The Centre for Vision Research
& Vision: Science to Applications

THE NEW VISTAS IN VISION RESEARCH

DECEMBER 4-7, 2023
SECOND STUDENT CENTRE
YORK UNIVERSITY

Thank You to Our Sponsors

WELCOME!

Welcome to the 2023 Centre for Vision Research (CVR) international conference, 'The New Vistas in Vision Research'! At the CVR our mission is to pursue world-class, interdisciplinary research and training in visual science and its applications. The CVR brings together researchers from psychology, electrical engineering & computer science, biology, physics, digital media, philosophy, kinesiology & health science, and other disciplines in a highly interdisciplinary and collaborative centre. The CVR has held biennial conferences since 1991, each focussed on a current topic in vision science from a multidisciplinary perspective. The 2023 conference is presented in partnership with the Vision: Science to Applications (VISTA) program and celebrates the success of this signature research program funded by the Canada First Research Excellence Fund. At the time it was announced, VISTA was the largest research award ever held at York University. VISTA focusses on the translation of basic vision science across biological and computational vision to impacts in the real world. Opening and closing talks from Rick Wildes and Doug Crawford, who were Associate Director and Director of VISTA through most of the program, will set the tone for the conference and position the VISTA program in the context of the broader vision science community. The conference scientific program and the oral sessions over the next four days will cover many aspects of vision science with a focus on the link between basic science and its ability to transform everyday life. Three poster sessions will expand on these topics and showcase diverse research in all areas of vision science. We hope you enjoy the conference and your time in Toronto. Thanks to the program committee and session chairs for curating such an exciting and intellectually stimulating program. Thanks also to the organizing committee for their attention to detail and for planning and coordinating the events and activities. We are very pleased to be able to offer two BIPOC travel awards with the financial support of Meta and VISTA that encourage the participation of attendees who identify as either Black, Indigenous, or a person of colour. Finally, special thanks to our silver sponsor AMD for their generous support.

Robert Allison
CVR Director

Centre for Vision Research

VISTA VISION: SCIENCE TO APPLICATIONS

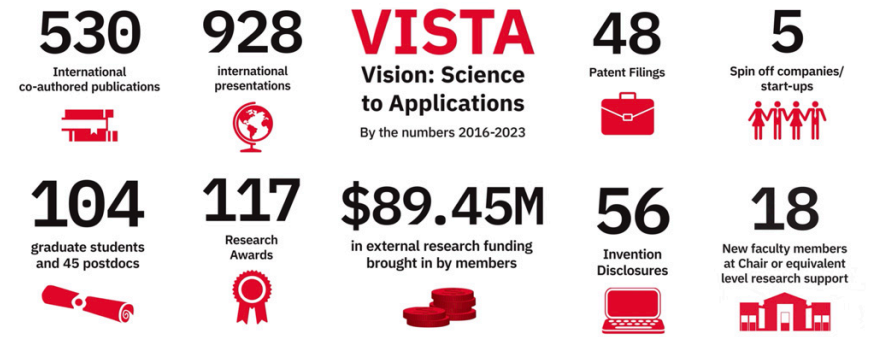
YORK UNIVERSITY

A message from VISTA's new Scientific Director, Laurie Wilcox

Vision Science to Application (VISTA) is a collaborative program funded by the Canada First Research Excellence Fund (CFREF, 2016-2023) that builds on York's world-leading inter-disciplinary expertise in biological and computer vision. In collaboration with over 50 academic, public, and for-profit partners from around the world, VISTA's mandate was, and remains, to establish Canada as a global leader in the vision sciences by integrating visual neuroscience with computer vision to drive innovation.

Our overarching aim is to advance visual science through research that spans computational and biological perspectives and results in real-world applications. Our inter-disciplinary approach, spanning visual neuroscience to computer vision and beyond, has created impact through strategic collaboration with our partners from around the globe.

Building on the foundation provided by the CFREF grant and generous continued support from York University's office of the VPRI, the VISTA program has entered a second phase sustaining and growing vision science as part of the Centre for Vision Research. In the coming years we will continue to support innovative applied vision science through graduate student entrance scholarships, member-initiated workshops and industry-academia focussed events. Our efforts will increasingly focus on supporting our graduate trainees to achieve the highest quality research in collaboration with leading industries and organizations for maximum impact.



CONFERENCE COMMITTEE

Robert Allison
Director, Centre for Vision Research

Caitlin Mullin
Program Director, VISTA & Connected Minds

Irit Printz
Coordinator, Centre for Vision Research

Herb Goltz
Partnerships Manager and Sustainability Officer, VISTA

Sana Saleh Bacha
Financial Analyst, VISTA & Connected Minds

Hadiza Mohammad Saidu
Administrative Assistant , VISTA & Connected Minds

OPENING KEYNOTE

Richard Wildes

Associate Professor, York University

Richard P. Wildes received the PhD degree from the Massachusetts Institute of Technology in 1989. Subsequently, he joined Sarnoff Corporation as a Member of the Technical Staff in the Vision Technologies Lab. In 2001, he joined the Department of Electrical Engineering and Computer Science at York University, Toronto, as an Associate Professor. He also serves as a consultant to the Samsung Artificial Intelligence Center. Previously, he was Associate Director of York's Vision: Science to Applications (VISTA) program and Associate Director of York's Centre for Vision Research. His research interests are in computational vision, especially video understanding and machine vision applications, as well as artificial intelligence. Awards and honours include the IEEE D.G. Fink Prize Paper Award, Sarnoff Corporation Technical Achievement Award, Lassonde School of Engineering Graduate Mentorship Award, holding a Tier 1 York Research Chair and twice giving invited presentations at the USA National Academy of Sciences.

The role of motion in action recognition: Man, machine and application

Intuition might suggest that motion and dynamic information are key to video-based action recognition. In contrast, there is evidence that state-of-the-art deep-learning based video understanding architectures are biased toward static information available in single frames. Resolving this apparent discrepancy has been challenging due to lack of a methodology and corresponding dataset to isolate the effects of dynamic information in video. In response to this situation, my colleagues and I have developed the Appearance Free Dataset (AFD) for action recognition. AFD is devoid of static information relevant to action recognition in a single frame. Modelling of the dynamics is necessary for solving the task, as the action is only apparent through consideration of the temporal dimension. We evaluated 11 contemporary action recognition architectures on AFD as well as its related colour (RGB) video. Our results show a notable decrease in performance for all architectures on AFD compared to RGB. We also conducted a complimentary psychophysical study with humans that shows their recognition accuracy on AFD and RGB is very similar and much better than the evaluated architectures on AFD. Our results motivate a novel video understanding architecture that revives explicit recovery of optical flow, within a contemporary design, for best performance on both AFD and RGB. We also present an application of AFD to privacy preserving action recognition.

Opening Keynote Session, Monday, December 4, 1:00 pm

CLOSING KEYNOTE

Dr. Doug Crawford

Distinguished Research Professor in Neuroscience, York
Research Chair in Visuomotor Neuroscience, York University

Doug Crawford completed his PhD in Physiology at Western University in 1993 and did a post-doc at the Montreal Neurological Institute, before joining York's Department of Psychology in 1995. His research at the York Centre for Vision Research has focused on the control of visual gaze in 3D space, eye-hand coordination, and spatial memory during eye movements. This has resulted in over 170 papers in publications such as Nature, Science and Annual Review of Neuroscience, and has garnered numerous awards, including the 2004 Steacie Prize and the 2016 Canadian Physiological Society Sarrazin Award. He has trained over 60 graduate students and post-doctoral fellows, guiding more than 30 of these into long-term research, clinical and teaching positions. He has founded various groups and programs, including the CFREF-funded VISTA program, which he directed for 7 years until 2023. He currently Directs the CFREF-funded 'Connected Minds' program at York and Queen's Universities.

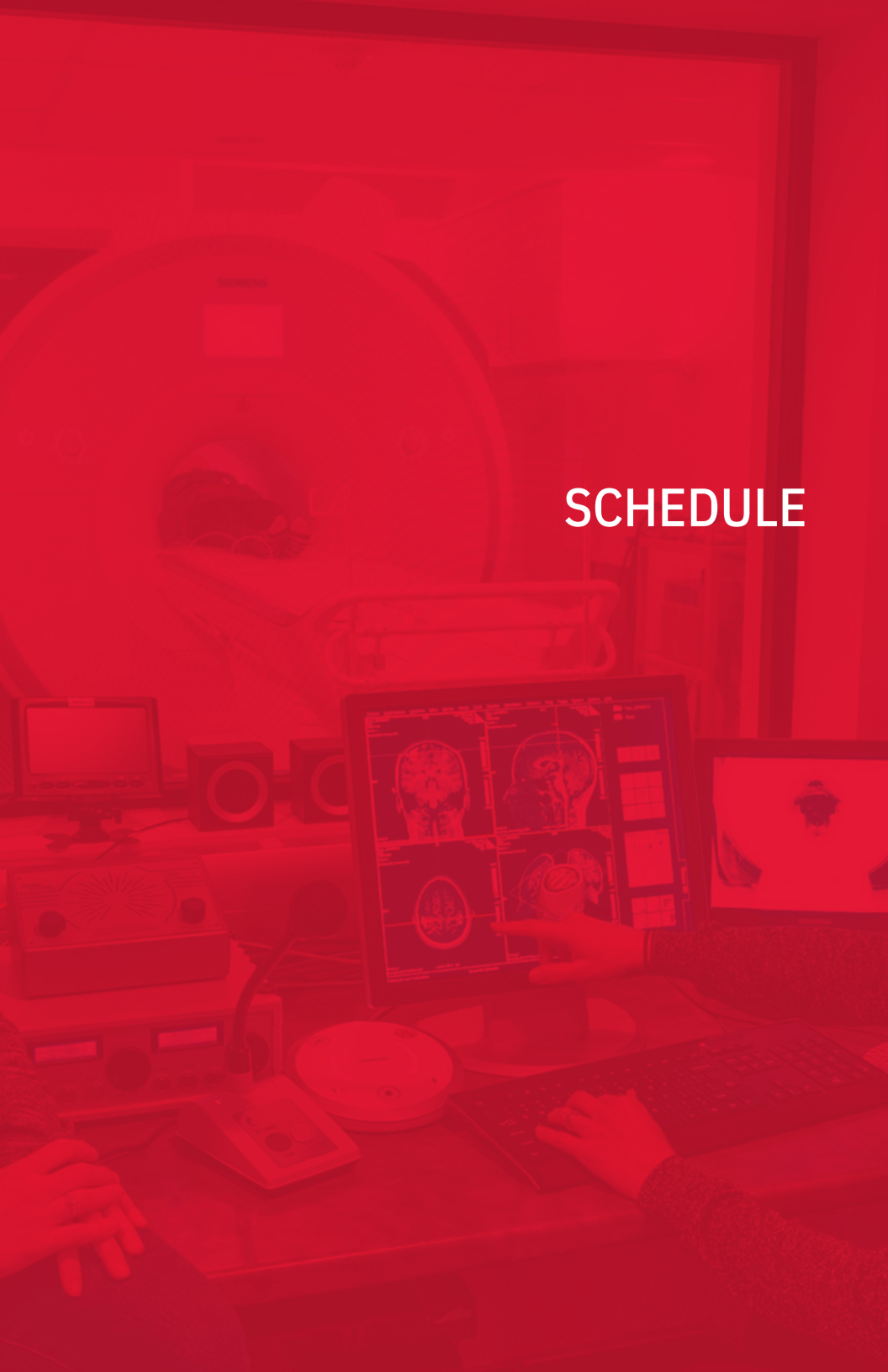
Cortical mechanisms for ego-allocentric integration in the gaze system

The visual system can use either egocentric (relative to the self) or allocentric (relative to the external world) cues for spatial location. This distinction is important for the 'two streams' hypothesis of cortical vision, but until recently was not emphasized in sensorimotor studies. Behavioral studies suggest that ego / allocentric cues are optimally integrated for goal-directed action (Byrne & Crawford) and neuroimaging studies suggest this might happens in frontoparietal cortex (Chen and Crawford 2020) but could not reveal cellular mechanisms. Here, we trained two monkeys to direct gaze toward remembered targets in the presence of a large visual landmark, while recording in the frontal (FEF) and supplementary (SEF) eye fields. Most visual response fields preferentially coded target location relative to the eye, but a substantial minority (30% FEF, 5% SEF) encoded landmark location (Schütz et al., 2023). Further, cells that coded both (targets and landmarks) showed a shift in target coding toward an intermediate target-landmark frame. To test how this information influences behaviour, we surreptitiously shifted landmark position during the memory delay. In both FEF and SEF (Bharmauria et al. 2020, 2921) this produced a partial shift in the spatial code of delay activity, which then became fully integrated with the eye-centered gaze code in the motor response. These results show that prefrontal cortex retains and integrates visual landmark signals with eye-centered target signals to generate the spatial code for gaze commands, potentially providing a general mechanism for optimal ego/allocentric integration in goal directed action.

Closing Keynote Session, Thursday, December 7, 5:00 pm

CONTENTS

SCHEDULE



MONDAY, DECEMBER 4

11:00 am	Registration
	Welcome Remarks
	Robert Allison, CVR Director
1:00 pm	Opening Keynote Richard Wildes - York University *The role of motion in action recognition: Man, machine and application*
	Talk session 1- Visual Cognition Chairs: Dale Stevens and Shayna Rosenbaum
2:00 pm	Zeynep Saygin - The Ohio State University *Innate neural scaffolds for mental function*
2:45 pm	Break
3:00 pm	Charan Ranganath - University of California, Davis *Cognition without walls: the role of structure in perception, episodic memory, and spatial navigation*
3:45 pm	Break
4:00 pm	Alex Martin - National Institutes of Health *Perception from the inside out*
4:45 pm	Break
5:00 - 7:30 pm	Reception / Poster Session 1

TUESDAY, DECEMBER 5

Talk session 2- Visually Guided Reach
Chairs: Jonathan Michaels and Lauren Sergio

9:00 am	Gunnar Blohm - Queen's University *Decoding cortical movement planning computations from magnetoencephalography (MEG)*
9:45 am	Break
10:00 am	Hans Scherberger - German Primate Center *Hand movement control in the primate brain*
10:45 am	Break
11:00 am	Miriam Spering - University of British Columbia *Multisensory integration in visually-guided action*
11:45 am - 1:30 pm	Lunch Break
1:00 - 2:30 pm	Poster Session 2

Talk session 3- Accuracy & Illusion
Chairs: Kevin Lande and Jake Beck

2:30 pm	Fiona Macpherson and Derek Brown - University of Glasgow *Illusions exist: naïve realism is false*
3:15 pm	Break
3:30 pm	Bence Nanay - University of Antwerp *Eye movements and the feeling of presence*
4:15 pm	Break
4:30 pm	EJ Green - Massachusetts Institute of Technology *The complexity of appearances*

WEDNESDAY, DECEMBER 6

9:00 am - 4:00 pm	Industry Day
	Talk session 4- Active Vision Chairs: Erez Freud and Kohitij Kar
9:00 am	Grace Lindsay - New York University *Modeling the neural mechanisms of attention in artificial neural networks*
9:45 am	Break
10:00 am	Maryam Vaziri-Pashkham - University of Delaware *Vision for real-time interactions with objects and people*
10:45 am	Break
11:00 am	Katharina Dobs - Liebig University Giessen *How do real-world task demands shape the functional architecture of the visual system?*
11:45 am - 1:30 pm	Lunch Break/Industry Recruitment/Demos
	Talk session 5- Machine Learning & AI Chairs: Joel Zylberberg and Michael Brown
1:30 pm	Blake Richards - McGill University *A brain-inspired measure for assessing the quality of deep net representations*
2:15 pm	Break
2:30 pm	Chris Pal - École Polytechnique de Montréal *From deep learning, visual perception and language understanding to AI*
3:15 pm	Break
3:30 pm	Eva Dyer - Georgia Institute of Technology *Toward a multi-task, multi-source foundation model to advance large-scale neural data analysis*
4:15 pm	Break
4:30 - 6:30 pm	Lab Tours

THURSDAY, DECEMBER 7

	Talk session 6- Creative Visualizations Chairs: Shital Desai and Laura Levin
9:00 am	Miriam Bopp - Universität Marburg *Augmented reality in neurosurgery*
9:45 am	Break
10:00 am	Jane Tingley - York University *From roots to pixels: Foresta-inclusive*
10:45 am	Break
11:00 am	Michael Proulx - Meta Reality Labs *Visual interactions in Extended Reality (XR)*
11:45 am - 1:30 pm	Lunch Break
12:30 - 2:00 pm	Poster Session 3
	Talk session 7- Guidance & Control of Gaze Chairs: Jeff Schall and Liya Ma
2:00 pm	Eileen Kowler - Rutgers University *Anticipatory smooth pursuit eye movements: Why do we need them, how do they work, and what could possibly go wrong?*
2:45 pm	Break
3:00 pm	Suliann Ben Hamed - Centre national de la recherche scientifique *Dynamics of attentional control at multiple time scales*
3:45 pm	Break
4:00 pm	Greg Zelinsky - Stony Brook University *Reward controls goal-directed attention*
4:45 pm	Break
5:00 - 6:00 pm	Closing Keynote Doug Crawford - York University *Cortical mechanisms for ego-allocentric integration in the gaze system*
6:30 - 9:30 pm	Conference Banquet

SPEAKERS & ABSTRACTS

SPEAKERS

Dr. Suliann Ben Hamed

Research director at CNRS (France), Group leader at Institut des Sciences Cognitives Marc Jeannerod, Adjunct director of Institut des Sciences Cognitives Marc Jeannerod, Institut des Sciences Cognitives Marc Jeannerod

Suliann Ben Hamed has an initial training in mathematics, physics and biology. She is alumina of the École Normale Supérieure of Paris. She defended a PhD in neurosciences from the University Pierre et Marie Curie in 1999, on the visuo-oculomotor functions of the parietal cortex. After a postdoctoral training in computational neurosciences, she was recruited at the CNRS and joined the Institute of Cognitive Sciences Marc Jeannerod in Lyon, France. She is now a CNRS research director, and directs a research group investigating the neural bases of spatial cognition and action. Her research involves comparative behavioral, electrophysiology and functional imaging studies in humans and nonhuman primates, to question the neural bases of attention, perception and multisensory space representation in relation with action. She also serves as adjunct director of the Institute of Cognitive Sciences Marc Jeannerod since 2021.

Dynamics of attentional control at multiple time scales

Attention is a crucial cognitive function that guides gaze through the selection of relevant information and the suppression of irrelevant information. Attentional control is implemented by the frontal eye fields. Recent evidence demonstrates that attentional control is highly flexible and dynamic. I will present electrophysiological evidence from the frontal eye fields describing attentional control dynamics at multiple temporal scales from several cycles per second to a few cycles per hour. I will first describe attentional saccades that are organized at an alpha rhythm, and that subserve the top-down control of attentional exploration of space, the selection of relevant information, as well as both proactive and reactive suppression of irrelevant information. I will then describe, at the scale of several minutes, a two-dimensional neuronal state associating three behavioral states characterized respectively by optimality in task, distractibility and impulsivity. These neuronal states are independent from the neuronal state implementing attention but can interfere with it. Last, I will show that prefrontal attentional control consistently fluctuates at a rhythm of circa 5 cycles per hour, impacting both overall prefrontal information as well as the phase locking between local (MUA alpha) and distal (LFP theta) processes. These fluctuations in attentional control in the range of a few cycles per hour are impacted by noradrenergic neuromodulation. Overall, I will thus show that attentional control that guides gaze should be viewed as a highly dynamic and flexible process encompassing multiple neuronal mechanisms.

Session: Guidance & Control of Gaze, Thursday, December 7, 3:00 pm

Dr. Gunnar Blohm

Professor, Queen's University

Gunnar Blohm (PhD in Engineering & Neuroscience; http://compneurosci.com/) is a Professor for Computational Neuroscience at Queen's University and Vice-Director of the CFREF program "Connected Minds". He is best known for his ground-breaking research in sensory-motor control investigating arm movement planning and the coordination of different oculomotor systems as well as related perceptual phenomena. In his lab, he uses computational modeling and machine learning approaches in conjunction with behavioural measures, psychophysics, brain imaging, brain stimulation and patient work. As an award-winning teaching innovator, he has co-founded (2020) and is co-directing Neuromatch Academy, offering accessible hands-on courses in Computational Neuroscience, Deep Learning and Computational Climate Science.

Decoding cortical movement planning computations from magnetoencephalography (MEG)

To plan a visually guided movement, the brain must choose an effector, calculate an extrinsic movement vector, then convert this into intrinsic muscle commands for current posture. These processes have been studied extensively using neurophysiology and fMRI, but the whole-brain temporal dynamics are unclear. Here, we exploited the spatiotemporal resolution of MEG. Human participants performed a delayed pro/anti wrist pointing task with the left or right hand and 3 different forearm postures. We then computed cortical source activity in 16 previously identified bilateral cortical areas (Alikhanian, et al., 2013) and carried out three analyses: (1) We compared pro/anti trials to identify brain areas coding for stimulus direction vs. movement direction. Sensory activity in α / β bands progressed from posterior to anterior cortical areas, culminating in a β-band movement plan in primary motor cortex. During the delay, movement codes then retroactively replaced the sensory code in more posterior areas (Blohm, et al., 2019). (2) We then asked how effector information (left vs. right hand) was coded and integrated into the movement plan. We found overlapping but distinct networks coding hand use vs. hand-specific motor plans (Blohm, et al., 2022). (3) Last, we contrasted opposing wrist postures to test when and where the extrinsic-to-intrinsic transformation occurred. We found another distinct pair of overlapping networks coding for posture (in γ band) vs. posture-specific movement plans (β). Overall, these results demonstrate several overlapping but distinct cortical networks that carry out different spatiotemporal computations for movement planning.

