and right path orientations and conditions were counterbalanced across participants. Results showed that the multivoxel pattern of the OPA distinguished between scenes based on the participants' biased navigation goals, as revealed by linear SVM classification(N=16). These results suggest that the OPA encodes navigational affordances beyond those computable directly from visual path orientations and can distinguish scenes with visually identical paths based on goal-directed navigational experiences.

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP)(NRF-2023R1A2C1006673)

56.479 NEURAL CORRELATES OF BOUNDARY PROMINENCE AND TYPE IN THE HUMAN BRAIN

Jieun Park[1] (gkrwl8577@yonsei.ac.kr), Soojin Park[1]; [1]Department of Psychology, Yonsei University, Seoul, Korea

What are the systematic cues or transitions in the world that allow us to break down the seemingly unbroken flow of navigation into discrete locations, spaces, and key turning points? Past studies have shown that structural boundaries such as a wall or a doorway help chunk the memory of the events. Here, we asked how the structural boundary prominence and type are represented in the brain. Boundary prominence was determined by the prominence of a perceptual boundary: Hard Boundary (HB) showing a large boundary transition within scenes such as a clear walk-out from an indoor space to an outdoor space; Soft Boundary (SB) showing a subtle transition such as an outdoor patio to an outdoor space; and No Boundary (NB) with an absence of any perceptible boundary or transition. For each boundary prominence condition, four boundary types (Indoor-Outdoor, Doorway, Landmark, Turn) were presented in six second movie clips depicting a boundary transition episode. During scanning, participants (N=10) performed a color-dot judgment task to compare whether the color of the two dots presented sequentially on top of the video stimuli were the same or different. In half of the trials, the two dots were presented within the same boundary, in other half, they were presented across the boundary. The average inter-stimulus interval between the two dots were matched. Results show that the PPA demonstrated a stair-cased activity pattern, with the highest activity for the hard boundary, followed by soft and no boundary. Both the RSC and OPA showed a main effect of boundary prominence, driven primarily by differences between hard and no boundaries. Boundary prominence interacted with type in the PPA and RSC, but not in the OPA. Preliminary behavioral analysis suggests hard boundaries may enhance dot color comparison accuracy across boundaries when dots are presented in segregation across the boundary.

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP)(NRF-2023R1A2C1006673)

56.480 FROM VISION TO DECISION: HOW TASKS TRANSFORM DYNAMIC BRAIN REPRESENTATIONS

Yaocong Duan[1] (yaocong.duan@glasgow.ac.uk), Jiayu Zhan[2], Joachim Gross[3], Rachael E. Jack[1], Robin Ince[1], Philippe G. Schyns[1]; [1]University of Glasgow, [2]Peking University, [3]University of Muenster

The occipital cortex is crucial for early visual feature representation during categorization tasks. Task-irrelevant features are quickly reduced within 120–150ms, while task-relevant features are selectively transmitted into the ventral visual pathway for further processing that influences behavior. The transformation of these task-relevant feature representations along the occipito-ventral pathway, however, remains unclear. To investigate this, a 10-participants study was conducted, where they performed four different 2-Alternative-Forced-Choice categorizations using the same set of 64 base images depicting a realistic city scene over 1,536 trials. These images included embedded targets such as eight face identities varying by gender and expression, and the presence of vehicles and pedestrians. Each trial began with a fixation cross followed by a 150ms presentation of a base image with randomly sampled features using the Bubbles technique to identify critical features for categorization. Participants' brain activity was recorded using MEG, alongside their categorization responses. Using information-theoretic co-information, we examined whether features are represented similarly or differently in MEG activity as they evolve dynamically across occipito-ventral pathway sources. Occipital cortex sources that maximally represented each stimulus feature were identified within 50–120ms and 160–200ms post-stimulus. Co-information was computed to determine if features were represented redundantly (similarly) or synergistically (complementarily) across source pairs over time. We found that task-relevant feature representations change dynamically across space-by-time. Around 100ms, features are redundantly represented across occipital cortex sources. By ~180ms, a new redundant representation of the same features emerges at the junction of occipital, ventral, and parietal regions, indicating a dynamic transition in feature representations. The initial representations interact with parietal regions, while the second set interacts with the ventral pathway, peaking ~300ms, suggesting that only the second representation is involved with internal decision-making. These results highlight the dynamic nature of feature transformation for flexible, task-specific visual categorization decisions in the brain.