Session: Visually Guided Reach, Tuesday, December 5, 9:00 am

PD Dr. Miriam Bopp

Head of Medical Technology Laboratory, Department of Neurosurgery, University Hospital Marburg

2003 – 2009 Studies in Computer Science
University of Karlsruhe Germany
2009 Diploma in Computer Science
University of Karlsruhe Germany
since 2009 Researcher (Doctoral Candidate, PostDoc)
Department of Neurosurgery, University Hospital Marburg
2014 Doctorate (Medical Sciences, Dr. rer. med.)
Department of Neurosurgery, University of Marburg
since 2016 Head of Medical Technology Laboratory
Department of Neurosurgery, University Hospital Marburg
2022 Habilitation (Imaging Neuroscience)
Department of Neurosurgery, University of Marburg

Augmented reality in neurosurgery

The introduction of image-guided surgery and neuronavigation in the 1990s had a major impact on neurosurgical practice and is now an indispensable tool for a wide variety of cranial and spinal neurosurgical procedures. First proposed in the 1980s, the commercialization of head-up display (HUD) microscopes in the 1990s introduced microscope-based augmented reality (AR) to the broader neurosurgical community by superimposing tumor outlines derived from imaging data onto the surgical microscope view. AR thereby encompasses the merging of data from realworld environments with virtual information and vice versa. Microscope-based AR complements standard surgical navigation by providing a real-time 3D virtual model that overlays anatomically relevant information such as target and risk structures onto the surgical field, facilitating the transfer of pertinent information between image space and the surgical field, thus reducing the need for attention shifts and improving surgeon comfort. Thus far, AR support in microscope-based AR has primarily served as an intraoperative guidance tool. However, there is potential to use microscope-based AR to evaluate navigation accuracy and update navigation, or to go beyond microscope-based AR by using alternative devices.

Session: Creative Visualizations, Thursday, December 7, 9:00 am

Dr. Derek H. Brown

Senior Lecturer, University of Glasgow

Derek H. Brown is a Senior Lecturer in the Department of Philosophy at the University of Glasgow. He works in philosophy of mind and perception, with particular interest in philosophy of colour, perceptual constancies, 'indirect' approaches to perception, the nature of illusion, imagination, and perceptual demonstratives. He recently co-edited (with F. Macpherson) The Routledge Handbook on the Philosophy of Colour (Routledge, 2021). He is Deputy Director of the Centre for the Study of Perceptual Experience and Director of Postgraduate Teaching for Philosophy at the University of Glasgow.

Illusions exist: naïve realism is false

By definition, perceptual illusions occur when perceptual experiences misrepresent perceived objects in one's environment. For example, if perceptual experience represents an object as being red or round, when it is green or square, then one undergoes a perceptual illusion. There is dispute about how many purported illusions are in fact illusions, or indeed whether any exist at all. For example, rather than appealing to perceptual misrepresentation, perhaps some, or all, alleged instances of illusion can be explained instead by appeal to perception of environmental oddities that we accurately experience, or to post-perceptual cognitive states that misrepresent, as opposed to perceptual ones, or to some combination of these and other options. Naïve realists are largely responsible for developing these strategies to try "explain away" alleged examples of illusions. This is because according to that theory there can be no perceptual illusions. In this paper we examine some of these strategies with regard to colour illusions. While some of these strategies persuade us that some purported colour illusions are in fact not illusions, we pinpoint where the explanatory purchase of the strategies breaks down. We conclude that examples of genuine illusions remain—indeed many of them—and therefore that perceptual experience can and does misrepresent perceived objects. In so doing, we argue that naïve realism is false. Thus, the nature and existence of illusion remains, as it has long been, at the heart of philosophy of perception.

Session: Accuracy & Illusion, Tuesday, December 5, 2:30 pm

Katharina Dobs

Research Group Leader, Justus Liebig University Giessen

Katharina Dobs is a research group leader in the Department of Psychology at Justus Liebig University Giessen (JLU) in Germany. Specializing in cognitive computational neuroscience, her research combines machine learning with cognitive neuroscience to better understand human visual perception in the mind and brain. She was recently awarded an ERC Starting Grant focused on using deep neural networks to understand functional specialization in the human visual cortex. Before establishing her lab at JLU, Dr. Dobs was a postdoctoral researcher at the Center for Brains, Minds, and Machines at MIT, working with Prof. Nancy Kanwisher. She also held a postdoctoral fellowship with Dr. Leila Reddy at CerCo-CNRS in France. Dr. Dobs earned her PhD from the Max Planck Institute for Biological Cybernetics and holds dual diplomas in Computer Science and Psychology from Philipps-University Marburg.

How do real-world task demands shape the functional architecture of the visual system?

Vision scientists have long characterized the functional architecture of the visual system, and they are succeeding increasingly in deciphering how it performs various tasks. However, the question of to what extent the functional architecture is shaped by the task requirements of visual tasks in the real world is less explored, largely due to limitations in human testing. In this talk, I will argue that the new ability to optimize artificial neural networks (ANNs) for performance on human-like visual tasks now enables us to approach this question. Specifically, when a particular behavioral or neural phenomenon spontaneously emerges in ANNs optimized for a task, it suggests this phenomenon may result from the brain's optimization for that same task. I will highlight the recent success of this strategy in exploring how the human face perception system is shaped by real-world task demands, at both behavioral and neural levels.

Session: Active Vision, Wednesday, December 6, 11:00 am

Dr. Eva Dyer

Associate Professor, Georgia Institute of Technology

Eva Dyer (she/they) is an Associate Professor in the Coulter Department of Biomedical Engineering at the Georgia Institute of Technology. Dr. Dyer's research lies at the interface of artificial intelligence and neuroscience, where she aims to both use AI to understand neural computation (AI for Neuro) and use insights from the nervous system to develop robust and lifelong AI systems (Neuro for AI). Eva completed all of her degrees in Electrical & Computer Engineering, obtaining her Ph.D. and M.S. from Rice University and a B.S. from the University of Miami. She is the recipient of a Sloan Fellowship in Neuroscience, NSF CAREER Award, Next Generation Leader Award from the Allen Institute, a McKnight Foundation Technological Innovations in Neuroscience Award, and a CIFAR Azrieli Global Scholar Award.

Toward a multi-task, multi-source foundation model to advance large-scale neural data analysis

Despite the rapid growth of datasets in neuroscience, it has been difficult to integrate many diverse datasets into models of brain function due to variability across experiments and shifts in our measurements of the brain's state. In this talk, I will outline our efforts to consolidate datasets from various tasks, brain regions, and experimental conditions into a cohesive "neurofoundation" model. By leveraging extensive pretraining on neural data, we aim to enable robust generalization across different modalities and species. Further, I will delve into how this foundational model promises to improve data efficiency, expand the capabilities of brain-machine interfaces and neural decoders, and provide advanced, user-friendly tools to the wider neuroscience community, setting the trajectory for a more integrated approach to neural data analysis.

Session: Machine Learning & AI, Wednesday, December 6, 3:30 pm

E. J. Green

Associate Professor, Massachusetts Institute of Technology

E. J. Green is an Associate Professor in the Department of Linguistics and Philosophy at MIT. His research examines questions at the intersection of philosophy of mind and cognitive science, with a significant focus on perception. Recent topics include the nature of spatial perception (within and across sense modalities), the contents of perceptual experience, the format of perceptual representation, and the border between perception and cognition. Before arriving at MIT, he earned his PhD at Rutgers University and then spent a year as a postdoc at NYU. Starting in January 2024, he will take up a position at Johns Hopkins as Miller Associate Professor of Philosophy, with a secondary appointment in Psychological and Brain Sciences.

The complexity of appearances

The philosophy of perception contains a familiar argument that spatial appearances must be complex. Roughly, when you view a circular coin at an oblique angle, there is, many argue, a sense in which it appears circular, but also a sense in which it appears elliptical (or perhaps “elliptical from here”). So, shape perception has both a “constant” aspect that remains relatively stable across perspectives, and a “perspectival” aspect that varies across perspectives. This talk argues that the appearance of shape is considerably more complex than the traditional argument envisages. Specifically, perception science indicates that the constant aspect of shape appearance is layered at varying levels of abstraction. I motivate this view in the context of contemporary skeletal models of shape representation. I then argue that this layered view of shape appearance provides us with the resources to rebut a classic and enduringly influential argument that perception does not put us in touch with the objective spatial properties of objects, such as their intrinsic shapes and sizes, but only with certain “appearance” properties (such as visual angle or some more esoteric alternative). Irrespective of the classic argument’s virtues when it comes to our perception of highly precise or determinate shape properties, I argue that it does not threaten our perceptual awareness of abstract shape properties.

Session: Accuracy & Illusion, Tuesday, December 5, 4:30 pm

Prof. Eileen Kowler

Distinguished Professor, Rutgers University-Piscataway, NJ

Eileen Kowler is a Distinguished Professor of Psychology at Rutgers University. Her research has focused on the importance of high level visual and cognitive processes in the control of eye movements, including anticipatory smooth eye movements, links between oculomotor control and spatial attention, saccadic adaptation, localization, search, and eye movements while reading graphs or text. She has served as President of the Vision Sciences Society (2022), as an Associate Editor of the Journal of Vision, and Oculomotor Section Editor of Vision Research. Kowler also directed an NSF-funded graduate training program on Interdisciplinary Perceptual Science, and served as Senior Associate Dean for Academic Affairs of the Rutgers School of Graduate Studies (2018-2023). She is an elected fellow of the Society of Experimental Psychologists and AAAS, and was the inaugural winner of the Vision Sciences Society's Davida Teller Award.

Anticipatory smooth pursuit eye movements: why do we need them, how do they work, and what could possibly go wrong?

Smooth pursuit eye movements are evoked by anticipated target motion, reducing time delays and improving pursuit accuracy. Anticipatory pursuit may be based on memory for recent target motions, or on perceptual cues signaling when or where a target is likely to move. Rubinstein et al., for example, examined pursuit of clear and noisy random-dot kinematograms (RDKs) whose directions were selected from Gaussian distributions with small or large SDs (i.e., narrow or wide priors). Pursuit directions were initially based solely on the prior. The influence of the immediate RDK motion emerged earlier relative to the start of RDK motion for the wide prior and clear RDKs, a result qualitatively consistent with Bayesian principles. Bayesian model simulations provided an effective, but not complete, accounting of performance. J.Z. Wang et al. found that anticipatory pursuit of a target whose motion was initiated by a collision was biased to follow the path derived from knowledge of Newtonian mechanics. These results attest to the breadth and effectiveness of memory or perceptual cues in triggering anticipatory smooth pursuit, but leave open central questions about scope. At some point the high-level resources needed to compute and store future motion paths based on cues or past history could exceed reasonable limits and begin to interfere with concurrent, task-related operations. A complete understanding of the resources demands of anticipatory pursuit may shed light on the more general question of how the brain monitors effort to produce what it deems to be acceptable levels of performance.

Session: Guidance & Control of Gaze, Thursday, December 7, 2:00 pm

Grace Lindsay PhD

Assistant Professor, Psychology and Data Science,
New York University

Grace Lindsay is an Assistant Professor of Psychology and Data Science at New York University. She holds a B.S. in Neuroscience from University of Pittsburgh and a Ph.D. in Neurobiology and Behavior from Columbia University. Her research focuses on the use of artificial neural networks as models of biological information processing. With her group she studies the relationship between attention and learning and investigates the tools of systems neuroscience. As a faculty member in Data Science, she also works on applied machine learning to tackle climate change. Prior to NYU, Lindsay was a postdoc at the Gatsby Computational Neuroscience Unit at University College London and spent a year as a research fellow at the Bernstein Center for Computational Neuroscience in Freiburg, Germany. She is the author of the popular science book, Models of the Mind: How Physics, Engineering and Mathematics Have Shaped Our Understanding of the Brain.

Modeling the neural mechanisms of attention in artificial neural networks

Countless behavioral studies have demonstrated how validly-cued visual attention can enhance performance on challenging tasks. Neural recordings taken during such conditions of attention have identified the mechanisms that may support this enhanced performance. Given that convolutional neural networks can be thought of as neurally-plausible task-performing models of visual cortex, we can use them to connect neural mechanisms of attention to increases in task performance. I will present a series of completed and ongoing studies that demonstrate this principle. Topics covered will include spatial and feature attention, the relationship between attention and learning, the circuit mechanisms needed for attention, and how these findings can extend to the auditory system as well.

Session: Active Vision, Wednesday, December 6, 9:00 am

Professor Fiona Macpherson

Professor of Philosophy and Director of the Centre for the Study of Perceptual Experience, University of Glasgow

Fiona Macpherson, FRSE, MAE, is Professor of Philosophy and Director of the Centre for the Study of Perceptual Experience, University of Glasgow. She is president of the British Philosophical Association, trustee of the Kennedy Memorial Trust, and member of the UK Research and Innovation Creative Industries Advisory Group. She is currently a Leverhulme Research Fellow. She has published over 40 papers on the mind and perception, and edited/co-edited eight volumes, on topics including colour perception, hallucination, the senses, sensory substitution and augmentation, and imagination and memory. She is co-academic lead on Dreamachine, a live touring show, that received a 5* review in The Guardian, the 2023 Lumen Prize for Immersive Environment and the CogX award for Best Innovation in Creative Arts. She won a 2021 Hidden REF Award for the online Extreme Imagination Exhibition. She is co-lead on The Perception Census, a citizen science project investigating perceptual diversity.

Illusions exist: naïve realism is false

By definition, perceptual illusions occur when perceptual experiences misrepresent perceived objects in one's environment. For example, if perceptual experience represents an object as being red or round, when it is green or square, then one undergoes a perceptual illusion. There is dispute about how many purported illusions are in fact illusions, or indeed whether any exist at all. For example, rather than appealing to perceptual misrepresentation, perhaps some, or all, alleged instances of illusion can be explained instead by appeal to perception of environmental oddities that we accurately experience, or to post-perceptual cognitive states that misrepresent, as opposed to perceptual ones, or to some combination of these and other options. Naïve realists are largely responsible for developing these strategies to try "explain away" alleged examples of illusions. This is because according to that theory there can be no perceptual illusions. In this paper we examine some of these strategies with regard to colour illusions. While some of these strategies persuade us that some purported colour illusions are in fact not illusions, we pinpoint where the explanatory purchase of the strategies breaks down. We conclude that examples of genuine illusions remain—indeed many of them—and therefore that perceptual experience can and does misrepresent perceived objects. In so doing, we argue that naïve realism is false. Thus, the nature and existence of illusion remains, as it has long been, at the heart of philosophy of perception.

Session: Accuracy & Illusion, Tuesday, December 5, 2:30 pm

Dr. Alex Martin

NIH Senior Investigator, NIH/NIMH

Dr. Martin received his Ph.D. from the City University of New York. He is Acting Chief of the Laboratory of Brain and Cognition, and Chief of the Cognitive Neuropsychology Section within the Intramural Research Program of the National Institute of Mental Health. His lab focuses on understanding the neural organization of human memory, perception, and social processing systems and how they breakdown in neuropsychiatric disorders. Dr. Martin is an elected Fellow of the American Association for the Advancement of Science, the Association for Psychological Science, and the American Psychological Association.

Perception from the inside out

Current models of object perception maintain that perception arises from a series of hierarchical stages proceeding from sensory input to conscious perception. That is, from the outside in. This idea runs counter to the neural architecture of perceptual systems as recurrent networks of feedforward and feedback connections. In this talk I'll present neuropsychological and neuroimaging data in support of the idea that object perception occurs mainly from the inside out, as William James's first proposed in 1890.

Session: Visual Cognition, Monday, December 4, 4:00 pm

Bence Nanay

Professor of Philosophical Psychology, University of Antwerp

Bence Nanay is professor of philosophical psychology at the University of Antwerp in Belgium. He has five books published with Oxford University Press, the most recent one, Mental Imagery: Philosophy, Psychology, Neuroscience published in 2023. Three more books with OUP and one with W. W. Norton are forthcoming.

Eye movements and the feeling of presence

Empirical studies show great variation in the smoothness of tracking eye movements in various perceptual situations. In this talk, I examine the relation between the smoothness of tracking eye movements and the feeling of presence in the case of perception, dreaming, visualising, stereograms, concave mirrors, and virtual reality, and argue for a very minimalist and mechanistic explanation of the feeling of presence.

Session: Accuracy & Illusion, Tuesday, December 5, 3:30 pm

Professor Christopher Pal

Full Professor, Distinguished Scientist, Polytechnique Montreal, Mila and ServiceNow Research

Christopher Pal is a Canada CIFAR AI Chair, Full Professor at Polytechnique Montreal, and an Adjunct Professor at the UdeM DIRO. He is also a Distinguished Scientist at ServiceNow Research. He has been involved in Artificial Intelligence and Machine Learning research for over 25 years, often publishing work on large scale language modelling methods and generative modelling techniques. He received his PhD in Computer Science from the University of Waterloo.

From deep learning, visual perception and language understanding to AI

Deep Learning has transformed the field of computer vision over the past decade. More recently the power of neural language models has brought deep learning to the attention of the public in ways that would have been difficult to imagine a decade before. In this talk I'll provide a brief tour of key developments and insights that have led the community to the kinds of architectures that are now widely used for computer vision and language processing, and I'll survey several contributions from the broader Mila community and my own group ranging from: neural architectures for video description to early examples of large language models to new ideas at the intersection of transformers and neural cellular automata. I'll present two recent contributions from my group in more detail: 1) Autobots: a latent variable sequential set transformer for motion understanding in complex scenes, suitable for guiding both autonomous vehicles and robots, and 2) our work on masked conditional video diffusion. Placed in the broader context of the rapid advances in Artificial Intelligence (AI), I'll close my presentation with some remarks on the implications of extremely capable AI systems.

Session: Machine Learning & AI, Wednesday, December 6, 2:30 pm

Dr. Michael Proulx

Research Scientist, Reality Labs Research (Meta)

Dr Michael J. Proulx is a Research Scientist on the Eye Tracking Research team at Meta Reality Labs Research (formerly Oculus Research), where he is working to advance eye tracking and other applications for augmented and virtual reality. He is a Fellow of the Society for Experimental Psychology and Cognitive Science of the American Psychological Association, and Fellow of the Royal Institute of Navigation. He is also currently a Reader in Psychology and Director of the Crossmodal Cognition Lab at the University of Bath. He is co-founder and Co-Director of the REVEAL Research Centre (REal & Virtual Environments Augmentation Labs), Co-Investigator for CAMERA 2.0 (Center for the Analysis of Motion, Entertainment Research and Applications), and a doctoral supervisor in the Centre for Digital Entertainment and ART-AI (Accessible, Responsible, and Transparent Artificial Intelligence) in the Department of Computer Science.

Visual interactions in Extended Reality (XR)

Eye tracking is a tool for understanding human cognition, and a technology that can enable novel interactions with current and future computing platforms. Advances in computational modelling and lower-cost wearable eye trackers provide opportunities for impact in human-computer interaction in Extended (Augmented and Virtual) Reality, with important implications for understanding human perception at a fundamental level, too. I will explore these issues through a few case studies of our research, including: the use of eyetracking with interactive virtual environments, gaze dynamics in real-world tasks, and novel applications of gaze for interaction in virtual reality.

Session: Creative Visualizations, Thursday, December 7, 11:00 am

Professor Charan Ranganath PhD

Professor, University of California, Davis

Charan Ranganath is Director of the Memory and Plasticity Program and a Professor of Psychology and Neuroscience at the University of California at Davis. He uses functional magnetic resonance imaging (fMRI), intracranial and scalp electroencephalography (EEG), neuropsychology, and computational modeling to study the neural mechanisms of episodic memory. Dr. Ranganath has been recognized with several awards, including the Young Investigator Award from the Cognitive Neuroscience Society, the Mid-Career Award from the Psychonomic Society, and a Guggenheim Fellowship. His forthcoming book on the neuroscience of human memory, entitled “Why We Remember,” will be released by Doubleday Press in February, 2024.

Cognition without walls: the role of structure in perception, episodic memory, and spatial navigation

Most research in cognitive and systems neuroscience (including much of my past work) has focused on the roles of specific brain areas within specific domains of research like perception, language, and memory, using paradigms with randomly ordered stimuli and brief trial periods. I will explain how and why this approach has missed important aspects of brain function, and present work based on controlled research paradigms with stimuli that include structured events that play out over long timescales. Results from these studies emphasize the roles of gated interactions between the hippocampus and neocortical networks in comprehension and memory for events and also goal-directed navigation.

Session: Visual Cognition, Monday, December 4, 3:00 pm

Dr. Blake Richards

Associate Professor, McGill/Mila

Blake Richards is an Associate Professor in the School of Computer Science and Department of Neurology and Neurosurgery at McGill University and a Core Faculty Member at Mila. Richards' research is at the intersection of neuroscience and AI. His laboratory investigates universal principles of intelligence that apply to both natural and artificial agents. He has received several awards for his work, including the NSERC Arthur B. McDonald Fellowship in 2022, the Canadian Association for Neuroscience Young Investigator Award in 2019, and a Canada CIFAR AI Chair in 2018. Richards was a Banting Postdoctoral Fellow at SickKids Hospital from 2011 to 2013. He obtained his PhD in neuroscience from the University of Oxford in 2010 and his BSc in cognitive science and AI from the University of Toronto in 2004.

A brain-inspired measure for assessing the quality of deep net representations

Self-Supervised Learning (SSL) with large-scale unlabelled datasets enables learning useful representations for multiple downstream tasks. However, assessing the quality of such representations efficiently poses nontrivial challenges. Existing approaches train linear probes (with frozen features) to evaluate performance on a given task. This is expensive both computationally, since it requires retraining a new prediction head for each downstream task, and statistically, since it requires task-specific labels for multiple tasks. This poses a natural question, how do we efficiently determine the "goodness" of representations learned with SSL across a wide range of potential downstream tasks? In particular, a task-agnostic statistical measure of representation quality, that predicts generalization without explicit downstream task evaluation, would be highly desirable. In this work, we take inspiration from recent neuroscience findings to develop a metric for representation "goodness". Specifically, we observe that the decay coefficient for the eigenspectrum of the empirical feature covariance (α) serves as a measure of representation quality, with networks whose α value is closer to that of the brain showing better downstream generalization. We show how this can be used to tune the hyperparameters of SSL models.

Session: Machine Learning & AI, Wednesday, December 6, 1:30 pm

Zeynep Saygin

Associate Professor, Ohio State University

I'm an Associate Professor at Ohio State University's Psychology Department and director of the Z-lab. I received my Sc.B. in Neuroscience from Brown University, and Ph.D. in Systems Neuroscience from MIT. I'm interested in the origins of the human mind and I investigate how neuroanatomy determines the functional organization of the human brain, and how the existing architecture changes with typical or adverse experience. I use longitudinal neuroimaging to measure brain activity and connectivity in infants and young children, and computational modeling to predict brain and behavioral outcomes. My work in atypical development and brain injury involves using early brain data to predict susceptibility to brain injury, behavioral consequences of head impact and injury, and recovery outcomes.