This research was funded by the Wellcome Trust [Senior Investigator Award, UK; 107802] awarded to PGS; the European Research Council (ERC) grant [759796] awarded to REJ; the Wellcome Trust [214120/Z/18/Z] awarded to RI.

56.481 ATYPICAL PLACE SELECTIVITY IN THE RETROSPLENIAL COMPLEX IN INDIVIDUALS WITH AUTISM SPECTRUM DISORDER

Taylor L. Li[1] (taylor.li@nih.gov), Andrew S. Persichetti[1], Cassandra M. Levesque[1], W. Dale Stevens[2], Adrian W. Gimore[1], Alex Martin[1]; [1]NIMH (National Institute of Mental Health), [2]York University

A small, but intriguing, behavioral literature on individuals with autism spectrum disorder (ASD) has shown that they are impaired when navigating using an allocentric reference frame (i.e., map-based navigation), but are not impaired when navigating using an egocentric reference frame (i.e., visually guided navigation). Meanwhile, there is some neuroimaging evidence in typically developing individuals (TD) demonstrating that the retrosplenial complex (RSC) is part of a map-based navigation system, while the occipital place area (OPA) is part of a visually-guided navigation system. The key identifying feature of the RSC and OPA (putative place-processing cortical regions) is that they respond significantly more to pictures of places compared to pictures of objects or faces (i.e., place selectivity). Therefore, we predicted that in individuals with ASD, the RSC would show significantly weaker place selectivity when compared to a TD control group, while the OPA would not show such a difference between the groups. To test our prediction, we used functional MRI (fMRI) to scan groups of ASD and TD individuals that were matched on age, IQ, in-scanner head motion, and tSNR. During the scan, participants viewed pictures of unfamiliar places, faces, and objects while performing a one-back task. We operationalized place selectivity as the difference between the brain responses to places and the average responses to faces and objects in each region. As predicted, independent-samples t-tests showed that place selectivity was significantly lower in the RSC, but not the OPA, in the ASD group compared to the TD group. Crucially, we also found a significant ROI x Group interaction in a two-way ANOVA. These results suggest that impaired map-based navigation in individuals with ASD may be due to atypical functioning in the RSC.

56.482 CHARACTERIZING THE PROBABLE LOCATION OF SCENE PERCEPTION AND PLACE MEMORY AREAS ALONG THE CORTICAL HIERARCHY – A PUBLICLY AVAILABLE RESOURCE

Adam Steel[1,2], Caroline Robertson[3]; [1]University of Illinois, [2]Beckman Institute, University of Illinois, [3]Dartmouth College

The human brain must integrate visual perception with spatial memory for effective navigation. Recent work has identified place memory areas (PMAs) that process remembered spatial information, positioned anterior to scene perception areas (SPAs) that analyze visual scenes. However, the relationship of the PMAs to the broader cortical hierarchy remains unclear due to limited group-level characterization. We examined PMA and SPA locations across three fMRI datasets (44 participants) that used different acquisition parameters. We identified SPAs in all participants using a standard visual localizer where participants viewed scenes versus faces, and we localized PMAs using a memory task where participants recalled personally familiar places versus familiar faces (Datasets 1-2) or places versus multiple categories (familiar faces, bodies, and objects, and famous faces; Dataset 3). We found PMAs could be consistently localized across datasets and maintained a systematic anterior position relative to SPAs. The relative displacement between PMAs and SPAs was highly reproducible, suggesting a fundamental feature of cortical organization. Group analyses revealed PMAs fall at the boundary between externally-oriented networks (dorsal attention) and internally-oriented networks (default mode), and at an inflection point along the cortical hierarchy between unimodal sensory and amodal, apical regions. Additionally, while SPAs overlapped with retinotopic maps, PMAs were consistently located anterior to mapped visual cortex. These results suggest PMAs represent a transition zone between perceptual and mnemonic systems. We have released probabilistic parcels of these regions to facilitate future research into their roles in spatial cognition.