Innate neural scaffolds for mental function

What determines the functional specialization of the human brain? I'll present experiments that investigate how early developing connectivity drives visual specialization and relevant behavior in typical and atypical development. First, I'll demonstrate how a pre-reader's brain connectivity scan can pinpoint exactly where that same child's visual word form area (VWFA) will develop 2 years later, after they learn how to read. I'll also show how the VWFA does not emerge from the FFA (or other category-selective cortex) but that its continued functional specialization for visual words may drive the right-laterization of the FFA. Then, I'll discuss experiments in neonates that further demonstrate how connectivity (specifically with language cortex) earmarks the eventual location of the VWFA and other visual cortex, and the differential gene expression that provides a potential explanation for the emergence of this basic proto-organization/proto-connectivity of the brain. I'll then focus on the specific connectivity of language cortex with the VWFA and ask whether the VWFA is simply another language region in adults, and whether the absence of canonical language cortex from birth prevents the formation of a VWFA. Finally, I'll present past and ongoing experiments that try to use connectivity as a potential biomarker to predict individual differences in future behavior. These experiments underscore the importance of language cortex and its connectivity in biasing the emergence of other uniquely human cognition. More broadly, they demonstrate how predicting later brain function and behavior from early MRI data can offer powerful strategies for understanding the emergence and plasticity of high-level representations.

Session: Visual Cognition, Monday, December 4, 2:00 pm

Hans Scherberger

Professor for Primate Neurobiology, German Primate Center, Göttingen, Germany

Hans Scherberger received his Master in Mathematics (1993) and his Medical Doctor (1996) from Freiburg University, Germany. He currently heads the Neurobiology Lab at the German Primate Center and is Professor for Primate Neurobiology at Göttingen University (since 2008). He was trained in systems electrophysiology at the University of Zurich (1995-1998) and the California Institute of Technology (1998-2003) before leading a research group at the Institute of Neuroinformatics at Zurich University and ETH (2004-2009). His research focuses on neural coding and decoding of hand movements and their interactions with sensory systems.

Hand movement control in the primate brain

Hand function plays an important role in all primate species, and its loss is associated with severe disability. Grasping movements are complex actions for which the primate brain integrates sensory and cognitive signals to generate meaningful behavior. To achieve this computation, specialized brain areas are functionally connected, in particular in the parietal (anterior intra parietal area, AIP), premotor (area F5), and primary motor cortex (M1 hand area). This presentation will highlight recent experimental results in non-human primates to characterize how individual neurons in these cortical areas interact in order to generate grasping movements on the basis of sensory signals, and how such neuronal population signals can be used to decode hand actions, e.g., for operating a neural prosthesis.

Session: Visually Guided Reach, Tuesday, December 5, 10:00 am

Dr. Miriam Spering

Associate Professor, University of British Columbia

Miriam Spering is Associate Professor in Ophthalmology and Visual Sciences at the University of British Columbia, where she is also Associate Dean of Graduate and Postdoctoral Education in the Faculty of Medicine. She completed her PhD in Psychology at the University of Giessen, Germany (with Karl Gegenfurtner) in 2006 and subsequently conducted postdoctoral work at the Center for Neural Science at New York University (with Marisa Carrasco) before joining UBC in 2011. Research in her lab focuses on the interaction between vision and movement in healthy adults and patients with movement disorders. Her group uses eye movements as a model system for the integration of motor, cognitive, and sensory signals across the senses. For her research work and contributions to mentorship, in 2022, Miriam was awarded a Humboldt Fellowship for experienced researchers and the mentorship award of Women in Cognitive Science Canada.

Multisensory integration in visually-guided action

Seeing and perceiving the visual world is an active and multimodal process during which the eyes continuously scan the visual environment to sample information. My research group uses human eye movements as sensitive indicators of perception and performance in naturalistic tasks such as catching prey or hitting a ball. These tasks require prediction of an object's trajectory from a brief glance at its motion, and an ultrafast decision about whether, when and where to intercept. In the real world, tasks like this rely on both vision and hearing. Here, I will present results from two research programs that use continuous eye movements as a readout of the brain processes underlying the integration of information across the senses. The first series of studies investigates go/no-go decisions in a simulated baseball environment, combining visual and auditory batting cues. Even though auditory cues were not informative, observers' interceptive actions were biased by sound. Continuous eye movements uncovered that sound affected interception within 125 milliseconds of visual target onset, suggesting that auditory and visual signals are combined earlier than previously thought. The second set of studies uses an ultra-rapid orienting task, relying on observers' ability to orient and inhibit eye movements in response to sudden-onset distractors. Combining visual and auditory distractors revealed an additive response mode at latencies as short as 100 milliseconds. These studies demonstrate that eye movements are an excellent model system that reveals the time course, computations, and brain mechanisms underlying multisensory integration.

Session: Visually Guided Reach, Tuesday, December 5, 11:00 am

Jane Tingley

Associate Professor, York University

Jane Tingley is an artist, curator, and Assistant Professor at York University in Toronto, Canada. Her studio work combines traditional studio practice with new media tools - and spans responsive/interactive installation, performative robotics, and telematically connected distributed sculptures/ installations. She has participated in exhibitions and festivals in the Americas, the Middle East, Asia, and Europe - including translife - International Triennial of Media Art at the National Art Museum of China, Beijing, Gallerie Le Deco in Tokyo (JP), Elektra Festival Montréal (CA) and the Künstlerhause in Vienna (AT). She received the Kenneth Finkelstein Prize in Sculpture in Manitoba, the first prize in the iNTERFACES – Interactive Art Competition in Porto, Portugal, and has received support from a number of funding agencies including the Arts Councils of Manitoba, Ontario, and Québec, the Canada Council for the arts, Canadian Foundation for Innovation, and the Social Sciences and Humanities Research Council of Canada.

From roots to pixels: foresta-inclusive

In this talk, Jane Tingley will introduce her project Foresta-Inclusive: (ex)tending towards, which is an artwork comprised of a networked sensing infrastructure (Foresta-Inclusive) and an in-gallery art installation ((ex)tending towards). The sensing infrastructure is comprised of several ecosensors installed in a forest that sense phenomenon (air and soil temperature & humidity, Volatile Organic Compounds, particulate matter (.1-10 micrometers), wind, C02 and rain) and send live data to a cloud platform, which can be then used in different artworks. For this discussion Jane Tingley will talk about the data that was collected at the rare Charitable Reserve in Cambridge, ON in the summer of 2022 and the artwork that visualizes this data entitled (ex)tending towards. This work not only visualizes the hidden experience of the tree and its surrounding ecology, but also the installation and manner of interaction creates an embodied and exploratory space where the deep time of a tree's life is remembered, and the human body is slowed in order to engage – reminding us, that to meaningfully engage with the more-than-human, we ourselves must be prepared to shift our behaviour and decentre ourselves.

Session: Creative Visualizations, Thursday, December 7, 10:00 am

Dr. Maryam Vaziri-Pashkam

Assistant Professor, University of Delaware

Maryam is a cognitive neuroscientist, interested in the intersection of visual cognition and action. She has an M.D. from Tehran Medical University and a Ph.D. in cognitive psychology from Harvard University. After her PhD she worked as a postdoctoral fellow at Harvard University and subsequently as a research fellow at the National Institute of Mental Health. She is currently an assistant professor at the Department of Psychological and Brain Sciences at the University of Delaware and the PI of the Movement and Visual Perception lab. Her research aims to understand the computational and neural mechanisms that enable real-time interaction with objects and people.

Vision for real-time interactions with objects and people

Living in an environment filled with objects and people, we immediately know how to move our bodies to interact with our surroundings. The underlying computational challenge is to dynamically transform the visual input into the continuous space of possible movements. In this talk, I will review my research studying the visual processing that allows interaction with objects and people. In the domain of interaction with objects, I will present behavioral findings that suggest that to grasp objects, humans rely on features that are distinct from those useful for object categorization. Using neuroimaging results, I will then establish the existence of robust and abstract representations for objects in the human parietal cortex, a region that bridges visual and motor areas of the brain that may have a role in extracting object features relevant to grasping. In the domain of interactions with people, bringing evidence from immersive behavioral experiments, I show that humans use their knowledge of the kinematics of the human body to predict others' goals during real-time interactions. Using neuroimaging results, I will then demonstrate that the human parietal cortex may be involved in extracting the kinematics of movements suggesting the role of this region during real-time interactions. Gathering evidence from behavioral and neuroimaging experiments, the research sketched in this talk aims to link the visual input and the continuous output space of movements.

Session: Active Vision, Wednesday, December 6, 10:00 am

Greg Zelinsky

Professor of Psychology and Computer Science,
Stony Brook University

Dr. Zelinsky is a professor of psychology and computer science at Stony Brook University who studies human visual attention from both behavioral and computational perspectives. He and his colleagues design and train deep neural network models to explain the different roles that attention plays in facilitating our recognition of objects and in controlling our goal-directed behavior, particularly the eye movements that people make as they search for objects in the world. Current research projects include: (1) predicting the eye fixations that people make as they search for target objects in natural images, (2) understanding the mechanism of object-based attention and how attention spreads throughout objects, (3) exploiting advances in generative AI modeling to study how attention similarly generates object percepts in our peripheral vision for hypothesis testing and training feedback, and (4) characterizing how the attention behavior of non-experts differs from experts, specifically digital pathologists reading cancer images.

Reward controls goal-directed attention

We show that attention control can be understood as seeking out goal-related reward. Our premise is that achieving a goal is rewarding, and that attention controls goal-directed behavior by computing expected reward and maximizing its receipt. To recover goal-related reward we use inverse-reinforcement learning, a data-driven method for learning a reward policy by training on expert behavior. The goal-directed behavior that we study is visual search, making the expert behavior the eye fixations made by people searching for targets. We trained our model on ~300K fixations from 10 people searching 6202 images for 18 target categories (the COCO-Search18 dataset). We found that predictions of search-fixation scanpaths from our reward map model approached the noise ceiling imposed by agreement in searcher behavior and achieved performance close to state-of-the-art in computer vision for search-scanpath prediction. However, the predictions of reward maps are unique in that they are interpretable, with the model's biological and cognitive plausibility already yielding insights into: (1) how inhibitory tagging could be implemented as a withdrawal of reward from previously searched locations, (2) how the termination of target-absent search can be understood as a depletion of reward, and (3) how a context created by the non-target objects in a scene can be learned and used to facilitate search efficiency. These findings, combined with the reward map model's predictive success, provide converging support for our conclusion that the pursuit of goal-related reward is a plausible theory for how human attention controls goal-directed behavior.

Session: Guidance & Control of Gaze, Thursday, December 7, 4:00 pm

SESSION CHAIRS

SESSION CHAIRS

Jacob Beck

Associate Professor, YRC Tier 2
Philosophy

Michael Brown

Professor, CRC Tier 1
Electrical Engineering & Computer Science

Shital Desai

Assistant Professor YRC Tier 2
Design
Digital Media

Erez Freud

Assistant Professor
Biology
Psychology

Kohitij Kar

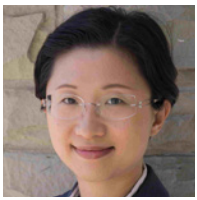
Assistant Professor
Biology

Kevin Lande

Assistant Professor
Philosophy

Laura Levin

Associate Professor
Associate Dean, Research



Liya Ma

Psychology
Cognitive Neuroscience
Perception & Cognition
System Neuroscience

Jonathan Michaels

Assistant Professor
Kinesiology & Health Science

Shayna Rosenbaum

Professor, YRC Tier 2
Biology
Psychology

Jeff Schall

Professor
Biology
Psychology

Lauren Sergio

Professor, YRC Tier 1
Kinesiology & Health Science

Dale Stevens

Associate Professor, MRI Director
Psychology

Joel Zylberberg

Assistant Professor, CRC Tier 2
Physics & Astronomy

POSTER SESSIONS

POSTER SESSION 1 December 4, 5:00 - 7:30 pm

Presenter & Other Authors	Title	Board
Aalim Makani Katharina Pöhlmann, Behrang Keshavarz, Raheleh Saryazdi	Pupil dilation in virtual reality and its relation with measures of visually induced motion sickness and vection	1
Amin Fadaeinejad Abdallah Dib, Marcus Brubaker, Nikolaus Troje	Talking heads for game and telecommunication systems	2
Arefeh Farahmandi Parisa Abedi, Gunnar Blohm	An alternative model to divisive normalization for multisensory integration	3
B. Marius 't Hart Patrick Cavanagh	Frame induced position shifts extend outside the frame in space but not in time	4
Camille Proszanski Brittney Hartle, Laurie Wilcox	Does reaching aid the scaling of stereoscopic depth?	5
Elizaveta Yakubovskaya Hamidreza Ramezanpour, Kohitij Kar	Robustness of object position representation across spatial and temporal context in the macaque inferior temporal cortex	6
Emily D'Alessandro Danai Kokkinopoulou, Jasmine Mosavi, Karolina A. Bearss, Ashkan Karimi, Rachel Bar, Joseph FX DeSouza	Dance improves affective state in individuals with Parkinson's Disease	7
Gaelle N. Luabeya Ada Le, Erez Freud, Simona Monaco, J. Douglas Crawford	Cortical integration and functional connectivity in reach/grasp fMRI study	8
Hamid Ramezanpour Sachi Sanghavi, Kohitij Kar	Selective modulation of visual attention and response inhibition by transcranial ultrasound stimulation of pulvinar and globus pallidus internus	9
Hamidreza Dastmalchi Seyed Nima, Tayarani Bathaie, Aijun An	Incorporating GAN and codebook priors for robust blind face restoration	10
Harshitha Koppisetty Robert Allison, Laurie Wilcox	Perception of materials in virtual reality based on their audiovisual properties	11
Helio Perroni Filho Mohammad Akhavan, Nizwa Javed, Arian Haghparast, James Elder	AirChair: smart wheelchair convoys for airports	12
Jason J. Yu Fereshteh Forghani, Konstantinos G. Derpanis, Marcus A. Brubaker	Long-term photometric consistent novel view synthesis with diffusion models	13
Jiali Song Anureet Jeji, Avery Chua, Cristeidy Gonzalez, Benjamin Wolfe	Are unreliable car warnings still helpful to the driver? The effect of auditory cues on road hazard detection in dynamic road scenes.	14
Jinani Sooriyaarachchi Curtis L. Baker, Chang'an Zhan	Cortical state effects vary across different neuronal subclasses in the primary visual cortex	15
John Kennedy Marta Wnuczko	Mirror-image pointing illusion: visible arms, occlusion and overshooting	16

Presenter & Other Authors	Title	Board
Khushbu Patel Laurie M. Wilcox, Laurence T. Maloney, Krista A. Ehinger, Jaykishan Y. Patel, Emma Wiedenmann, Richard F. Murray	Lightness constancy in real and virtual environments	17
Lina Musa Amirhossein Ghaderi, Ying Chen, J. Douglas Crawford	Hierarchical functional modularity of brain networks for egocentric and allocentric memory-guided reaching	18
Matthew C. Macdonald-Dale Stefania S. Moro, Peter J. Kohler, Jennifer K. E. Steeves	Retinotopic mapping of primary visual cortex of people with one eye	19
Matthew Kowal Mennatullah Siam, Md Amirul Islam, Neil D. B. Bruce, Richard P. Wildes, Konstantinos G. Derpanis	Quantifying and learning static vs. dynamic information in deep spatiotemporal networks	20
Onoise Gerald Kio Robert Allison	Data-driven prediction of outcomes of a target search task	21
Parham Eftekhar Gene Cheung	Image satellite inpainting	22
Ricky Chow Stevenson Baker, Deena Herman, Shimin Mo, Jennifer A. Bugos, Claude Alain, R. Shayna Rosenbaum	Mismatch negativity as an index of auditory pattern separation in aging	23
Seohee Han Morteza Rezanejad, Dirk B. Walther	Memorability of line drawings of scenes: the role of contour properties	24
Simran Rooprai Harsimran Dogra, Joseph FX DeSouza, Jenna Smith-Turchyn, Emily D'Alessandro, Nicole Anderson, Karolina Bearss	Assessing gait and cognitive outcomes in individuals with Parkinson's Disease after a six-year dance intervention	25
Tasfia Ahsan Erez Freud	From 2D to 3D: importance of depth information to high level vision	26
Teodora Neagu Rebecca L Hornsey, Arleen Aksay, Laurie M. Wilcox	Integration of motion parallax and stereopsis for surface segmentation	27
Tim Henley Jason Pina, Josh Tindell, Jérôme Lecoq, Natalia Orlova, Shiella Caldejon, Blake Richards, Joel Zylberberg	Reconstructing pyramidal neurons recorded from awake behaving subjects	28
Viswajit Vembukumar Nikolaus Troje	Evaluating gaze perception with simulated motion parallax	29
Xianze Meng Alexander Barnett, Joseph FX DeSouza	Structural & functional connectivity correlation between PMN and Hippocampus during segmentation and recall of naturalistic stimuli	30
Yeganeh Gharedaghi Gene Cheung, Xianming Liu	Retinex-based image denoising / contrast enhancement using gradient graph laplacian regularizer	31
Zainab Haseeb Silvia Guidi, Anna Kosovicheva, Benjamin Wolfe	Stimuli for sentence readability – a new, validated corpus for studies of digital reading	32

POSTER SESSION 2 December 5, 1:00 - 2:30 pm

Presenter & Other Authors	Title	Board
Ahmed Nadeem Björn Jörges, Laurence R. Harris	Perception of motion in depth during visually simulated self-motion	1
Alban Flachot Jaykishan Patel, Khushbu Patel, Tom S. A. Wallis, Marcus Brubaker, David H. Brainard, Richard F. Murray	Can deep neural networks for intrinsic image decomposition model human lightness constancy?	2
Anaa Salim Zafer Sara Djambazovska, Hamidreza Ramezanpour, Gabriel Kreiman, Kohitij Kar	The impact of scene context on visual object recognition: comparing humans and monkeys	3
Andrew Chaston Naomi Thomas, Ewa Niechwiej-Szwedo	Investigating eye-hand coordination during the performance of Fitts' task	4
Andrew King Laura Mikula, Shanaathanan Modchalingam, Jacob Boulrice, Denise Henriques	Can tools be used as cues for visuomotor adaptation in virtual reality?	5
Elena McKee Lisa Christian, Ewa Niechwiej-Szwedo	Exploring relationships in typically developing children through vision and motor tests	6
Faruq Afolabi Xue Teng, Rob Allison, Laurie Wilcox	Understanding affordances and the effects of distortions in shared mixed reality environments	7
Hongyi Guo Alexander Schütz, Robert Allison	Assimilation of optic flow in saccadic eye movements	8
Jaykishan Patel Alban Flachot, Javier Vazquez-Corral, David H. Brainard, Thomas S. A. Wallis, Marcus A. Brubaker, Richard F. Murray	A deep convolutional neural network trained to infer surface reflectance is deceived by mid-level lightness illusions	9
Julie Ouerfelli-Ethier Pisella, Laure, Khan, Aarlenne Z.	Spatiotemporal competition resolution during anti-saccades	10
Khushi Patel Laurie M. Wilcox, Laurence T. Maloney, Krista A. Ehinger, Jaykishan Y. Patel, Emma Wiedenmann, Richard F. Murray	Effects of different highlighting styles on reading speed and comprehension	11
Krista Mitchnick R. Shayna Rosenbaum, Boyer D. Winters	Object, tactile, and spatial oddity judgements are impaired in DG-compromised rats but enhanced in CA1-compromised rats	12
Lynn Schmittwilken Paul Jolly; Michele Rucci	Foveal perimetry: towards mapping visual sensitivity across the foveola	13
Maren Wehrheim Na Yeon Kim, Ralph Adolphs, Kohitij Kar	Probing the link between dynamics of "face-selectivity" in macaque IT cortex and facial emotion discrimination behavior	14
Marta Wnuczko John M. Kennedy, Selene Carboni	Directions yes, screens no: how perspective is used by an early-totally-blind man	15
Mohamed Abdelhack Akio Murakami, Keita Suzuki, Fan Cheng, Kei Majima, Yukiyasu Kamitani, Hidehiko Takahashi	Uncovering the visual processing imbalance in schizophrenia using deep neural network representations	16

Presenter & Other Authors	Title	Board
Na Yeon Kim Kushin Mukherjee, Shirin Taghian A lamooti, Ralph Adolphs, Kohitij Kar	Leveraging artificial neural networks to enhance diagnostic efficiency in autism spectrum disorder: a study on facial emotion recognition	17
Naila Ayala Ewa Niechwiej-Szwedo, Suzanne Kearns, Elizabeth Irving, Shi Cao	Exploring Gender Differences in Gaze Behaviour During Simulated Flight	18
Nicole Smeha Diana J. Gorbet, Lauren E. Sergio	Visuomotor control networks change as a function of hormone levels in working-aged women.	19
Nikita Klimenkov Svetlana Kovalenko, Elena Gorbunova	The influence of perceptual load on cognitive fatigue	20
Nupur Katyare R. Wong, S. Everling, Liya Ma	Prefrontal ensemble dynamics in spatial working memory in marmosets	21
Parsa Balalaie Kayne Park, Stephen H. Scott, Jolande Fooken	Investigation of eye-hand coordination during rapid interception task	22
Patrick Cavanagh Stuart Anstis	The frame effect is suppressed for stationary probes	23
Pranavan Thirunavukkarasu Steven P. Errington, Amirsaman Sajad, Jeffrey D. Schall	Laminar architecture of visual responses in supplementary eye field of macaques	24
Remy Cohan Jennifer Steeves	The effect of continuous theta burst stimulation to primary visual cortex on binocular rivalry	25
Romesa Khan Matthias Niemeier, Hongsheng Zhong, Jack Cai	Predictive coding dynamics in a grasping neural network increase robustness to noise	26
Sanjana Kapisthalam Martina Poletti	Temporal modulations of extrafoveal sensitivity to changes during fixation	27
Sara Djambazovska Gabriel Kreiman, Kohitij Kar	Visual angle and image context alter the alignment between deep convolutional neural networks and the macaque ventral stream	28
Shanaathanan Modchalingam Andrew King, Denise Y. P. Henriques	Effects of sensory prediction errors and visual environment cues on internal model updating and switching during motor learning	29
Shirin Taghian Alamooti Hamidreza Ramezanpour, Na Yeon Kim, Ralph Adolphs, Kohitij Kar	Developing a non-human primate model to dissect the neural mechanisms of facial emotion processing relevant in autism spectrum disorder	30
Shouyu Ling Max Kramer, Sophia Robert, Christina Patterson, Michael Granovetter, William Welch, Arish Alreja, Avniel Ghumann, Marlene Behrmann	Neural Dynamics of Category-Specific Visual Perception in Adolescents: Insights from sEEG	31
Yousif Kashef Alghetaa Simon Kornblith, Kohitij Kar	Quantifying Alignment between Human and Machine Explanations: A Novel Approach Using Explanation Masked Images	32

POSTER SESSION 3 December 7, 12:30 - 2:00 pm

Presenter & Other Authors	Title	Board
Arleen Aksay Brittney Hartle, Robert S. Allison, Elizabeth L. Irving, Sion Jennings, Laurie M. Wilcox	Visual Illusions in Aviation: Simulating the Black Hole Phenomenon	1
Ashley Clark Martina Polet	Eccentricity driven modulations of visual crowding across the central fovea	2
Bjoern Joerges Laurence Harris	Estimating Speed and Time-to-Contact During Visually Simulated Lateral Self-Motion	3
Brandy Murovec Behrang Keshavarz	Can cognitive factors influence illusory self-motion perception (vection)? Exploring the role of image realism and expectation in younger and older adults.	4
Claudia Damiano Maarten Leemans, Johan Wagemans	Exploring the semantic inconsistency effect in scenes using a continuous measure of linguistic similarity	5
Devin Heinze Kehoe Mackenzie Burgeon R Becket Ebitz	Overt attentional strategies during multi-dimensional value-based decision-making	6
Hamid Ramezanpour Ghazaleh Darmani, Regina Annirood, Can Sarica, Talyta Grippe, Artur Vetkas, Andres Lozano, Samuel Pichardo, Robert Chen	Leveraging computational and animal models of vision to probe atypical emotion recognition in autism	7
Harrison Ritz Jonathan Pillow, Jonathan Cohen	Task Preparation is Reflected in Neural State Space Dynamics	8
Jennifer Lin Hongying Wang, Saihong Sun, Xiaogang Yan, John Douglas Crawford	Influence of a visual landmark shift on memory-guided reaching in monkeys	9
Krista Kelly Dorsa Mir Norouzi, Norah Nyangau, David Stager, Cynthia L Beauchamp, Prashanthi Giridhar	Ocular motor function and postural control during static balance in children with amblyopia	10
Lynn K. A. Soerensen James J. DiCarlo, Kohitij Kar	Probing the effects of object category learning on the macaque inferior temporal cortex	11
Maria Orlando Alberto Umiltà, Federico Fornaciari, Elisa Ciaramelli, R. Shayna Rosenbaum	Examining the Effects of Real-World Experience on Lab-Based Scene Memory	12
Maryum Khan Shanaathanan Modchalingam, Andrew King, Marius 't Hart, Denise Henriques	Effects of tool use and perturbation during motor adaptation on hand localization	13
Mohad Nasir Xiaoxin Chen, Andrew Silva, Ewa Niechwiej-Szwedo	Investigating the Association between Sensory Eye Dominance and Fixation Stability	14
Oluwaseyi Elizabeth Shodipe Robert Allison	Modelling the relationship between the objective measures of car sickness	15
Raphael Gastrock Edward Ody, Denise Y. P. Henriques, Bernard Marius 't Hart	Motor adaptation versus de novo learning: Comparing neural markers of movement preparation and outcome error processing between two distinct visuomotor tasks	16

Presenter & Other Authors	Title	Board
Rezaul Karim Mennatullah Siam, Richard P. Wildes	MED-VT: Multiscale Encoder-Decoder Video Transformer for Video Segmentation	17
Rupsha Mutsuddi Melanie Baljko, Shital Desai	Daily Rituals in Everyday Activities for Community Dwelling Adults with Dementia for Assistive Prompting Technology Design – Understanding the Current Research Landscape	18
S.M. Hossein Hosseini Aleksander Trajcevski, Stefan de Lasa, James Elder	Monocular Depth Estimation Using Semantics and Geometry	19
Sara Chaparian Jeffrey Schall, Peter Kohler	Relating Variability in Scalp EEG to Variability in Cortical Morphology	20
Shaya Samet Baker, N, Freud, E, Elder, J, Kohler, PJ	EEG Studies of Configural Shape Perception	21
Shenoa Ragavaloo Peter Kohler	Brain Responses to Symmetries in Naturalistic Novel Three-Dimensional Objects	22
Takao Fukui Mingze Zhang, Akira Hasegawa	Effects of short-term chopsticks usage on kinematics of subsequent reach-to grasp and uplift movements: Relation to object familiarity and autistic traits	23
Veronica Nacher Parisa Abedi-Khoozani, Harbandhan Arora, Xiaogang Yan, Hongying Wang, John Douglas Crawford	A cortical mechanism for eye-head-hand coordination: lPFC 'gaze' signals encode future head and hand motion during visually guided reach.	24
Wanyi Lyu Jeffrey Schall	Resolving stages of processing in visual search: Investigation of behavioral errors and the underlying covert operations with frontal eye field neurophysiology double factorial design	25
Xue Teng Laurie Wilcox, Robert Allison	Increasing motion parallax gain compresses space and 3D object shape	26
Yael Goldstein Marcusohn Rahaf Asaad, Leen Asaad, and Erez Freud	Shape sensitivity in the dorsal pathway does not depend on attention	27
Yara Mary Iskandar Christopher Lee, Sebastian Bosse, Peter J Kohler	Spatial Mechanisms Mediating Visual Responses to Symmetries in Textures	28
Zoe Stearns Martina Polet	Comparing the temporal dynamics of pre-microsaccadic and presaccadic vision	29
Zoha Ahmad Krista Kelly, Erez Freud	Reduced perception-action dissociation in children with amblyopia	30

LAB TOUR OPTIONS

LAB TOUR OPTIONS

Demos	Title	Location
James Elder	The AirChair project and the SnapBot project	Conference Centre Lobby
Lauren Sergio	Computer vision and augmented reality for mobility and skilled performance assessment	Conference Centre Lobby
Shital Desai	Daily rituals in everyday activities for community dwelling adults with dementia – early prototypes of rituals embedded in assistive technologies for everyday life	Conference Centre Lobby

Lab Tours - Group 1

Denise Henriqes	VR motor learning tasks	Calumet College 304
Doug Crawford	TMS of cortical targets for visuospatial representation	Lassonde 0009 M/N
Rob / Laurie	3D perception in virtual environments	Lassonde 0001 / 3005
Niko Troje	Eye contact behaviour during video communication	Life Science Building 093

Lab Tours - Group 2

Dale Stevens	fMRI facility tour	Sherman 1009
Jennifer Steeves	Image-guided non-invasive brain stimulation with TMS	Sherman 1004
Laurence Harris	Tumbling Room tour	Sherman 1026
Peter Kohler	EEG lab tour	Sherman 1014
Taylor Cleworth	Neuromechanics of balance and mobility	Sherman 2018

Sign up for a tour at the registration desk


Centre for
Vision Research


VISTA
VISION: SCIENCE
TO APPLICATIONS
YORK UNIVERSITY


YORK
UNIVERSITÉ
UNIVERSITY

Presenting author	e-mail address	Additional authors	Poster Title	Poster abstract
Aalim Makani	katharina.pohlmann@uhn.ca	Aalim Makani, Raheleh Saryazdi, Behrang Keshavarz, Katharina Pohlmann	Pupil dilation in virtual reality and its relation with measures of visually induced motion sickness and vection	Visually Induced Motion Sickness (VIMS) and the illusion of self-motion (vection) are two closely related phenomena often experienced in virtual environments. Vection has been shown to enhance experiences in Virtual Reality (VR), by increasing enjoyment, sense of presence, and immersion, whereas VIMS has shown to lead to negative experiences. VIMS and vection are generally measured using subjective rating scales or questionnaires highlighting a need for objective measures. The present study explored the role of pupil dilation on VIMS and vection. Although there is evidence suggesting that VIMS might be related to increased pupil diameter as well as irregular patterns of pupillary rhythms, these findings are not well established. Additionally, only very little is known about the relationship between pupil dilation and vection. The goal of the present study was to address this gap in knowledge. A total of 48 participants (Age: M = 25.52, SD = 5.18; males: 22) were immersed in a virtual scene of outer space for 15 minutes using a VR headset (HP Omnicept Reverb G2). Participants were asked to virtually touch passing objects while being passively moved through the VR scene. Pupil dilation was recorded via an eye tracker incorporated inside the VR headset throughout stimulus exposure and was investigated as a potential physiological correlate of VIMS and vection. We measured participants' level of VIMS during the VR task using the Fast Motion Sickness Scale (FMS), a verbal rating scale ranging from 0 (no VIMS at all) to 20 (severe VIMS). At the end of the VR task, participants indicated on a binary scale whether they had experienced any VIMS during VR exposure. Based on this binary response, participants were either assigned to a VIMS (n = 30) or a no-VIMS (n = 18) group. We also measured vection intensity (0-10) and vection duration (0-100%) at the end of the VR task. Average pupil diameter throughout VR exposure did not differ between the VIMS (M = 4.67, SD = 0.87, IQR = 1.25) and the No-VIMS group (M = 4.81, SD = 0.47, IQR = 0.59; p = .471). Although there was no significant difference in overall pupil diameter between VIMS and No-VIMS group, we did observe differences in the spread of pupil dilation, with a larger spread found for the VIMS group, as measured by Bartlett test ($p < .001$) and as indicated by a larger SD and IQR compared to the No-VIMS group. In contrast, no meaningful correlation between average pupil dilation and the FMS ratings as well as vection intensity or duration was found. Our results suggest that the experience of VIMS during VR exposure may be reflected by fluctuations in pupil diameter rather than the overall pupil size, whereas one's experience of vection does not seem to be linked to changes in pupil dilation at all. This finding can build the foundation for future work looking into pupil dilation as a potential marker of VIMS potentially in combination with additional physiological signals.

Ahmed Nadeem	ahmednad@my.yorku.ca	Björn Jörges, Laurence R. Harris	Perception of motion in depth during visually simulated self-motion	Objects moving at a constant velocity in depth (either towards or away from the observer) produce retinal motion from which their direction and speed can be deduced (Regan and Beverley 1979; Regan et al. 1979; Palmisano 1996). When an observer views an object moving in depth while they themselves are moving it is necessary to parse the resulting movement into components created by the self-motion and those created by the object's movement. Although this parsing has been well demonstrated for sideways movement (Evans et al. 2020), the situation during motion in depth is less clear (Warren and Rushton 2008). Here, we explore whether humans can successfully parse the components of optic flow during motion in depth. 41 participants judged whether the speed of a ball moving in depth at 4 m/s was moving faster or slower than a cloud of small balls moving laterally while they were either (a) stationary or in the presence of visually stimulated self-motion at 4m/s either (b) in the same or (c) in the opposite direction of the target ball. We used the method of constant stimuli with the ball cloud moving between 0.2 and 8.0 m/s in 8 logarithmically equal steps. Participants were able to correctly parse out the self-motion component and consistently match the speed of the ball cloud to the motion of the target ball with a gain of 0.5, undeterred by their simulated self-motion either in the same or opposite direction to the ball. We conclude that for this range of target and self-motion speeds, parsing can be achieved during self-motion in depth.
Alban Flachot	flachot.alban@gmail.com	Jaykishan Patel, Khushbu Patel, Tom S. A. Wallis, Marcus Brubaker, David H. Brainard, Richard F. Murray	Can deep neural networks for intrinsic image decomposition model human lightness constancy?	A challenge in vision science is understanding how the visual system parses the retinal image to represent intrinsic properties of scenes, such as surface reflectance and lighting. Deep learning networks have provided successful new approaches to inferring intrinsic images, and here we investigate these networks as models of human lightness constancy. We examined two state-of-the-art architectures for intrinsic image decomposition (Yu & Smith, 2019; Li et al., 2020), trained on photorealistic images of synthetic scenes. To compare network and human performance, we measured the networks' estimates of surface reflectance using Mondrian patterns embedded in an indoor scene. A reference patch was shown under a fixed illuminant, and multiple test patches were shown under five different illumination levels. At each illumination level, we rendered 17 reflectance levels of the test patch and interpolated the networks' estimates for each to find a reflectance match to the reference, thus probing the networks' lightness constancy. We repeated this procedure for three different reference reflectances. We also tested human observers in a corresponding lightness matching task, using the same stimuli presented with a virtual reality display. Human observers showed good lightness constancy, with an average constancy index (CI) of 0.81 across all stimuli. They were also consistent across conditions, with CI standard deviations around 0.10 across reflectance and lighting conditions. The deep learning networks, however, showed poor reflectance constancy, with an average CI of 0.19. Qualitative analysis suggests that the networks often misinterpreted lighting changes as reflectance changes. The networks were also less consistent than humans, with CI standard deviations of 0.34 and 0.21 across reflectance and lighting conditions, respectively. These results show that these deep learning networks do not fully model human lightness constancy. We will discuss potential strategies to address this shortcoming, as well as proposals for further benchmarking such networks.

Amin Fadaeinejad	afadaei@yorku.ca	Abdallah Dib, Marcus Brubaker, Nikolaus Troje	Talking Heads for game and Telecommunication Systems	Accurate and high-quality face modeling holds immense significance across various industries, including gaming, animation, telecommunications, and cinema, where lifelike representations of human faces are crucial. The demand for face models capturing even the minutest of details, such as pores and wrinkles, continues to grow, particularly in the dynamic gaming industry. Additionally, we aspire for these faces to be rigged, enabling their use in the animation of dynamic scenes, making them capable of expressing movement and emotion. However, the challenge extends beyond realism; it is equally essential for these models to exhibit diversity, ensuring that characters remain distinct and engaging. Existing pipelines, such as the Basel Face Model and FaceWarehouse, utilize statistical variations to represent facial shapes and textures through principal component analysis (PCA). While these models serve as valuable resources, they come with limitations in terms of representing novel and diverse faces. Additionally, these pipelines fall short in capturing and showcasing the finest details. Our motivation was to design a tool that empowers artists and designers to seamlessly control the appearance of a high-quality, detailed face instance, encompassing characteristics such as gender, ethnicity, and age. However, designing such a tool poses numerous challenges, including enhancing controllability while maintaining high-quality outputs. To address these challenges, we harnessed the capabilities of generative adversarial networks (GANs). By leveraging the power of GANs, specifically StyleGAN2 in our case, we achieved remarkable results in the context of high-quality outputs. Furthermore, we implemented several mechanisms to control the details of the face instance, such as selecting gender and race. By capitalizing on the potential of GANs in this context, we significantly enhanced the diversity andrealism of face instances, facilitating their use across various applications in computer graphics,animation, and the gaming industry.

Anaa Salim Zafer	anaazaf@yorku.ca	Sara Djambazovska, Hamidreza Ramezanpour, Gabriel Kreiman, Kohitij Kar	The Impact of Scene Context on Visual Object Recognition: Comparing Humans and Monkeys	Context matters. In the visual domain, our interpretation of the natural world depends not only on the contents within the foveated region but also on the surrounding scene elements and our prior experiences. Beyond psychophysical demonstrations of how context can impact vision, we know very little about the mechanisms that help integrate objects and surrounding information during scene understanding. “Low-level” contextual effects have been studied extensively, including extra-classical receptive fields, temporal adaptation, and surround suppression. However, major lacunae remain in our understanding of how context impacts “higher-level” visual recognition discrimination task. To understand the neural processes behind these contextual influences, we require a more detailed examination of the neural networks involved. Rhesus macaques constitute an ideal animal model due to their similar visual processing circuits to humans. Nevertheless, it is essential to: 1) develop reliable behavioral metrics to quantify contextual effects in humans, 2) determine if macaques exhibit comparable contextual effects. We measured behavior across 90 human participants (on Amazon Mechanical Turk) during a binary match to sample object discrimination task with images having varying context (full context, incongruent context, no context, etc). Not surprisingly, our results show that varying the context of the image significantly changed the performance of the human participants. The effect of contextual manipulations resulted in a consistent pattern of behavior across context categories (with a trial-split reliability of ∼0.8). This was critical to ensure that such effects can be compared with macaques. Once the monkeys (n=2) were fully trained (i.e., reached ≥80% performance) on object categorization (for images in full context), in their home-cages (Giverin, Ramezanpour, & Kar, 2023), we presented them with the same contextually manipulated images as humans.. We observed that humans and monkeys share a significant level of behavioral variance during the image-level contextual variations (~31%). This shared variance was not predicted by low-level image based factors (e.g. object size, object position, contrast, spectral mean, etc.). Interestingly, naive macaque IT neural responses (collected in one monkey) does not explain the human-monkey shared variance (13% of image-level explained shared variance), suggesting that these observed effects may predominantly stem from learning processes and feedback mechanisms rather than image-evoked IT response statistics (shared between naive and trained macaques). Our findings underscore the importance of context in real world object recognition. We establish the limits of using rhesus macaques as an animal model to study the effect of scene context in human visual object recognition, and lay the groundwork for further exploration of the neural mechanisms behind contextual modulation.

Andrew Chaston	achaston@uwaterloo.ca	Naomi Thomas, Ewa Niechwiej-Szwedo	Investigating eye-hand coordination during the performance of Fitts' task	Eye-hand coordination is fundamental for a myriad of activities of daily living, from grasping objects to fine motor control in writing. The ability to perform these tasks requires a tight coordination between eye and hand movements to allow processing of visual information to guide action execution. The role of vision in guiding movements increases as tasks become more difficult, contributing to the speed-accuracy trade-off known as the Fitts' Law. The Fitts' task involves tapping a stylus rhythmically between targets on a screen and has been used to investigate visuomotor control strategies. However, very few studies to date recorded eye and hand movements concurrently. Therefore, the aim of this study was to determine how task difficulty affects eye-hand coordination strategies within the Fitts' task by recording hand kinematics and eye movements concurrently. Twelve participants (6 males; age 24.25 +/- 3.19 yrs) with normal or corrected to normal visual acuity and a stereoacuity performed the Fitts' tapping task. The experiment begun with a familiarization period, followed by 18 trials with 6 levels of index of difficulty (ID's: 2.28, 2,76, 3.50, 4.24, 5.41 & 6.97) which were performed 3 times in a block randomized design. Participants were instructed to "tap" the stylus between the two targets as fast as possible while remaining accurate, while eye movements were recorded with the Eye-Link II system and hand kinematics were captured an Optotrak Motion Capture system. Results were consistent with the literature showing a positive linear relationship between movement time and task ID ($p<0.001$). Higher ID was associated with lower movement peak velocity and long deceleration interval ($p<0.0001$). The mean latency difference between the eye and hand was 68ms (SD 41) with the hand motion onset generally preceding saccade onset. The effect of task difficulty on eye-hand latency difference was trending towards significance ($p= 0.06$). Higher task ID was associated with longer eye-hand latency difference (82ms vs 53ms). In conclusion, examining eye-hand coordination during the performance of a Fitts tapping task revealed that the hand movement typically precedes the eye movement, contrary to the well-established temporal coordination for externally triggered responses where the eye typically precedes the hand by approximately 100ms. Task difficulty had a relatively small influence over the temporal coordination pattern, suggesting that the coordination pattern is stable and relatively hardwired. These results extend our understanding of visuomotor coordination for internally triggered responses.

Andrew King	kinga21@yorku.ca	Laura Mikula, Shanaathanan Modchalingam, Jacob Boulrice, Denise Henriques	Can tools be used as cues for visuomotor adaptation in virtual reality?	Humans are experts at designing and utilizing unique tools to accomplish various tasks, like wielding an axe to chop wood. Humans are also capable of using different tools to accomplish opposing motor tasks simultaneously (using a fork and knife to cut food). While a lot is known about motor adaptation with changed visual feedback of the hand, we rarely consider how adaptation may differ in novel situations requiring pairs of tools with similar or different movement patterns. Here we test whether having two tools that require different movements to accomplish a similar goal would serve as sufficient cues for dual tool-use adaptation, akin to lead-in movements in dual motor adaptations. We ran an immersive virtual reality experiment where participants (n = 164) used one of two sets of tools; a motor incongruent pair, comprised of a paddle (forward motion) and a slingshot (backward motion), or a motor congruent pair, where participants used a paddle and a curling tool (both requiring forward motions). Participants swapped between tools every 8 trials. After a familiarization phase, we added visually opposite perturbations to the ball after it was launched from each tool (30o clockwise or counterclockwise rotation). In addition to the motor congruent and incongruent groups, we collected participants for multiple control conditions to further extrapolate the relationship between these cues and motor learning. Firstly, we collected participants for a single adaptation version of the task, which involved the presentation of each tool successionally, meaning that only one tool and its corresponding perturbation were learnt at a time. We added a "control" pairing of tools composed of functionally identical red and blue coloured paddles to assess the necessity of both cue types on learning. Finally, we included an "invisible" ball condition, where we evaluated the aftereffects of adaptation by making the ball vanish after participants launched their shot. We found that the motor incongruent, motor congruent, and single adaptation groups had reduced angular error in the exposure phase, suggesting that participants were able to form distinct motor memories for both pairs of tools. However, the reduction of angular error in the motor incongruent group was significantly larger than that of the congruent group, which reinforces the notion that internal cues are important markers for the formation and retrieval of motor memories. Additionally, there was no significant reduction of angular error in the coloured paddle condition, supporting previous findings in the literature that visual cues are not strong enough to facilitate the encoding and retrieval of motor memories. This is further reinforced by the results of the invisible ball condition, which found large washout effects following rotated trials. Overall, these findings suggest separate motor memories are more reliant on the movement profile of a perturbation than any associative visual cues.