56.483 HIGH-DIMENSIONAL STRUCTURE UNDERLYING INDIVIDUAL DIFFERENCES IN NATURALISTIC VISUAL EXPERIENCE

Chihye Han[1] (chan21@jhu.edu), Michael Bonner[1]; [1]Johns Hopkins University

Recent work in neuroscience challenges the traditional view that visual cortex reduces sensory inputs to low-dimensional representations, suggesting instead that neural population codes operate on high-dimensional manifolds. While previous work has demonstrated reliable individual differences in low-dimensional cortical responses, the structure of individual variation across higher-dimensional spaces remains poorly understood. Here, we investigated whether individual differences in visual experience are encoded in high-dimensional cortical response patterns during naturalistic movie viewing. We first replicated and extended previous work by showing that the reliable variance in movie-evoked fMRI responses is distributed across the full spectrum of latent dimensions in cortical activity, following a power-law distribution over latent-dimension ranks (here detected across hundreds of orthogonal dimensions). These effects were observed throughout visual cortex, including occipital, ventral temporal, and lateral temporal regions as well as in higher level semantic regions, including the supramarginal and angular gyri. We next constructed individual difference matrices (IDMs) to characterize intersubject variability in these high-dimensional representations. Our IDM analysis revealed reliable individual differences across multiple orders of magnitude of latent-dimension ranks. These high-dimensional differences persisted after accounting for voxelwise inter-subject correlations, indicating that they reveal aspects of individual variability that cannot be detected with conventional analysis methods. Further, we found that the IDMs vary across latent-dimension ranks, suggesting that distinct patterns of individual variability can be observed when considering different subsets of latent dimensions. Remarkably, we found that these high-dimensional differences in cortical representations correlate with differences in how subjects verbally described their visual experiences, quantified through semantic embedding distances between recall transcripts. Together, these results reveal a rich landscape of individual differences in visual processing that extends far beyond previously studied low-dimensional representations, suggesting that variation in subjective visual experience is encoded within largely unexplored high-dimensional neural structure.

56.484 NEURAL ACTIVITY RESOLVED IN SPACE AND TIME THROUGH FUSION OF LARGE-SCALE EEG AND FMRI DATASETS.

Ian Charest[1], Peter Brotherwood[1], Mathias Salvas-Hebert[1], Kendrick Kay[2], Frédéric Gosselin[1]; [1]Département de psychologie, Université de Montréal, Montréal, Québec, Canada, [2]Center for Magnetic Resonance Research (CMRR), Department of Radiology, University of Minnesota, Minneapolis, MN, USA

Understanding spatio-temporal dynamics of neural signals at the scale of the whole brain is essential for understanding brain computation. We present a novel method in which large-scale electroencephalography (EEG) and ultra-high field functional magnetic resonance imaging (fMRI) datasets are combined, yielding a full spatio-temporal trajectory of visual processing in the human brain. We leverage data from the Natural Scenes Dataset (NSD; Allen et al., 2022), which includes single-trial 7T fMRI responses, and the NSD-EEG companion dataset (Brotherwood et al.; 2024), recorded with high-density (164-channel Biosemi) electrodes during viewing of the same 73000 natural scene images (albeit by different participants). EEG channel amplitudes were used as predictors in a cross-validated fractional ridge regression model (Rokem and Kay; 2020) to predict voxelwise fMRI activity. This process is carried out time point by time point, revealing temporal information contained within each fMRI voxel. Our results reveal a distinct spatio-temporal pattern: significant prediction first emerges in the early visual cortex, consistent with initial feedforward visual processing. Over time, predicted activity patterns spread ventrally, dorsally, and peaking in prediction accuracy in parietal regions (Pearson correlation = 0.46), stabilizing roughly 300 ms after stimulus onset and persisting until 750 ms. This data-driven, voxel-level “EEG-to-fMRI” mapping effectively performs an empirical source reconstruction, linking channel-level EEG measurements to the underlying cortical generators. The observed pattern of activation, and its timing, fits well with the known biophysics of EEG—where the measured signals are spatiotemporal mixtures of cortical sources diffused through tissue. This method, when applied to sufficient amounts of high-quality data, provides a promising new avenue for understanding the dynamic neural representations with unprecedented spatio-temporal precision.