Arefeh Farahmandi	21afna@queensu.ca	Parisa Abedi, Gunnar Blohm	An Alternative Model to Divisive Normalization for Multisensory Integration	The integration of multiple sensory inputs is essential for human perception and action in uncertain environments. Studies have shown multisensory integration in humans is consistent with Bayesian optimal inference . However, neural mechanisms underlying this process are still debated. Different population coding models have been proposed to implement probabilistic inference. This includes a recent suggestion that explicit divisive normalization accounts for empirical principles of multisensory integration. However, whether and how divisive operations are implemented in the brain is not well understood. More importantly, all existing models suffer from the curse of dimensionality and thus fail to scale to real-world problems. Here, we propose an alternative model for multisensory integration that approximates Bayesian inference. We used a multilayer-feedforward neural network to carry out a multisensory integration (MSI) task. The task was to estimate the position and associated variability of the hand across different eye positions when visual and proprioceptive sensory information of hand position as well as proprioceptive information of eye position are available. The network learned to estimate the hand position and the variance based on the analytical Bayesian solution. Neurobiological studies have shown that the responses of neurons in multisensory brain regions are either enhanced or depressed with respect to their unimodal responses. We quantified these neuronal behaviors in our model by using response enhancement and response additivity indices. The Multisensory Layer of our network showed similar behavior to that reported in VIP neurons in the brain. Furthermore, our model reproduced super-additivity, inverse effectiveness, and gain-like modulations without explicit divisive normalization. Our model achieved this without neatly organized and regular connectivity structure between contributing neurons, such as required by explicit divisive normalization. Overall, we show that simple feedforward networks of purely additive units can approximate optimal inference across reference frames through parallel computing principles. This suggests that it is not necessary for the brain to use explicit divisive normalization to achieve multisensory integration.

Arleen Aksay1	aaksay@yorku.ca	Brittney Hartle, Robert S. Allison, Elizabeth L. Irving, Sion Jennings, Laurie M. Wilcox	Visual Illusions in Aviation: Simulating the Black Hole Phenomenon	Pilots may rely on their own vision to provide the inputs to control aircraft trajectory rather than instrumentation either by choice (e.g., landing in good conditions) or necessity (e.g., instruments are unavailable). During visually-guided flight, pilots are more susceptible to spatial disorientation caused by visual illusions. Pilots must recognize and avoid the danger posed by disorientation to maintain safe flight. However, opportunities to experience and learn to counter these illusions are limited during flight training. The goal of this project was to evaluate the effectiveness of a medium-fidelity flight simulator in generating common visual illusions that are encountered in the aviation environment. To accomplish this goal, we simulated one of the most frequently reported visual illusions, the black hole illusion, using the simulator at the University of Waterloo's Institute for Sustainable Aeronautics (WISA). The black hole illusion occurs at night and can be disorienting during the approach-and-land phase of flight. This illusion typically occurs over a featureless terrain, under an unlit sky, when only the runway lights are visible. In the absence other visual cues, the only information available to land is runway size and shape. As a result, pilots overestimate the aircraft's altitude, and initiate an aggressive descent with a shallow final approach. In addition to evaluating whether it is possible to generate this illusion in a simple modern flight simulator, we aimed to evaluate the relative susceptibility of men and women to this illusion. Most research on visual illusions are anecdotal reports, incident or accident reports, and the experimental studies are often restricted to men, or fail to specify pilot gender. With growing initiatives to create equal gender representation among the aviator population, the lack of data from other genders may affect training and risk assessment. To measure the black hole illusion, we simulated an approach-and-land scenario at a rural northern Ontario airport in Fort Severn using an ALSIM AL250 FSTD simulator at the WISA facility. The flight scenario was first completed at night followed by a daytime control condition at starting distances of 3.5 and 5NM. We measured the black hole illusion by evaluating the deviation of altitude and attitude from the ideal 3 degree glideslope. All participants held a Private Pilot License and a Transport Canada Medical Category 1 certificate. A total of 21 pilots (12 men and 9 women) completed the flight scenarios. Preliminary results showed that a significant proportion of pilots showed a shallow glidepath at the final approach at both distances, consistent with the effects of the black hole illusion. Interestingly, some pilots adopted strategies in their approaches to counteract the effects of the illusion, particularly if they had knowledge from ground school to recognize the illusion and its impact on perceived altitude. To date our results have shown that it is possible to produce the black hole illusion in a relatively low-cost aviation simulator. We will discuss how this phenomenon manifests, and preliminary analyses of the differences between men and women in these scenarios.

Ashley Clark	aclark43@ur.rochester.edu	Martina Poletti	Eccentricity driven modulations of visual crowding across the central fovea	It is well established that surrounding a stimulus with flankers decreases acuity, a phenomenon known as visual crowding. While it has been widely documented that the magnitude of crowding increases with eccentricity, it remains unknown if this increase starts already within the central 1-deg foveola, and if so, what is the rate of growth at this scale. Addressing this question is important as foveal vision is often confronted with crowded stimuli. We measured subjects' (N=6) visual acuity in a 4AFC task. Stimuli were viewed monocularly in either isolation or with surrounding flankers and were presented at different foveal and extrafoveal eccentricities. Stimuli were presented in Pelli font, designed for testing foveal crowding, and their widths, ranging from 0.4' to 4.5', were adjusted using an adaptive procedure. Eye movements were measured with high-precision using digital Dual Purkinje Image eye-tracking. To limit visual stimulation around the desired foveal eccentricity, we used a state-of-the-art custom-made gaze-contingent display system allowing for more accurate gaze localization and retinal stabilization. Our results show that the impact of crowding increases with eccentricity already within the foveola; acuity in the presence of flankers decreases as a function of foveal eccentricity, with a decrease of 32% at the center of gaze, and an additional 24% drop as early as 10' away. Further, crowding increases with eccentricity at a rate that is three times slower in the foveola (from 0-0.4 deg) than extrafoveally (from 1-6 deg). These findings reveal that visual crowding does not affect the whole foveola equally; its impact significantly increases even with minute changes in eccentricity. Therefore, under normal viewing conditions with crowded foveal stimuli, acuity likely drops considerably even a few arcminutes away from the preferred locus of fixation.
B. Marius 't Hart	thartbm@gmail.com	Patrick Cavanagh	Frame induced position shifts extend outside the frame in space but not in time	When two probes are flashed at the same physical location within a moving frame, their perceived location can be offset by as much as the frame moves (Özkan et al, PNAS, 2021). Here we examine the extent of the frame's influence in space and time. First, we positioned the flashed probes in front of or behind the frame in depth (or both) using red/cyan anaglyph glasses. The illusion strength was not affected by these depth mismatches. In contrast, placing the flashed probes outside the frame did influence the illusion. The illusion strength dropped to 50% magnitude once the probes were 5.4 dva to the left of the frame. In the vertical direction, the 50% decrease required an offset of 6.9 dva between the frame and the probes. Offsets in time from the presentation of the frame caused a complete loss of the illusion. The frame was presented for one to three cycles of left-right motion and when the probes were flashed before or after the frame presentation, there was no illusion, no matter how long the frame had been present. This suggests that the illusion depends on immediately present sensory information without any influence of the frame's motion before or after its actual presence on the screen. In conclusion, the frame effects do extend outside the boundaries of the frame in space but not in time.

Bjoern Joerges	bjoerges@yorku.ca	Laurence R. Harris	Estimating Speed and Time-to-Contact During Visually Simulated Lateral Self-Motion	We have previously shown that perceived speed can be biased in the presence of visual self-motion due to incomplete parsing of the visual motion into external and self-generated components [1]. Here, we investigate to what extent such biases in perceived speed might translate to biases in perceived time-to-contact in a motion extrapolation task. We further investigate whether the perception of speed and time-to-contact might be correlated, revealing a shared mechanism underlying both. Having preregistered our hypotheses and methodology as a Registered Report, we tested 40 participants in two tasks administered in a virtual reality environment: in the speed estimation task, they compared the speed of a single sphere moving laterally to that of a sphere cloud in a two-alternative forced-choice task. While the single sphere was displayed, the viewpoint could be moved laterally to simulate self-motion either in the same or the opposite direction of the sideways moving sphere. In the prediction task, they were shown only the single sphere which disappeared after 0.5s and had to indicate with a button press when the now-invisible sphere would reach a rectangular target. We replicated our earlier finding that visual self-motion in the opposite direction of the moving sphere increased its perceived speed but were unable to find a symmetrical bias during simulated self-motion in the same direction as the sphere. In the prediction task, on the other hand, we found convincing evidence that time-to-contact was underestimated for the opposite directions of sphere and self-motion and overestimated when they were in the same direction. In line with these results, we found that biases between the speed estimation and time-to-contact estimations were correlated when the sphere and self-motion were in opposite directions but not when they were in the same direction. While these results cast some doubt on the notion that a shared mechanism underlies speed perception and time-to-contact estimation in the presence of visual self-motion, we present some potentially unifying explanations for these seemingly disparate results. [1] Jörges B, Harris LR. Object speed perception during lateral visual self-motion. Attention, Perception, Psychophys 2022; 84: 25–46.

Brandy Murovec	bmurovec@torontomu.ca	Behrang Keshavarz	Can cognitive factors influence illusory self-motion perception (vection)? Exploring the role of image realism and expectation in younger and older adults.	Vection, defined as the sensation of self-motion in the absence of physical movement, is a critical component for an immersive Virtual Reality (VR) experience. Typically, vection occurs while observing dynamic visual content on large-scale displays (such as VR), producing an illusory sensation of self-motion in the opposite direction of the visual stimulus. Research in this domain has primarily focused on strategies to modulate vection by manipulating physical properties of the stimulus (e.g., speed, size of the field of view) with younger-to-middle aged adult samples. That said, little research has been done investigating the role of top-down, or cognitive factors on vection, and further how these manipulations may differentially influence participants of older age. The aim of the present study was to address this gap in knowledge by investigating the relative contribution of expectation and image realism as cognitive factors to the vection experience in participants of different age groups. To achieve this, 48 younger (15 male, 26 female, Mage=25) and 46 older (20 male, 26 female, Mage= 71.9) participants' expectations about vection were systematically manipulated via contextual information (i.e., half were told that the stimulus was a powerful vection inducer, while the other half were told the opposite). During the experimental task, participants observed a rotating visual stimulus presented on an array of three monitors, which created a virtual scene that induced circular vection about the yaw axis. The visual stimulus varied in terms of realism (photorealistic image vs. scrambled version of the same scene) and speed (fast, slow). During each trial, three measures of vection were recorded: onset latency, duration, and intensity. Results indicated that fast-moving and realistic stimuli considerably increased vection relative to the other stimulus conditions for all three measures. No evidence was found to support a meaningful role of expectation. In fast speed conditions, younger adults experienced vection quicker (but neither longer nor more intensely) than older adults. Taken together, we found evidence that both bottom-up and top-down factors can influence vection independently as well as synergistically, and that the age of participants may not be a meaningful contributor to vection in adulthood.

Camille Proszanski	c23am@yorku.ca	Brittney Hartle, Laurie M. Wilcox	Does reaching aid the scaling of stereoscopic depth?	Depth estimation from stereopsis is biased under many viewing scenarios and for a range of estimation methods, particularly for virtual stimuli. These distortions are often attributed to misestimates of viewing distance that result in incorrect scaling of binocular disparity. Most research on depth scaling has considered only visual cues to distance. However, when exploring the world, we often interact with objects and, in this way, may have access to proprioceptive distance cues. Previous studies have shown that stereopsis aids actions such as reaching and grasping, but is the reverse also true? We assessed the impact of proprioceptive distance from arm's reach on stereopsis using a ring game that is contingent on accurate distance perception. Observers used a handheld motion controller connected to their index finger to move the rings onto a peg in a virtual environment rendered in an Oculus Quest 2 headset. They completed the task as quickly as possible while avoiding touching the side of the ring to the peg. Errors were signalled by controller vibration. After each block of 5 trials, observers were given feedback regarding their completion time and accuracy. To evaluate the impact of reaching in depth, we assessed depth magnitude estimation before and after the completion of the ring task. Observers were asked to estimate the depth between a rectangle and a reference frame located at the same distance as the peg. To control for other visual cues that could contribute to the rescaling of perceived depth, we included a no-reach control condition in which observers completed the ring placement task using the controllers' joystick to place the rings without reaching in depth. This no-reach ring placement task had all the same visual information as the reaching task without the contribution of the physical movement of the reach. Our results were consistent with systematic distortions of perceived depth and distance reported in other virtual environments, where perception of distance is compressed, and depth estimates at these viewing distances are overestimated. We found that there was a small improvement in the scaling of depth estimates on subsequent test sessions; however, this improvement was likely due to practice and not dependent on performing the reach. Overall, there was no significant difference in the accuracy of depth judgements between the reach and no-reach tasks. Our results suggest that while stereopsis plays a critical role in reaching actions, distance information from reaching in depth is a relatively weak cue and does not aid in the rescaling of binocular disparities.

Claudia Damiano	claudia.damiano@utoronto.ca	Maarten Leemans, Johan Wagemans	Exploring the semantic inconsistency effect in scenes using a continuous measure of linguistic similarity	People use contextual information to visually explore complex scenes. Object recognition is facilitated by exploiting object-scene relations (i.e., which objects are expected in a given scene) and object-object relations (i.e., which objects are expected based on the occurrence of other objects). Semantically inconsistent objects deviate from these expectations, and thus they tend to capture viewers' attention more than an expected object in the same location would. This is the semantic inconsistency effect. Some objects fit the identity of a scene more or less than others, yet semantic inconsistencies have thus far been operationalized as binary (consistent vs. inconsistent). In an eye-tracking experiment, participants (N = 21 adults) viewed images of real-world scenes that contained either a consistent or inconsistent target object, according to the pre-labeled Scene Grammar (SCEGRAM) image database. Here, we studied the semantic inconsistency effect in a continuous manner by using the linguistic similarity of the target object to the scene category and to other objects in the scene, rather than the binary consistent or inconsistent label. Linguistic similarity between object- and scene-labels was calculated by taking the cosine similarity of a pair of vectors consisting of 300 values each that represent the meaning of a label in a high-dimensional space, obtained from a database called ConceptNet Numberbatch. Using this continuous measure, we found that the semantic inconsistency effect was negatively related to linguistic similarity between the target object label and the scene category label ($r = -0.37$, $p < 0.01$), as well as between the target object label and other objects within the scene ($r = -0.35$, $p < 0.01$), revealing that the (in)consistency effect is more than a simple binary classification. In other words, the lower the linguistic similarity score (i.e., the semantic consistency) between the target object and the scene category or other objects in the scene, the more people fixated on the target object. In an exploratory analysis, we also included all objects within the scene. We found that both highly consistent and highly inconsistent objects were fixated more frequently and longer than other objects (U-shaped relationship), showing that our continuous measure not only applies to object-scene (in)consistencies but also to object-object relationships.

Devin Heinze Kehoe	dhkehoe@gmail.com	Mackenzie Burgeon R. Becket Ebitz	Overt attentional strategies during multi-dimensional value-based decision-making	We live in an uncertain world where some choices are more rewarding than others. Very often, an option's visual features indicate its reward value (e.g., color may predict taste). Visual features can thus guide our decision-making processes, as we learn to associate them with reward. However, when options have multiple features that interact to determine value, the number of associations that must be learned and remembered increases exponentially. In this context, it seems likely that decision-makers would need to implement some kind of simplifying strategy to guide their choices to the best option. Here, we sought to determine if this is the case by looking at how overt visual attention changes as a function of learning in a multi-dimensional value-based decision-making task. In our experiment, humans made a series of decisions in a dense visual environment made up of options with color and orientation feature conjunctions. Critically, because our task dissociated overt attentional selection (saccades) from the final choice (button press), we could separate the exploration strategies used to investigate options from the strategies that guided the economic choice. Each option's unique combination of features determined its level of reward and reward-feature contingencies occasionally and unexpectedly changed. The change points allowed us to determine how people explored this complex environment both when they understood the feature associations, and when they were learning them. In a preliminary sample of 10 participants, we observed a few intriguing results. First, we found that participants consistently investigated more options after the change points than before. Follow up work will determine whether this is because they continued to have an attentional bias towards previously rewarded options or if they were simply more interested in evaluating more options when they were uncertain. Participants also had biases for both color and orientation, but surprisingly, these biases were not maximal at the change points, but instead most obvious when the participants had already committed to a target. We speculate that this could be due to a censoring process in which participants were only able to successfully find the correct target when it was aligned with pre-existing biases. Additional analyses will detail (1) the strategies participants use to visually explore the feature space during periods of both certainty and uncertainty, (2) the factors that determine when participants will switch between these strategies, and (3) how previously learned associations influence subsequent decisions, be those decisions attentional or economic.

Elena McKee	elena.mckee@uwaterloo.ca	Lisa Christian, Ewa Niechwiej-Szwedo	Exploring Relationships in Typically Developing Children Through Vision and Motor Tests	Vision and fine motor skills, known as hand-eye coordination, are associated with activities of daily living and are an essential component for academic and social development in children. Studies have shown visual impairments such as amblyopia and poor fine motor skills, are associated with reduced reading speed in typically developing children, however, this association has not been fully examined in other neurodevelopmental disorders, such as developmental coordination disorder (DCD). DCD is characterized by reduced gross and fine motor performance, but the mechanism is not fully understood. The overarching goal of this study is to characterize the visual function and motor performance in children with DCD to examine if visual processing deficits contribute to reduced motor performance. To effectively investigate this association, this study explores measures of visual function and motor skills in typically developing children and examines the association between binocular vision and fine motor skills. Twenty-six typically developing children (13 boys, 13 girls; aged 7-14 (9.62 years, SD=2.00) participated in this study. An optometric visual assessment was conducted to measure visual acuity, stereoacuity, vergence facility, accommodative facility and amplitude of accommodation. Motor skills were examined using the Movement Assessment Battery for Children (M-ABC-2). Other tests included the Beery Buktenica Test of Visuomotor Integration (Beery-VMI) and the Test of Word Reading Efficiency (TOWRE-2). All participants had normal visual acuity (-0.04 logMAR, SD=0.06) and stereoacuity (25.19 arc sec, SD=10.24; range 20-70). Vergence facility ranged between 6.00 and 30.50 cycles per minute (13.81 cpm, SD=4.98), mean binocular accommodative facility ranged between 2.50 to 14.00 cpm (8.60 cpm; SD=2.67). Mean amplitude of accommodation was 12.20 cm (SD 2.86). Motor skills scores were evaluated as standard component scores for the three sections and total scores of the M-ABC. Regression analysis revealed that accommodative facility explained 10% of variance in the total M-ABC score. Examining the three test components, revealed that accommodative facility explained 16% of variance in the balance score, and vergence facility explained 10% of variance in the manual dexterity. In contrast, aiming was not significantly associated with any measure of visual function. Preliminary findings from this study highlight a significant link between aspects of visual function and fine motor skills in typically developing children and serve as a key element in examining the association between vision and fine motor performance in children with DCD.

Elizaveta Yakubovskaya	eyakub@my.yorku.ca	Hamidreza Ramezanpour, Kohitij Kar	Robustness of object position representation across spatial and temporal context in the macaque inferior temporal cortex	Understanding visual scenes requires concurrent knowledge about the identity and position of objects. While prior studies have investigated the representation of object identity information along the ventral visual pathway, culminating at the inferior temporal (IT) cortex in primates, the mechanisms of object position representation within IT still need to be further explored. The common functional segregation of the primate ventral and dorsal pathways as the "what" and the "where" pathways suggest that information regarding object position must be available in the dorsal pathway. However, interestingly, Hong et al. (2016) provided evidence for the hypothesis that the population activity in the IT cortex (ventral stream) represents an object's spatial position in a behaviorally relevant manner. In this study, we further stress-test this hypothesis and estimate the robustness of the IT population activity in representing explicit object position information during varying spatial and temporal contexts. We performed large-scale neural recordings using chronic multi-electrodes (Utah arrays) in 2 rhesus macaques (~200 sites) to measure the neural responses across the IT cortex. We trained the monkeys to passively fixate on a monitor screen while we presented images at their central field of view (~8 deg) for 100 ms each. We presented 600 images while systematically changing the spatial context in which the objects are embedded. Our imageset consisted of 60 base images (full context) and nine other variations of each image (with the object in the same position but varying backgrounds). The background variations included incongruent context, no context, no object, blurred context, blurred object, blurred incongruent context, minimum context, jigsaw context, and textured context. We leveraged a motion adaptation paradigm to test the effect of temporal context on object position representations. We presented an oriented grating stimulus moving in one of four directions (for 600 ms), immediately followed by a base image (containing one of eight objects in an uncorrelated background). A visual inspection of our stimuli indicates a significant motion after-effect. We plan to quantify this behavioral phenomenon with psychophysics experiments in the future. While preliminary results suggest that the IT cortex robustly represents object position under varying spatial contexts, a more detailed data analysis will be necessary to quantify the changes. This research is pivotal in establishing the correlative link between distributed IT population activity and object position perception, expanding our comprehension of the ventral visual pathway's functional role, and laying the groundwork for the next generation of primate vision models.

Emily D'Alessandro	emdal10@my.yorku.ca	Danai Kokkinopoulou, Jasmine Mosavi, Karolina A. Bearss, Ashkan Karimi, Rachel Bar, Joseph FX DeSouza	Dance improves affective state in individuals with Parkinson's Disease	Parkinson's disease (PD) is a neurodegenerative disease marked by the death of dopaminergic neurons in the substantia nigra. While motor symptoms are the focus of neurorehabilitation for individuals with PD, non-motor symptoms (NMS) are often considered as precursors and can have debilitating effects on the individual's affective state. Neurorehabilitative methods that can target NMS is relevant to improve these individuals' quality of life. An fMRI study by DeSouza et al. (2022) demonstrated a reduction in depression scores using GDS, correlated with reduced BOLD signals in the subcallosal cingulate gyrus (SCG) over an 8-month period of participation in dance classes. Our aim is to demonstrate that dance improves symptoms of both generalized anxiety disorder (GAD) and major depressive disorder (MDD) through changes in the participants PANAS-X scores and fMRI activity before and after Dance for PD classes, collected from 2013-2019. The PANAS-X affect categories were divided based on DSM-IV diagnostic criteria for GAD and MDD. The fMRI data collected shows the BOLD signal of participants visualizing their learned dance in the scanner. Brain regions of interest for GAD are the orbitofrontal cortex (OFC), bilateral medial frontal gyrus and superior frontal gyrus, the middle temporal gyrus, inferior parietal lobe and the amygdala and regions of interest for MDD are the medial and dorsolateral prefrontal cortex, OFC, SCG, supplementary motor area, and the inferior and superior parietal lobe. We aim to see a BOLD reduction in activity in these areas at the end of the 8-month dance period, along with an improvement in the affective scores.
Faruq Afolabi	gbolayo13@gmail.com	Xue Teng, Rob Allison, Laurie Wilcox	Understanding Affordances and the Effects of Distortions in Shared Mixed Reality Environments	There is growing interest in Augmented and Virtual Reality (AR/VR), with attention on shared interactions outside the context of games. Effective interpersonal interactions give users the ability to communicate more effectively and collaborate on certain tasks that may be limited by how we use our current technologies. Compared to their real physical counterparts, outputs from AR/VR tend to be subject to distortions, and little is known about how these distortions affect user experiences. The problem also extends to the case of shared assets and experiences where users are expected to build on their shared perceptions of their environment to interact with each other and achieve their goals, in these cases, it is unclear how users come up with an understanding of their shared assets and environments, and how those distortions affect said understanding. To better understand this, using a shape sorting game and the Unity3D game engine, we designed a simple object posting motor task, where one participant holds a cube, and another carries out the posting task, the task was designed to be executed either completely physically or virtually using only the physical and virtual assets respectively or partially using both physical and virtual assets, where we make use of AR/VR headsets to display the virtual perspectives. By carrying out user studies, we will be able to observe how the users interact and come up with their shared perceptual understanding of the presented assets and environment. Additionally, by varying the perspective each participant carries out their part of the task, i.e., physical vs virtual, we can identify how distortions affect users and come up with ways to mitigate those effects. The research aims to better inform the creation of shared experiences in Mixed Realities ensuring consistency for different users, improving experiences, and furthering the adoption of these technologies.

Gaelle N. Luabeya	gaellenl@yorku.ca	Ada Le, Erez Freud, Simona Monaco, J. Douglas Crawford	Cortical Integration and Functional Connectivity in Reach/Grasp fMRI Study	Daily tasks, such as picking up a cup, require the integration of two successive movements: an initial reach toward the target and a final grip around the cup. While the cortical mechanisms for reach and grasp were explored extensively, the mechanisms that integrate these components are not yet established. Here, we used functional magnetic resonance imaging to investigate which brain areas are involved in integrating information about object location and grasp orientation into the movement plan using a cue-separation paradigm. Participants grasped vertically or horizontally a cubic object presented to the left or the right of their body midline. Two successive cues preceded the grasping movement onset: a visual cue of the target location (L) and an auditory cue of the target orientation (O). In the O-L condition, the orientation cue was followed by the location cue; in the L-O condition, the location cue was followed by the orientation cue. A delay period followed each cue presentation. Whereas the first delay only required participants to remember one cue, the second delay required participants to remember two cues and integrate them as they prepared to initiate the reach-and-grasp movement. We conducted preliminary analysis on the conjunction analysis to assess which areas showed a greater response during the second than the first delay regardless of the order the sensory information was presented. We found increased activities in the left supplementary motor Area (SMA) and the bilateral Early Visual Cortex (EVC). Suggesting that not only are motor-related regions like SMA involved in action preparation, but the EVC also partakes in the final motor planning. This could indicate a top-down processing where higher-order motor regions interact with lower-order visual areas to gather information relevant to action preparation. Additional analysis will be conducted to look at the functional connectivity in these regions and other brain regions.
Hamid Ramezanpour	hamidram@yorku.ca	Sachi Sanghavi, Kohitij Kar	Selective Modulation of Visual Attention and Response Inhibition by Transcranial Ultrasound Stimulation of Pulvinar and Globus Pallidus Internus	Thalamus is traditionally seen as a relay for transmitting information to the cortex without active participation. Recent animal studies have challenged this view by showing that manipulations of the thalamus affect executive functions. In the past, it was not feasible to causally study potential roles of the human thalamus in executive functions as traditional non-invasive brain stimulation methods could not target this structure. Leveraging transcranial ultrasound stimulation (TUS), which has now made it possible to target deep brain areas, here we test the hypothesis that the human pulvinar, a posterior thalamic nucleus, plays an active role in the control of visual attention. To establish a comprehensive framework for interpreting our findings, we included a control paradigm involving the globus pallidus internus—a basal ganglia region implicated in response inhibition—and a stop signal task. This approach allowed us to discern between specific effects on attention and inhibition, thus unveiling the unique contributions of these regions. We used BabelBrain (a Python-based application) to plan an optimal trajectory for sonication, compensate for ultrasound losses due to skull barrier, and ensure the acoustic intensities and thermal rise in individual subjects are within the safety range. Our results revealed a double dissociation of effects: TUS of the pulvinar enhances visual attention without influencing response inhibition, whereas TUS of the globus pallidus internus selectively impairs response inhibition, leaving visual attention unaffected. This dissociation highlights the distinct cognitive functions of the pulvinar and the globus pallidus internus, while establishing a causal link between pulvinar activity and attentional modulation. Moreover, our study demonstrates that TUS has great potential to causally investigate functions of deep brain areas that cannot be targeted by other noninvasive brain stimulation techniques.

Hamid Ramezanpour	hamidram@yorku.ca	Ghazaleh Darmani, Regina Annirood, Can Sarica, Talyta Grippe, Artur Vetkas, Andres Lozano, Samuel Pichardo, Robert Chen	Leveraging computational and animal models of vision to probe atypical emotion recognition in autism	Recognizing others' emotions based on facial expressions is a core component of human social interactions. Previous studies (Wang and Adolphs 2017) have suggested that autistic individuals show differences in their facial emotion recognition compared to neurotypical adults. What are the neural mechanisms that account for these observed differences? Here we lay the groundwork for a new approach combining cutting-edge computational and empirical non-human primate work to test theories of atypical facial emotion recognition in autistic adults. In a recent study, the author(s) observed that artificial neural network (ANN) models of vision developed to achieve a myriad of visual objectives (e.g., object, emotion and face identification) could be fine-tuned to perform facial emotion judgments. Interestingly, the ANNs' image-level behavioral patterns better matched the neurotypical subjects' compared to autistic adults. This behavioral mismatch was most remarkable when the ANN behavior was constructed from units that correspond to the primate inferior temporal (IT) cortex. Here we directly test these two predictions in the rhesus macaques. First, we trained two macaques to perform a binary facial emotion (happy vs. fearful) discrimination task. Consistent with ANN predictions, the macaque image-level behavioral patterns better matched the behavior obtained in human Controls than in autistic individuals. Second, we implanted multi-electrode arrays in the IT cortex of two macaques and performed large-scale neural recordings while they fixated on images (used in the Wang and Adolphs study). Using the recorded neural multiunit spiking activity, we built regression models (165 IT-based models tested) to predict facial emotion ground truth ("level of happiness") on held-out images. Consistent with ANN-IT predictions, macaque IT population decodes of facial emotions better matched the neurotypical behavior compared to autistic individuals. Our results, therefore, establish the rhesus macaque as an appropriate species to further probe the neurobehavioral markers with ANN-guided hypotheses and experiment design.

Hamidreza Dastmalchi	hrd@yorku.ca	Seyed Nima Tayarani Bathaie, Aijun An	Incorporating GAN and Codebook Priors for Robust Blind Face Restoration	Face images captured in real-world scenarios often undergo multiple types of degradation, including low resolution, noise, blurring, and compression artifacts. Consequently, Blind Face Restoration (BFR) has garnered substantial interest, focusing on the task of reconstructing a corresponding high-quality (HQ) face image from a low-quality (LQ) input. BFR is an ill-posed inverse problem because multiple HQ images can result in the same LQ image due to loss of information, requiring prior information to limit the solution space. Current deep learning-based BFR approaches often rely on geometric and generative facial priors to constrain the solution space. However, they often struggle to maintain quality and fidelity when confronted with high-degradation scenarios. In our research, we propose a two-stage restoration approach that successfully harnesses the strengths of generative facial priors and codebook priors to overcome these limitations. In the first stage, we introduced the Generative Prior Restorer (GPR) network to reconstruct an initially restored image. This GPR network incorporates a deep Autoencoder with a StyleGAN generative model. The Autoencoder generates corresponding latent vectors and multi-resolution feature maps for the StyleGAN network to synthesize the initial restored image. Simultaneously, a lightweight Style-to-ID network introduces an identity-in-style loss, encouraging the inclusion of identity information within the style vectors. In the second stage, we employed a two-component framework called the Codebook Prior Restorer (CPR), which consists of a Vector Quantized AutoEncoder (VQAE) and a Feature Transformation Module (FTM). The VQAE network works to mitigate remaining distortion, while the FTM network focuses on enhancing fidelity, collectively enhancing the quality of the restoration. Extensive experiments conducted across diverse datasets confirm the superior performance of the proposed method in terms of quality and fidelity while also demonstrating robustness across a spectrum of degradation levels.

Harrison Ritz	hritz@princeton.edu	Jonathan Pillow, Jonathan Cohen	Task Preparation is Reflected in Neural State Space Dynamics	We can flexibly reconfigure our neural information processing to accommodate a wide array of different tasks. These task switches are increasingly being modeled as dynamical transition between task representations, however the empirical evidence informing these theories is controversial or incomplete. Here, we explore how the putative cognitive dynamics during task switching are reproduced in whole-brain neural dynamics. To study whole-brain neural dynamics, we re-analyzed a recent EEG experiment on task-switching by Hall-McMaster and colleagues (2019). Before each trial, human participants were cued to whether they would need to respond to the color versus the shape of a compound stimulus (50% switch rate). Borrowing emerging methods from systems neuroscience, we fit 'latent state space' models to whole-brain EEG activity during this pre-trial epoch. These models infer the linear recurrent neural dynamics and task encoding that gives rise to multivariate timeseries of electrode voltages. Fitting this model to trial-level EEG (N=30), we found that our approach was highly accurate at predicting EEG activity on single held-out trials. We also found that our state space models extended traditional multivariate encoding analyses by revealing rich interactions between task encoding and system recurrence. This experiment provides preliminary support for the dynamical systems approach to modeling task reconfiguration.Moreover, it provides a new empirical basis for developing richer process models of how our brains support cognitive flexibility.

Harshitha Koppisetty	harshithakoppisetty13@gmail.com	Robert Allision, Laurie Wilcox	Perception of Materials in Virtual Reality based on their Audiovisual Properties	Human perception of the world and the various objects in it requires integration of information from multiple senses. To replicate this experience in a virtual environment, ideally, stimuli for all senses would be provided. Current research focuses largely on visual and auditory stimuli. Previous studies on the perception of materials have shown that when the visual quality is low or the material properties are visually ambiguous, the auditory properties help bridge the gap and even reduce the amount of time needed to identify them. This suggests that humans are sensitive to the agreement between the two sensory inputs. But what would happen if the auditory properties conflicted with an object's visual properties? Would one sense be treated as more important than the other? Would this conflict give rise to any audiovisual illusions like the McGurk effect? And would there be a difference if the person is merely observing the object or interacting with it themselves? In this study, a Virtual Reality headset, the HTC VIVE Pro 1, was used to present a virtual target object that was struck with a rod, and an associated impact sound was played. The participant's task was to classify the material of the object. There were four materials to choose from: glass, metal, plastic, and wood. The associated impact sounds for these materials were carefully recorded under identical conditions in a controlled environment. Sixteen conditions were created by combining each sound with all materials. Each condition was repeated 20 times, for a total of 320 trials. To study the effect of participant interaction on their perception of a material, in half the trials, the participant observed an agent strike the block, and in the remaining trials the participant struck the object themselves using a virtual rod attached to the hand controller. The time it takes for them to classify was also recorded. Preliminary data from 9 participants shows that most participants tend to classify the target based on the auditory properties, with fewer participants tending towards the visual information. So far there appears to be no difference in classification accuracy between the agent interaction trials and the self-interaction ones, there is no major change in the classification itself. However, when observing an agent generating the sound, decision times were longer. Finally, in one of the sixteen conditions, we observe a McGurk-like phenomenon. Here, the target's texture was wooden and had a plastic impact sound and the participants chose to classify this target as metal more often than any other material. To conclude, our results suggest that the auditory signal is the strongest determinant of the material when conflicting stimuli are presented. Additional data and analysis will be needed to confirm the strength and to understand the source of the McGurk phenomenon seen here.
Helio Perroni Filho	helio@yorku.ca	Mohammad Akhavan, Nizwa Javed, Arian Haghparast, James Elder	AirChair: Smart Wheelchair Convoys for Airports	We present a platform for hands-free steering of a wheelchair convoy moving in single file. The first wheelchair in the convoy follows a human guide travelling on foot, the second follows the first and so on. Each wheelchair is controlled by its own visual tracking and navigation stack, and is able to independently dodge or stop for obstacles in order to avoid collisions. A mobile control device supporting a wireless messaging mechanism allows the guide to manage the convoy, and the wheelchairs to send alerts to the guide and each other. Real world experiments in moderately crowed environments demonstrate the architecture's effectiveness, while suggesting directions for further development.

Hongyi Guo	hguo06@yorku.ca	Alexander Schütz, Robert S. Allison	Assimilation of optic flow in saccadic eye movements	Most studies on oculomotor control are performed with stationary observers. However, when humans locomote the direction of features in the environment changes, creating an optic flow field. When executing a saccade during locomotion, the position of the saccade target changes during the saccade, so ideally our visual system should adjust the saccade parameters differently in different optic flow fields. We asked how optic flow affects the parameters of saccades to stationary targets. Each trial began with observers fixating a dot at the centre of a screen. Shortly after fixation, one of a variety of background stimuli were presented, including radial and rotational optic flow fields and a random dot kinematogram without coherent optic flow. A saccade target appeared at 10 degrees either to the left or to the right, and observers were asked to saccade to the target. As soon as the saccade was detected, the stimuli disappeared, in order to avoid feedback. Latencies and amplitudes of saccades were analyzed. The results showed a strong assimilation effect of optic flow, such that saccade amplitudes were biased in the direction of the optic flow. Moreover, saccade latencies were shorter in optic flow fields. These findings indicate that the oculomotor system partially accounts for the expected displacement due to optic flow in the planning of saccadic eye movements and are consistent with the assimilation of background motion in smooth pursuit eye movements.
Jason J. Yu	jjyu@yorku.ca	Fereshteh Forghani, Konstantinos G. Derpanis, Marcus A. Brubaker	Long-Term Photometric Consistent Novel View Synthesis with Diffusion Models	Novel view synthesis from a single input image is a challenging task, where the goal is to generate a new view of a scene from a desired camera pose that may be separated by a large motion. The highly uncertain nature of this synthesis task due to unobserved elements within the scene (i.e., occlusion) and outside the field-of-view makes the use of generative models appealing to capture the variety of possible outputs. In this paper, we propose a novel generative model which is capable of producing a sequence of photorealistic images consistent with a specified sequence and a single starting image. Our approach is centred on an autoregressive conditional diffusion-based model capable of interpolating visible scene elements and extrapolating unobserved regions in a view and geometry consistent manner. Conditioning is limited to an image capturing a single camera view and the (relative) pose of the new camera view. To measure the consistency over a sequence of generated views, we introduce a new metric, the thresholded symmetric epipolar distance (TSED), to measure the number of consistent frame pairs in a sequence. While previous methods have been shown to produce high quality images and consistent semantics across pairs of views, we show empirically with our metric that they are often in consistent with the desired camera poses. In contrast, we demonstrate that our method produces both photorealistic and view-consistent imagery.

Jaykishan Patel	jaykishanpatel96@gmail.com	Alban Flachot; Javier Vazquez-Corral; David H. Brainard; Thomas S. A. Wallis; Marcus A. Brubaker; Richard F. Murray	A deep convolutional neural network trained to infer surface reflectance is deceived by mid-level lightness illusions	A long-standing view is that lightness illusions are by-products of strategies employed by the visual system to stabilize its perceptual representation of surface reflectance against changes in illumination. Computationally, one such strategy is to infer reflectance from the retinal image, and to base the lightness percept on this inference. CNNs trained to infer reflectance from images have proven successful at solving this problem under limited conditions. To evaluate whether these CNNs provide suitable starting points for computational models of human lightness perception, we tested a state-of-the-art CNN on several lightness illusions, and compared its behaviour to prior measurements of human performance. We trained a CNN (Yu & Smith, 2019) to infer reflectance from luminance images. The network had a 30-layer hourglass architecture with skip connections. We trained the network via supervised learning on 100K images, rendered in Blender, each showing randomly placed geometric objects (surfaces, cubes, tori, etc.), with random Lambertian reflectance patterns (solid, Voronoi, or low-pass noise), under randomized point+ambient lighting. The renderer also provided the ground-truth reflectance images required for training. After training, we applied the network to several visual illusions. These included the argyle, Koffka-Adelson, snake, White's, checkerboard assimilation, and simultaneous contrast illusions, along with their controls where appropriate. The CNN correctly predicted larger illusions in the argyle, Koffka-Adelson, and snake images than in their controls. It also correctly predicted an assimilation effect in White's illusion. It did not, however, account for the checkerboard assimilation or simultaneous contrast effects. These results are consistent with the view that at least some lightness phenomena are by-products of a rational approach to inferring stable representations of physical properties from intrinsically ambiguous retinal images. Furthermore, they suggest that CNN models may be a promising starting point for new models of human lightness perception.
Jennifer Lin	linj68@yorku.ca	Hongying Wang, Saihong Sun, Xiaogang Yan, John Douglas Crawford	Influence of a visual landmark shift on memory-guided reaching in monkeys	The brain uses various sources of visual information, including both egocentric cues (e.g., object location relative to the eye) and allocentric (e.g., object location relative to other visual landmarks) to aim movements. It has been shown that humans optimally weigh egocentric and allocentric (landmark) cues when pointing (Bryne & Crawford 2010) but it is not known if monkeys do this. The main purpose of this study is to determine the influence of allocentric cue shifts on reaching responses in non-human primates. In order to do this, reach and gaze data were collected from one female Macaca mulatta monkey (ML) trained to perform a memory guided reaching task. The hand was initially placed at 1 of 3 varying locations of a waist level LED bar while gaze fixated centrally. A landmark (4 'dots' spaced 10° apart forming the corners of a virtual square) was then presented at 1 of 15 locations on a touch screen after a delay. A visual target then appeared transiently at a variable location within or outside this virtual square, followed by a visual mask. After the mask, the landmark either reappeared at the same location (stable landmark condition) or shifted by 8 degrees in one of 8 directions (landmark shift condition). The fixation light then extinguished, signalling a reach to the target. 'No-landmark' controls were the same, but without the landmark. Compared to gaze responses, reach had lower variance and decreased reaching error in the presence of a stable landmark. In the landmark shift condition, reaches shifted partially (mean=29%) with the landmark, Overall, this data suggests that the monkey is influenced by visual landmarks when reaching to remembered target in a similar way as humans.

Jiali Song	jiali.song@utoronto.ca	Anureet Jeji, Avery Chua, Cristeidy Gonzalez, Benjamin Wolfe	Are unreliable car warnings still helpful to the driver? The effect of auditory cues on road hazard detection in dynamic road scenes.	Road hazards can include everything from wandering deer to other vehicles, and to safely evade these hazards, drivers need to know where and when they appear. A warning system that can buy drivers more time to respond can be the difference between experiencing a collision or evading one. In a previous experiment, we found that an alerting auditory cue indicating an impending hazard speeded hazard detection by approximately 200ms (Song, Chua, et al., 2023). However, cues always accurately indicated the presence of a hazard in that experiment, whereas computerized systems that produce these warnings in real cars are unlikely to be perfect. In the current study, we asked how reliable auditory alerting cues need to be for drivers to make use of them, and how well drivers can tolerate wrong or misleading cues. To measure hazard detection performance, participants watched a brief segment of front-facing dashcam video taken from Road Hazard Stimuli, a dataset compiled from crowd-sourced dashcam footage (Song, Wolfe, et al., 2023) on each trial. Forty-eight licensed drivers were asked to press a key to indicate the presence of a hazard as soon as it appeared in each video. If they indicated the presence of a hazard, participants were then asked to locate the hazard (left or right side of the video). If participants did not see a hazard, they were asked to indicate the absence of a hazard after the end of the video. Across all trials, half of the videos contained hazards, and auditory cues were presented during half of all videos. Cues could be valid (presented during a video with a hazard, or not presented during a video with no hazard) or invalid (not presented during a video with a hazard, or presented during a video with no hazard). To examine the effect of cue reliability on hazard detection performance, we varied the overall proportion of valid cues between two groups; the high-reliability group received 80% valid cues, and the low-reliability group received 50% valid cues. We found small, non-significant effects of cue validity and reliability on both the speed and accuracy of hazard detection, suggesting that participants were able to ignore the cues in these conditions. However, for videos that did not contain hazards, the presence of invalid cues speeded correct responses compared to valid trials without cues ($F(1, 47) = 82.33$, $p < 0.001$). Hazard detection and localization accuracy remained approximately 90%, across all groups and conditions. These results suggest that, when hazard detection performance is higher than the proportion of valid cues, drivers coped with invalid cues by ignoring them altogether when responding to hazards. Furthermore, the speeded reaction time found for hazard absent videos with invalid cues are consistent with the idea that warnings temporarily increased alertness in drivers in the absence of hazards. These findings suggest that warnings for hazards that drivers can readily detect must have an accuracy of at least 80% to be useful.

Jinani Sooriyaarachchi	jinani.sooriyaarachchi@mail.mcgill.ca	Curtis L. Baker, Dr. Chang'an Zhan	Cortical State Effects Vary Across Different Neuronal Subclasses in the Primary Visual Cortex	Primary visual cortex neurons respond selectively to features of visual stimuli such as spatial frequency and orientation. Linear-nonlinear models containing a single receptive field filter can be used to model phase-sensitive simple cells. However, more elaborate models having a combination of nonlinear subunits are required to model phase-invariant complex cells. Further, cortical neurons' trial-to-trial response variability can make the parameter estimation of these models more difficult. This variability is thought to be caused by cortical state fluctuations and its effects are poorly understood, potentially compromising RF estimation. Here we utilized extracellular recordings from the primary visual cortex of anesthetized, paralyzed cats with 32-channel NeuroNexus probes in response to rapid sequences of natural images (Talebi and Baker, 2012). We developed a simple compact convolutional neural network method to estimate receptive field models for both simple and complex visual cortex cells from their responses to natural images (Nguyen et al, bioRxiv 2023). A single model parameter from a parameterized rectifier unit in the model determined the simple vs. complex nature of the receptive field. We further analyzed the interaction between a neuron's spiking response and cortical state inferred from local field potentials (LFPs) and multi-unit activities (MUAs), to obtain better quantitative estimates of neuronal RFs. For this purpose, we extended our model architecture by incorporating LFP and MUA-driven filtering pathways in parallel with the aforementioned stimulus-driven receptive field pathway. The model parameters were estimated with an iterative regression algorithm with L2 regularization, using TensorFlow and Keras on a training dataset, along with a validation dataset for regularization. The variance accounted for (VAF) was used to evaluate the model's predictive performance on a hold-back test dataset. Incorporating brain state effects provided significant improvements in VAF measures of predictive performance despite the additional model parameters. These improvements were dependent on the type of neuron, e.g. occuring more in complex than in simple cells, more in those with oriented than non- oriented RFs. These findings highlight the importance of incorporating the trial-to-trial variability due to brain state variations in studies of sensory coding at the single neuron level, and indicate that the nature of brain state variations may vary substantially across different types of neurons.

Joey Wong	joeywong@my.yorku.ca	Shital Desai	Enhancing Hand Function in Older Stroke Patients: Designing Instrumented Objects for Home-Based Stroke Rehabilitation	This research will address the need for more engaging home-based stroke rehabilitation interventions for older stroke patients. It aims to enhance patient motivation and adherence to stroke rehabilitation exercises through the design of an instrumented object that focuses on alleviating paresis of the hand. The current methods of stroke rehabilitation can be inaccessible as those living in rural communities can have difficulty accessing rehabilitation clinics. Exercises that are prescribed by healthcare professionals can be repetitive and can limit the transfer of skills during the patient's transition from clinical rehabilitation to home-based care. Often, tools developed for rehabilitation fail to incorporate input from various stakeholders which can make it difficult for patients to adopt to daily life. These factors can contribute greatly to a patient's lack of adherence to rehabilitative interventions. Patients have expressed their interest in exercises that connect rehabilitation to activities in their daily lives as they become more meaningful and enjoyable. The instrumented object offers exercises that can be transferred to daily tasks and will allow patients to enhance their grasping, twisting, stimulating motor skills and hand dexterity. The patient will be able to complete rehabilitative exercises while engaging in a creative activity that involves using smaller household objects to create new images. The object is also equipped with pressure sensors that have the potential to track patient progress over time and provide real-time auditory and haptic feedback. By developing a better understanding of the experiences of patients and stakeholders through their rehabilitative process, it further identifies opportunities for technology to support their rehabilitative journey in remote or home-based settings in more engaging and effective ways.
John Kennedy	kennedy@utsc.utoronto.ca	Marta Wnuczko	Mirror-image Pointing Illusion: Visible Arms, Occlusion And Overshooting	Observers are asked to use their right arm to point to their left shoulder's mirror-image. Unawares, they use their mirror-image finger to occlude the target, not their real finger. The real finger overshoots to the left of the target Besides normal mirrors, the illusion occurs with (A) a mirror to the side of the person pointing (B) a mirror tilted at 45º (C) a mirror lying on the floor (D) two mirrors forming a 90º concave corner and reversing left-right (E) a monitor image also left-right reversed. Prior to pointing, the illusion requires the pointing arm to be visible. The illusion is not due to mirror neurones.
Julie Ouerfelli-Ethier	j.o.ethier@hotmail.com	Pisella, Laure, & Khan, Aarlenne Z.	Spatiotemporal competition resolution during anti-saccades	Anti-saccade are eye movements directed in the opposite direction of a visual target. Because the visual target and the saccade goal are decoupled, it has been suggested that the competition between those two locations need to be resolved for the execution of a correct anti-saccade. However, it remains unclear how this competition is resolved spatially and temporally. To examine this, we tested 14 participants on a pro-saccade paradigm as well as three different spatial configurations of anti-saccades: 90° away across hemifields, 90° away within the same hemifield, and 180° away. We measured anti-saccade metrics, such as error rates, amplitudes and saccade endpoints, across different eccentricities (4° and 7°) as a function of reaction times. We observed a dynamic modulation of error rates as well as endpoints as a function of reaction times. Moreover, compared to pro-saccade endpoints, we found a spatial bias toward the visual target in the endpoints of anti-saccades across all anti-saccade paradigms. This bias was also more pronounced for the large 7° eccentricity condition compared to the 4° condition. Taken together, these results point to saccade goal selection and execution processes that vary dynamically across time and space during anti-saccades.

Khushbu Y Patel	khushbup@my.yorku.ca	Laurie M. Wilcox, Laurence T. Maloney, Krista A. Ehinger, Jaykishan Y. Patel, Emma Wiedenmann, Richard F. Murray	Lightness constancy in real and virtual environments	Virtual reality (VR) technology is being used in an increasing number of applications where performance is critical, including research applications where we wish to use VR to investigate performance in real scenes. However, previous work has shown that people often perceive surface properties differently in real and virtual environments, and so it is important to measure and characterize these differences. In the first of two experiments, we evaluated how well virtual platforms support realistic lightness perception. We measured lightness constancy on a custom-built physical apparatus, in VR, and on a 2D flat-panel display. Twelve observers participated in three conditions. In the physical condition, observers performed a lightness matching task where adjustable reflectance patches were visible through two 2-degree apertures. On each trial, the reference aperture was set to one of three reflectances (0.18, 0.39, 0.55). The match aperture had one of five illumination levels, between 1.25 and 3.05 times the illuminance at the reference aperture. Observers adjusted the reflectance at the match aperture until it appeared to match the reflectance at the reference aperture. In the VR condition, observers viewed an apparatus and room that approximately replicated the physical condition, rendered in Unity, on an Oculus Rift S headset. In the flat-panel condition, observers viewed an apparatus like the one in the physical condition, rendered on an LCD screen using Unity. Thouless ratios (a measure of lightness constancy) in the physical condition (mean 0.87) were significantly higher than in the flat-panel condition (0.79), indicating better constancy, but were not significantly different from the VR condition (0.83). We also found that variability across observers was significantly greater in VR than in the physical environment. Thus, in the simple scenes investigated here, lightness constancy was similar in physical and VR environments. In the second experiment, we evaluated lightness constancy using a more complex task that required the observers to compensate for the 3D orientation of test patches relative to a light source. Twelve observers performed a lightness matching task on a custom-built physical apparatus and in VR. On each trial, a reference patch was set to one of two reflectances (0.38, 0.47) and one of seven 3D orientations relative to a point light source. Observers adjusted the reflectance of a grey match patch to match the perceived grey of the reference patch. Unlike in the first experiment, here Thouless ratios were markedly and significantly better in real scenes (mean 0.44) than in VR (0.22), with relatively poor constancy in both conditions. Thus, with this more complex lightness matching task, lightness constancy was much weaker, and VR performance was notably inferior to that in the physical environment. In conclusion, we find significant task-dependent differences between lightness constancy in real and virtual environments. These discrepancies should be considered when developing applications where realistic performance is critical.

Khushi Patel	khu.patel@utoronto.ca	Laurie M. Wilcox, Laurence T. Maloney, Krista A. Ehinger, Jaykishan Y. Patel, Emma Wiedenmann, Richard F. Murray	Effects of different highlighting styles on reading speed and comprehension	In a digital format, reading rulers (i.e., underlining the current line being read) have been found to improve reading speed for both dyslexic and non-dyslexic readers. However, the mechanisms underlying this speed increase, and whether this increase in reading speed is accompanied by changes to comprehension are unknown. It is well-established that reducing visual crowding and improving parafoveal processing (wherein the reader can read the next word on the line, adjacent to the word they are currently looking at) can improve reading speed. To better understand the mechanisms by which these rulers may impact reading speed and comprehension, we examined a range of underlining and highlighting conditions similar to what readers with dyslexia find helpful, but manipulated the appearance of our interventions to help reveal how the visual appearance of the intervention might impact crowding and therefore reading. Specifically, our study examined participants' gaze behaviour while they read 18 short passages under five different highlighting conditions: a fixed black underline under the line of text being read, a yellow underline under the line of text being read, a lightbox highlighter (a white "box" around the line being read, with a low-contrast overlay over the rest of the text), a drop shadow covering the unread sections of the passage (approximating the bottom edge of a desktop screen), and yellow highlighting over the line being read. Participants also answered five comprehension questions per passage to assess understanding. Reading speed (words per minute) and comprehension were calculated for each condition. Pilot data from 9 young adults (age 18-28) revealed no significant differences in reading speed and comprehension scores across the different highlighting conditions. Additionally, we found that, compared to other conditions, participants' preference for a specific highlighting condition did not correspond to faster reading speed in that condition. Eye tracking data showed that the average number and duration of fixations, and the average number of regressions (backward eye movements that allow rereading text to facilitate comprehension) were lowest in the black underline condition, and highest in the drop shadow condition. Based on these preliminary findings, a fixed black underline may increase reading speed, while the drop shadow condition may be slowing down readers, as the unread sections of the text are hidden under an opaque shadow. Our pilot work hints that line-by-line highlighting styles may influence eye movements while we read, and our existing paradigm can help to elucidate the underlying mechanisms of these interventions.

Krista Kelly	krista.kelly@uwaterloo.ca	Dorsa Mir Norouzi, Norah Nyangau, David Stager, Cynthia L Beauchamp, Prashanthi Giridhar	Ocular motor function and postural control during static balance in children with amblyopia	Purpose : Children with amblyopia score lower on balance tasks using standardized tests of motor ability (Kelly et al, IOVS, 2020), and have ocular motor dysfunction such as fixation instability (Kelly et al, Exp Eye Res, 2018). Postural control during static balance is tightly linked to ocular motor control. Here, we examine the extent to which amblyopia impacts postural stability, and the contribution of ocular motor dysfunction to performance. Methods : Thirty children diagnosed with amblyopia and 31 control children ages 4 – 14 years were enrolled. Children were fitted with an iPod installed with the NIH Toolbox® Standing Balance Test that uses the built-in accelerometer to measure anterior-posterior (A-P) sway, and with a wearable eye tracker (Tobii Glasses 2) to record eye movements. Children completed 4-5 standing balance poses in sequential order; eyes open on flat surface, eyes closed on flat surface, eyes open on foam surface, eyes closed on foam surface, eyes open on flat surface with feet tandem (7+ years only). Children were instructed to stand as steady as possible for 50 seconds with feet together and arms crossed on the chest. For eyes-open poses, children were asked to stare at the center of a fixation cross. Outcome measures were overall A-P sway standard score, eyes open/eyes closed ratios (a measure of reliance on vision), and average fixation duration and total number of fixations during each eyes-open pose. Results : Compared with controls, amblyopic children had more A-P sway (amblyopia, 85±11 vs control, 91±9, p=0.02), and relied less on vision while standing on the flat surface (ratio, 1.15±0.06 vs 1.31±0.05, p=0.03) and foam surface (ratio, 1.33±0.06 vs 1.65±0.08, p=0.002). During all eyes-open poses, amblyopic children also had shorter average fixation duration (flat, 341±31 vs 656±108 msec, p=0.009; foam, 290±22 vs 483±65 msec, p=0.009; tandem, 309±27 vs 505±75 msec, p=0.02), and more fixations except for tandem (flat, 110±7 vs 79±8, p=0.004; foam, 123±6 vs 93±8, p=0.003; tandem, 96±7 vs 87±9, p=0.43) compared to controls. Conclusions : Postural stability during static balance is impacted by amblyopia. Further, these results indicate a role of ocular motor dysfunction in poor postural stability. Determining causes of poor balance may inform the development of interventions to help children with amblyopia succeed.

Krista Mitchnick	kam13@yorku.ca	R. Shayna Rosenbaum, Boyer D. Winters	Object, tactile, and spatial oddity judgements are impaired in DG-compromised rats but enhanced in CA1-compromised rats	The hippocampus (HPC) is necessary for supporting episodic memory; however, evidence suggests that the HPC subregions differentially contribute to various underlying component processes of episodic memory. For example, the dentate gyrus (DG) appears to be necessary for the orthogonal representation of similar information (i.e., pattern separation) to facilitate precise encoding. Given that this computational process occurs as information enters the DG (i.e., during encoding), we reasoned that the DG might additionally be involved in non-mnemonic, perceptual discrimination. Moreover, although prominent theories of HPC function predict a domain-specific role of the DG in processing spatial scenes, it is possible that the DG plays a domain-general role in discriminating other stimulus types that form rich, episodic memories. Indeed, we recently demonstrated that a unique individual (BL) with selective bilateral DG lesions, was impaired in his ability to detect an odd object amongst three identical distractors (i.e., oddity task), compared to matched controls. In contrast, perirhinal cortex (PRh) has been shown to be necessary for successful domain-specific oddity judgments involving objects. Therefore, we assessed performance on visual, spatial, and tactile oddity tasks in male rats with dorsal DG lesions, PRh inactivation, or dorsal CA1 inactivation, the latter included as a negative control. Our results demonstrate that dDG-lesioned rats exhibited impairments on all three tasks at an intermediate level of difficulty, but spared performance on the easier level. Conversely, PRh-inactivation produced deficits only on the visual oddity task (easy version). Unexpectedly, dCA1 inactivation enhanced discrimination in all modalities (intermediate and hard versions). These results further support the domain-general involvement of the DG in difficult perceptual discriminations compared to the domain-specific (object identity) involvement of PRh in this process. Furthermore, our results demonstrate that perceptual discrimination can be enhanced by inactivation of the CA1 subregion, possibly due to reduced interference from retrieval processes. These findings add to a growing literature documenting a role for the HPC in perceptual processes.

Lina Musa	lmusa09@yorku.ca	Amirhossein Ghaderi, Ying Chen and J. Douglas Crawford	Hierarchical Functional Modularity of Brain Networks for Egocentric and Allocentric Memory-guided Reaching	The brain can encode targets for reaching in egocentric and/or allocentric reference frames (Byrne and Crawford 2010). The differences in the cortical activation of these two representations has been described (Chen et al., 2014; Neggers et al., 2006). For example, Chen et al. (2014) identified egocentric directional selectivity in dorsal brain areas (the parieto-frontal cortex) versus landmark-centered directional selectivity in ventral brain areas (inferior temporal gyrus and inferior occipital gyrus) during a delayed reach task. However, differences in the functional organization of brain networks have not been studied. Here, we performed a secondary analysis of the event-related fMRI task from Chen et al. (2014), to distinguish human brain networks involved in egocentric versus allocentric spatial representation of reach goals. Based on their previous univariate analysis we expected that the functional brain networks will differ, with increased hubness in ventral brain regions in the allocentric task. The paradigm consisted of three tasks with identical stimulus display but different instructions: egocentric reach (remember absolute target location), allocentric reach (remember target location relative to a visual landmark), and a nonspatial control, color report (report color of target). We performed a graph theoretical analysis (GTA) on time series data recorded during the entire trial, contrasting egocentric and allocentric data versus controls and each other. The fractal organization of the network modules was determined using agglomerative and divisive clustering approaches. Modularity maximization and consensus partitioning were used to identify the scale of modules that best delineate egocentric and allocentric brain networks. A comparison with resting state network parcellations showed poor overlap, with decreased modularity in resting state networks. Instead, the data were largely segregated into dorsal and ventral modules, with similar organization in egocentric vs. allocentric trials. However, the egocentric network demonstrated significantly stronger modularity in parietal modules. Brain regions that were important for connecting different modules in the allocentric task were found ventrally in the parahippocampal cortex and inferior extrastriate visual cortex, whereas in the egocentric task those hubs were found in the superior extrastriate visual cortex and the interparietal lobule. We conclude that egocentric versus landmark-guided reach utilizes largely overlapping cortical networks, but with significant differences consistent with a dorsal-ventral specialization.

Lynn K. A. Soerensen	lynnka@mit.edu	James J. DiCarlo, Kohitij Kar	Probing the effects of object category learning on the macaque inferior temporal cortex	Like humans, adult non-human primates can learn to categorize visual objects. Much prior work shows that individual neurons in the inferior temporal (IT) cortex, critical for visual object recognition, modestly increase their selectivity to objects from learned categories. How do these neural changes quantitatively link to behavioral performance gains (“learning”) on object discrimination tasks? While the field now has relatively accurate models of image-driven IT responses, we do not yet have a similar understanding of adult IT plasticity. To begin to address this, we measured changes across the IT cortex due to object category learning and asked how those quantitatively relate to behavior. We performed multi-electrode recordings in two groups of macaques (3 monkeys/group), while monkeys viewed naturalistic images (8 categories, 80 images/category, 100 ms). Before recording, one group (naive) was only trained to fixate passively on images; the other group (trained) also learned to discriminate multiple object categories via operant conditioning. For comparing both groups, we randomly sampled 53 reliable, visually responsive sites from each monkey to construct two pools of IT activity (corresponding to the trained and naïve monkeys respectively, 159 sites per pool, 300 repetitions of the random sampling). First, consistent with previous studies, we observed a significant increase in the object-category selectivity of the trained IT responses compared to the naïve IT. Next, this increase in selectivity yielded a small, but statistically significant increase in the IT-based linear decoding accuracy for the learned object categories as well as an increase in the categorical representational similarity. Lastly, these subtle changes in the trained IT representations resulted in better predictions of image-level behavioral error patterns compared to linear decodes from the naïve IT population responses. How do these concurrent changes in IT and behavior arise? We provide a systems-level perspective on this by casting the monkey's category training as an extension of contemporary artificial neural networks (ANNs), the leading models of the primate ventral stream. Interestingly, we observed that for various task-optimized ANNs (with different architectures, pre-training objectives, and category learning schemes), the model layer matched to untrained IT responses showed monkey-IT-like increases in category information after training. Akin to IT, specific ANN-IT representations were also more predictive of monkey behavior after training. In sum, we provide empirical evidence that category learning produces behaviorally relevant, modest changes in the macaque IT cortex, enhancing category information readout. Furthermore, we introduce a computational framework to simulate these changes, enabling us to formulate testable hypotheses about the representational reconfigurations induced by category learning.

Lynn Schmittwilken	L.Schmittwilken@rochester.edu	Paul Jolly; Michele Rucci	Foveal perimetry: Towards mapping visual sensitivity across the foveola	Perimetry is a widely established diagnostic tool to detect local sensitivity changes across the visual field. It is essential for diagnosing and monitoring various medical conditions affecting the optic nerve, retina, and visual pathway. Standard perimetry tests typically examine large portions of the visual field, such as the central 60 degrees (standard automated perimetry) or 5-10 degrees (microperimetry) with a coarse resolution. However, these methods do not allow mapping of visual function within the foveola, the region responsible for high acuity vision. The foveola only covers the centermost one degree of the visual field. Despite its small size, the foveola is vital for many daily behaviors and its importance is also reflected in a consistent neural over-representation throughout the visual hierarchy. In existing perimetry tests, the entire foveola is represented by a single data point. Thus, little is known about how and whether visual sensitivity varies within the foveola. Testing visual sensitivity across the foveola is technically challenging because incessant, small eye movements prevent testing adjacent spatial locations in isolation. Compensating these eye movements is difficult because their magnitude is comparable to the uncertainty in gaze localization of standard eye trackers. Here, we developed an approach to map visual sensitivity across the foveola. We use a high-precision digital dual Purkinje image eye tracker with real-time control of the retinal input through a custom apparatus for gaze-contingent display. This system enables accurate localization of the line of sight. Observers were asked to detect a brief probe (50 ms duration, 5x5 arcmin size) presented at one of 13 predefined spatial locations within the centermost one degree (diameter) around the preferred fixation locus. Probes were presented over a uniformly gray display (luminance: 10 cd/m2). Observers responded whether they detected the probe via a button press and received auditory feedback. We tested the right eye of all observers while their left eye was patched. To validate our approach, we used the method of constant stimuli to rigorously measure individual psychometric functions at each probe location. Preliminary data from five expert human observers indicates considerable inter-individual variability in foveal sensitivity. Across observers, peak sensitivity is shifted towards the right visual field with higher sensitivity along the horizontal meridian. Our results show that mapping sensitivity across the foveola is possible with the proposed approach. Supported by NIH grant EY018363

Maren Wehrheim	marenwehrheim@gmail.com	Na Yeon Kim, Ralph Adolphs, Kohitij Kar	Probing the link between dynamics of “face-selectivity” in macaque IT cortex and facial emotion discrimination behavior	Recognizing facial emotions constitutes a core component of human social interactions. Previous research has discovered several cortical sub-regions ("face patches"), primarily localized in the primate inferior temporal (IT) cortex, that respond relatively selectively to faces. However, one can build algorithms that decode faces from the responses of more heterogenous samples of IT neurons. Here we test the link between psychophysical (behavioral) facial emotion discrimination ability and decoding models constructed from neurons with varying "face-selectivity". Adult human participants (n=12) discriminated the emotions (happy vs. fearful) in 80 images (8 identities) presented for 100 ms. We generated a hypothesis space linking IT activity and facial emotion behavior using convolutional neural network(CNN) models of primate vision (adapted to perform emotion discrimination). The estimate of "face-selectivity" for each model-IT neuron was proportional to the difference in its response to faces vs. other objects. We determined the goodness of a linking algorithm by correlating the CNN-IT-based emotional intensity predictions and the patterns of human behavioral discrimination. Interestingly, we observed that some CNNs (e.g., ResNets) generated stronger predictions when decoding algorithms contained more "face-selective" units (100 units per decoder). However, for many other models (e.g., Cornet-S), any random sample of their IT units demonstrated equally predictive decodes. To ask which of these hypotheses is consistent with primate visual processing, we performed large-scale recordings in the IT cortex of two macaques (~200 sites). Preliminary results show that macaque IT-based decodes incorporating more "face-selective" neurons better predict facial emotion discrimination. Interestingly, the predictions' strength and dependence on "face-selectivity" only prevailed during early (70-110ms) IT responses. In contrast, facial identity decodes were strongest during the late (150-200ms) IT responses. The high "face-selectivity" and early (putatively feedforward) IT response-based decoding that our results reveal, facilitate discriminating among and constraining the development of the next-generation brain models of facial emotion processing.

Maria Orlando	morlando@yorku.ca	Alberto Umiltà, Federico Fornaciari, Elisa Ciaramelli, R. Shayna Rosenbaum	Examining the Effects of Real-World Experience on Lab-Based Scene Memory	Boundary extension (BE) is as an error in scene memory, such that participants retrieve details beyond the given boundaries of a scene image. Boundary contraction (BC) is the opposite effect, whereby participants retrieve less context within the boundaries of a given scene image. In the BE literature, there is variability in the types of stimuli that are used, how BE is tested, and the proposed mechanisms underlying the phenomenon. Some research supports the view that BE reflects (re)construction of the scene from an internal representation that was formed, whereas other research supports the view that BE (and BC) emerge from image-based properties, including the number of central objects and whether an object is pictured in close range or from a wider angle. Assessing the effects of prior knowledge and experience of a scene on this bias can help disentangle the role of visual perception and scene construction. The current study tested the influence of familiarity on scene recognition through the comparison of lab-based encoding of images of pre-experimentally familiar (real-world) places with images of unfamiliar places. Participants used a continuous rating scale to indicate how they perceived the boundaries of a test image relative to a previously studied image. There was a tendency for BC across both image conditions, with evidence of maintained, and an instance of greater, BC for familiar than unfamiliar scene images. Importantly, the lack of evidence for increased BE with greater familiarity favours an image-based theoretical account of BE and BC.
Marta Wnuczko	marta.wnuczko@guelphhumber.ca	John M. Kennedy, Selene Carboni	Directions yes, screens no: how perspective is used by an early-totally-blind man	Drawing progresses from copying shapes to copying directions (Kennedy). The advance is a function of drawing experience for the blind and the sighted. The later stage uses linear perspective, and requires two conditions (1) awareness of the directions of objects (which blind people can show by pointing), (2) treating the picture surface as a projection screen (Arnheim). S.M., a totally early blind man is a drawing novice. He points correctly in the directions of objects from his vantage point but in drawing he copies the shapes of single or multiple aspects of objects in plans and elevations. We suggest S.M. uses perspective in pointing but not in free drawing, because for him the screen idea is absent. We propose that drawing practice would lead to the use of the screen concept.

Maryum Khan	mkhan36@yorku.ca	Shanaathanan Modchalingam, Andrew King, Marius 't Hart, Denise Henriques	Effects of tool use and perturbation during motor adaptation on hand localization	Our brain has a remarkable capacity for learning movements and adapting them to accomplish a motor goal. In many adaptation studies, participants move in a 2D plane while their hand is represented by a cursor. When visual feedback of hand position is misaligned, people can quickly compensate for this perturbation, show persistent reach aftereffects, and even misestimate the location of the unseen hand in the direction of previous visual training. However, it is unknown how well this generalizes to real-world settings or to the tools we use every day. Here we will use immersive virtual reality to test if end-effector shifts are also observed in more naturalistic virtual reality environments and if they extend to tools as end effectors. In Study 1, we replicated our previous work where we found shifts in end-effector localization after adapting reach movements to a 30° visuomotor rotation of the hand, showing a similar magnitude of both shifts in unseen hand location and reach aftereffects following training to the perturbation in the VR environment. In the next condition (Study 2), we extend this paradigm to investigate how well people can adapt when aiming with a common tool, like a pen, and whether the tool location is also recalibrated. Participants will reach to the same targets using both a physical and virtual pen, whose movements will also deviated by 30°, and we will measure the extent that the unseen location of hand-held tool, as well as the hand (in separate trials) recalibrates with adaptation. Our results will provide insight into the adaptive processes involved when learning to wield tools in more complicated, realistic environments.
Matthew C. Macdonald-Dale	mattmd@yorku.ca	Stefania S. Moro, Peter J. Kohler, Jennifer K. E. Steeves	Retinotopic mapping of primary visual cortex of people with one eye	The surgical removal of one eye (unilateral eye 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. People who have had one eye removed in the first months of life present with several long-term morphological and functional adaptations in adult cortical and subcortical visual, auditory, and multisensory brain regions. In the current study we use functional magnetic resonance imaging (fMRI) to map spatial coding of visual input in topographically organized visual cortex in people who have had one eye surgically removed. We asked whether retinotopic mapping might be altered to reflect the reduction in input to the visual system and how the developmental age at which the eye is lost might play a role. A standard checkerboard stimulus was used to map population Receptive Fields (pRFs) in adult participants who had undergone monocular enucleation, as well as binocularly intact controls. For each voxel, pRF mapping identifies the visual field location and receptive field size that accounts for most variance in the BOLD response of each voxel. Preliminary findings suggest cortical magnification is within the same range for both hemispheres, in both binocular controls and people who have one eye. Preliminary findings also indicate similar retinotopic mapping in V1 in people with only one eye compared to binocular controls. This indicates that the loss of one eye results in little functional change in early visual cortex; despite the deafferentation of half of visual sensory input, retinotopic organization remains stable.

Matthew Kowal	m2kowal@yorku.ca	Mennatullah Siam, Md Amirul Islam, Neil D. B. Bruce, Richard P. Wildes, Konstantinos G. Derpanis	Quantifying and Learning Static vs. Dynamic Information in Deep Spatiotemporal Networks	There is limited understanding of the information captured by deep spatiotemporal models in their intermediate representations. For example, while evidence suggests that action recognition algorithms are heavily influenced by visual appearance in single frames, no quantitative methodology exists for evaluating such static bias in the latent representation compared to bias toward dynamics. We tackle this challenge by proposing an approach for quantifying the static and dynamic biases of any spatiotemporal model, and apply our approach to three tasks, action recognition, automatic video object segmentation (AVOS) and video instance segmentation (VIS). Our key findings are: (i) Most examined models are biased toward static information. (ii) Some datasets that are assumed to be biased toward dynamics are actually biased toward static information. (iii) Individual channels in an architecture can be biased toward static, dynamic or a combination of the two. (iv) Most models converge to their culminating biases in the first half of training. We then explore how these biases affect performance on dynamically biased datasets. For action recognition, we propose StaticDropout, a semantically guided dropout that debiases a model from static information toward dynamics. For AVOS, we design a better combination of fusion and cross connection layers compared with previous architectures.

Mohad Nasir	m24nasir@uwaterloo.ca	Xiaoxin Chen, Andrew Silva, Ewa Niechwiej-Szwedo	Investigating the Association between Sensory Eye Dominance and Fixation Stability	Introduction Fixation stability (FS), a measure of oculomotor control, refers to an individual's ability to precisely visual fixate on a target. Poor FS has been implicated in a variety of visual disorders, including amblyopia and macular degeneration. It may also impact functional behaviors like reading and hand-eye coordination in normal observers. Therefore, it is important to develop a better understanding of the factors influencing fixation stability. Here, we investigated whether one's preference to use a particular eye over the other, called sensory eye dominance, impacts one's fixation stability. Methods Twelve healthy controls (age = 25.0, SD 2.39 years; 6 males, 6 females) with normal or corrected to normal visual acuity and stereopsis of at least 40 arc seconds participated. Fixation stability was assessed using a 3deg fixation target presented at the center of the monitor. An EyeLink 2 recorded eye movements. Participants completed the fixation task when viewing binocularly and monocularly with the left eye and right eye. An infrared filter was used to cover the participant's eye during monocular viewing, which allowed binocular eye position recording even in monocular viewing conditions. Fixation stability was quantified using bivariate contour ellipse area (BCEA) and by measuring the frequency and amplitudes of microsaccades. Finally, for sensory eye dominance, we manipulated the contrast of letters presented to both eyes to find the contrast balance point (i.e., the interocular contrast difference) in which both eyes are equally contributing to the observer's perception. Results During binocular viewing, the logBCEA (-0.848 ± 0.129) and microsaccade amplitude (0.192 +- 0.036 deg) were not statistically different between the left and right eyes. Fixation stability was significantly worse during monocular viewing (-0.652 +- 0.135), which was further reduced for the covered eye (-0.386 ± 0.130). The rate of microsaccades during binocular viewing was 0.63 (SD 0.26) and 0.93 (SD 0.29) during monocular viewing ($p < 0.05$). Microsaccade amplitude was 0.19 (SD 0.04 deg) during binocular viewing compared to 0.22 (SD 0.06) for the viewing eye and 0.27 (SD 0.06) for the covered eye during monocular viewing ($p < 0.05$). For sensory eye dominance, the mean point of subjective equality (PSE) was 0.50 ± 0.04 (range: 0.44 to 0.54) indicating no particular dominance by either eye. There was no significant correlation between fixation stability measures and sensory eye dominance. Next Steps This study demonstrates that in a cohort of control participants with normal vision, fixation stability is not associated with sensory eye dominance. These results serve to establish a normal range of sensory eye dominance and the corresponding fixation stability, which will be helpful when interpreting results from clinical populations.

Mohamed Abdelhack	mohamed.abdelhack@camh.ca	Akio Murakami, Keita Suzuki, Fan Cheng, Kei Majima, Yukiyasu Kamitani, Hidehiko Takahashi	Uncovering the Visual Processing Imbalance in Schizophrenia Using Deep Neural Network Representations	Schizophrenia is a psychiatric disorder that is characterized by the occurrence of hallucinations among many diverse symptoms. Recent hypotheses suggest that hallucinations are induced due to an imbalance between top-down and bottom-up signals in the perceptual processing pathway. The nature of this imbalance is still poorly understood. To uncover this, we measured brain activity by functional magnetic resonance imaging (fMRI) while schizophrenia patients and control subjects were viewing blurred images. We then decoded the features of a pre-trained deep neural network from brain activations in order to measure the feature gain. Feature gain quantifies the level of top-down sharpening effect that the visual cortex performs to improve the neural representations. We compared the feature gain in the schizophrenia patients with that in the control subjects. We found considerable differences in the feature gain between schizophrenia and control subjects especially in the lower visual areas. The Schizophrenia group showed a lower feature gain indicating a lower top-down effect than the control group. This result suggests that an excess bottom-up processing is occurring in the schizophrenia group which supports previous behavioral results about the schizophrenia patients depending on the stimulus signal more than prior information in their perceptual inference. They could shed new light on the processing imbalance problem that could lead to a better understanding of cognitive deficits in schizophrenia.

Na Yeon Kim	nayeon@caltech.edu	Kushin Mukherjee, Shirin Taghian Alamooti, Ralph Adolphs, Kohitij Kar	Leveraging Artificial Neural Networks to Enhance Diagnostic Efficiency in Autism Spectrum Disorder: A Study on Facial Emotion Recognition	Autistic individuals differ in social behaviors, often due to a unique way of processing emotions from faces compared to neurotypical (NT) individuals. Gaining insights into the computations underlying these differences can aid in identifying the behavioral markers of autism. Traditional limitations in understanding these behaviors have been the lack of image-computable models that map image-level properties in facial photographs to behavioral patterns. In recent years, advances in artificial neural networks (ANNs) have resulted in models that partially mimic human behavior in object recognition tasks and offer insights into neural activity patterns in the primate's ventral visual pathway. Kar (2022) observed that such brain-mapped ANNs can make human-like facial emotion judgments. Interestingly, these ANNs' behavioral patterns aligned more closely with the NT population than autistic adults (ASD). This study investigates whether these advanced ANNs can guide experimental design by precisely predicting the emotional judgment differences between the NT and ASD groups when presented with specific images. If successful, these predictions can be instrumental in creating more diagnostic stimuli, enhancing our understanding of the variances between these two groups. Utilizing methods introduced by Kar (2022), we adapted representations from IT-analogous layers of several ANNs to predict facial emotion judgments in ASD and NT groups. This was based on data from Wang and Adolphs (2017). For this purpose, we extracted activation from ANNs such as AlexNet, ResNet-50, VGG-19, ConvNext-base, ViTbase, CLIP-ViT, and CORNet-S. Particularly for the CLIP model, we focused on the feature activation from the visual layer of the visual encoder. Using these ANN models, we curated a selection of images from the Montreal Set of Facial Displays of Emotion (MSFDE), which showcases emotional facial expressions across various ethnicities. We aimed to identify images that produced significant differences in emotional judgment between NT and ASD participants. Subsequently, an in-lab (and online) experiment was conducted using these selected images, where participants were tasked with distinguishing between 'fearful' and 'happy' expressions. This in-lab study included 12 adults with ASD and 13 NT participants who met the DSM-5 diagnostic criteria for ASD and the Autism Diagnostic Observation Schedule-2 (ADOS-2; module 4) criteria and 13 NT subjects (2 females, mean age = 32.2, age range = 24-39 years) with no history of psychiatric or neurological disease and no family history of autism. All subjects had normal-range IQ and corrected-to-normal visual acuity. The results revealed that images identified by specific ANN models, especially CLIP and ViT, resulted in more pronounced behavioral differences between the ASD and NT groups than randomly chosen images from the same dataset. A positive correlation ($R = 0.7$, $p = 0.04$) existed between the selected images' diagnostic power and an ANN's ability to predict NT responses. In essence, the higher-performing ANNs proved more beneficial as diagnostic instruments in experimental design. In conclusion, these findings underscore the potential of high-performing ANNs as valuable diagnostic tools in autism research, providing a framework for refining experimental designs that seek to understand the intricate behavioral differences between autistic and neurotypical individuals.

Naila Ayala	nayala@uwaterloo.ca	Ewa Niechwiej-Szwedo Suzanne Kearns Elizabeth Irving Shi Cao	Exploring Gender Differences in Gaze Behaviour During Simulated Flight	Background: Piloting an aircraft is a complex task that relies on the processing of task relevant visual cues in the cockpit and external environments. Inadequate monitoring of these relevant visual cues can significantly impact pilot performance and situation awareness (SA: hazard identification, evaluation, and correction). Previous research has suggested that gender differences exist in gaze behaviours associated with cognitive and perceptual tasks. No research has examined if these gender differences in gaze behaviour are apparent during the operation of an aircraft or if they impact pilot performance and SA. Aim: To explore gaze behaviour, performance, and SA differences between male and female pilots during a flight simulation task. Methods: Twenty pilots completed 9 landing scenarios that took place in an ALSIM 250 flight simulation device. The recruited group included 10 female pilots that were flight hour and rating matched with 10 male pilots. The scenarios consisted of 4 easy (high visibility, no/low winds) and 4 difficult (high visibility, high winds) landing tasks, as well as a single emergency flight scenario (i.e., engine failure). SA was subjectively measured after every trial using the Situation Awareness Rating Technique (SART) questionnaire. SA was also measured objectively by examining response time to the emergency flight scenario. Landing performance, gaze behaviour, and SART responses were examined using a 2x2 mixed model ANOVA (Gender: male, female; Task Difficulty: easy, difficult). Emergency trial response time and SART scores were analyzed separately using an independent samples t-test (Gender: male, female). Results: Performance results demonstrated no main effect or interaction involving gender (ps>0.455). Dwell time %, dwell count, and average dwell duration also failed to demonstrate a main effect or interaction involving gender (ps>0.084). Secondary gaze metrics (SGE, GTE, number of bouts, average bout duration, total bout time) were not significantly influenced by gender (ps>0.279). SART questionnaire scores revealed a main effect of gender (p=0.029) with female pilots (22.1, SD= 4.5) having a higher subjective SA score than male counterparts (16.3, SD= 6.9). The attentional demand component of the SART demonstrated an interaction involving gender and task difficulty (p=0.003). Specifically, an increase in SA demand scores associated with increasing task difficulty was significantly larger in male pilots (easy: 8.9, SD= 2.9; difficult: 16.1, SD= 9.5) compared to female pilots (easy: 6.1, SD= 2.7; difficult: 8.9, SD= 2.9). Emergency SART scores revealed a main effect of gender (p=0.037), while response times did not (p=0.243).Regression analyses for the emergency probe demonstrated that males had reduced SGE/GTE patterns when performing optimally (p=0.012), while females had larger SGE/GTE patterns (p=0.019). Conclusion: The current study demonstrated no significant gender differences in gaze behaviours or flight performance in low-time pilots. Subjective SART scores show some gender differences that may reflect differences in task confidence as this was not reflected by the emergency response times. Optimal emergency scenario performances in male and female pilots were supported by different visual scanning strategies.

Nicole Smeha	nsmeha@my.yorku.ca	Diana J. Gorbet, Lauren E. Sergio	Visuomotor control network changes as a function of hormone levels in working-aged women	Interactions with our environment require intact connections between frontal, parietal, and subcortical brain regions. While some of these interactions are direct, others require cognitive-motor integration (CMI), where the guiding visual information and motor action are decoupled. Our group has shown sex differences in the networks controlling these visuomotor skills.1 However, most of this research is based on data from pre-menopausal college-aged women. To address the knowledge gap around the impact of sex hormones on the neural control of movement, we examined females between the ages of 34 and 57 in their pre, peri, and postmenopausal stages. We hypothesized that the neural activity underlying skilled movement control would differ as a function of hormone concentrations. Participants underwent MRI scanning during which they performed standard and CMI eye-hand coordination tasks using an MRI-safe touchscreen. Participants were trained on the tasks until performance reached greater than 90% success. Estrogen, progesterone, and testosterone levels were also collected. A generalized psychophysiological interaction (gPPI) analysis was implemented to examine how functional connectivity between regions known to be involved in our CMI task was modulated by sex hormone concentrations. Our preliminary linear regression analysis showed that lower levels of testosterone were associated with increased right prefrontal cortex-inferior parietal lobule functional connectivity (IPL; $p < 0.05$), and precuneus-right superior parietal lobule connectivity (SPL; $p < 0.05$) in the standard condition. Further, an elevated estrogen:testosterone (E:T) ratio was associated with increased bilateral prefrontal cortex-IPL connectivity ($p < 0.05$), as well as precuneus-left SPL connectivity ($p < 0.05$). In the CMI condition, lower levels of progesterone were associated with increased left prefrontal cortex-IPL ($p < 0.05$), precuneus-bilateral SPL ($p < 0.05$), and right-left IPL connectivity ($p < 0.05$). Finally, a smaller E:T ratio was associated with increased functional connectivity between the right prefrontal cortex and SPL ($p < 0.05$). Thus, with equivalent performance, we observe hormone-related differences in the connectivity strength of the brain networks for skilled performance. These data show that the neural control of visually-guided movement is differentially affected depending on sex hormone levels. We suggest that the structure of the networks required for accurate movement performance changes with the fluctuations in progesterone, estrogen, and testosterone that occur later in life. 1. Gorbet D & Sergio (2007). EJN 25(4), 1228-1239.

Nikita Klimenkov	klimenkovnkt@gmail.com	Svetlana Kovalenko, Elena Gorbunova	The influence of perceptual load on cognitive fatigue	Cognitive fatigue is usually defined as a functional mechanism of behavior regulation aimed at adapting resource-demanding cognitive processes to overload, including the perception of the overload itself. Cognitive fatigue reduces cognitive control, sustained attention, and the ability to ignore irrelevant information. The dominant approach to understanding cognitive fatigue is resource-based theories of attention. The main factors in cognitive fatigue are task duration and task complexity. Therefore, task difficulty at the perceptual level can be assessed using the concept of perceptual load. The ability to ignore irrelevant distracting stimuli is important for everyday life, as distraction can have a range of consequences, some of which are mortal. Theory of perceptual load proposes the idea of a hybrid mobile filter whose position will be determined by the parameters of perceptual processes and the availability of resources for information processing. When describing the complexity of a task, the term perceptual load should be used, which includes the number of stimuli, the similarity of these stimuli, and the number and complexity of perceptual operations in working memory. Since attention resources are limited, when they become insufficient to process relevant stimuli, processing of irrelevant ones decreases. If resources are insufficient to process relevant stimuli as well, overload occurs. When overload turns out, a state of fatigue occurs in order to adapt and form a different strategy to perform the task. However, it is unclear how the lack of attentional resources causes fatigue: whether it would be determined by the degree of distractor processing by the subject. Experiment was conducted to examine the effect of perceptual load. In both experimental conditions, only the presence of fatigue induction in the form of task difficulty — increasing perceptual load — was varied. The experiment consisted of three blocks, where behavioral measures would be compared between the first and last blocks, and the block between them served as the induction of condition. Each block consisted of 58 trials, where 2 stimuli are presented to the subject on the monitor, which are required to be memorized. Then, one stimulus disappears and after a fixation cross, 4 stimuli are presented, among which there is a disappeared stimulus. The subject is required to click on the disappeared stimulus. Letters of Armenian, Japanese and Hindi languages were used as stimuli with low perceptual load, Chinese letters were used as stimuli with high perceptual load. Also, before and after the experiment, subjective indicators of fatigue were measured using a questionnaire. Significant statistical differences were found in the subjective feeling of fatigue and in the number of errors in the group with high perceptual load in the second block of the experiment. However, no significant differences in reaction time were found between the first and the last block. In the group with low perceptual load in the second block, no differences between the first and the last block were found for any of the dependent variables.

Nupur Katyare	nupur00@yorku.ca	R. Wong, S. Everling, Liya Ma	Prefrontal ensemble dynamics in spatial working memory in marmosets	Neurons in the lateral prefrontal cortex (lPFC) are known to play a role in spatial working memory. Single unit recordings from the lPFC of marmosets during a delayed-match-to-location task confirmed that like Macaques, marmoset lPFC neurons exhibit directional tuning, which correlated with task performance (Wong et al., 2022) . Not much, however, is known about population coding in the lPFC of marmosets. Here, we re-analyzed data from Wong et al. (2022), focusing on the coding of spatial information in neuronal ensembles using support vector machines, using 2 different approaches: Best unit and Optimized (Leavitt et al., 2017). In the Best-unit approach, we first sorted neurons based on their decoding accuracy. We then calculated decoding accuracy of top-N neuron ensembles at each size. In the Optimized approach, we identified the N-1 neurons that together with the top neuron, achieved the best performance. As previously reported in macaques, the Optimized approach significantly out-performed the Best-unit approach; and ensembles of size 5-15 performed superior to full ensembles. We next analyzed the ensemble performance separately for broad-spiking and narrow-spiking neurons (BSNs and NSNs). Although BSNs dominated the ensembles, the NSNs contribute substantially to the decoding, especially in the Optimized ensembles, indicating a significant role of interneurons in spatial coding. Analysis of signal and noise correlations revealed a population of non-tuned neurons that were present in Optimized but not in the Best-unit ensembles, and 80-100 % of these exhibited significant noise correlations with tuned neurons. This observation supports a role of noise correlations in enhancing spatial coding. Finally, we observed that the average inter-unit distance for a 10-neuron ensemble was significantly shorter than the average inter-unit distance among the population. This indicates colocalization of neurons that play key roles in spatial coding in lPFC. Our results provide the first evidence that population-level coding properties in the lPFC can be generalized across primate species. References Leavitt ML, Pieper F, Sachs AJ, Martinez-Trujillo JC (2017) PNAS 114:E2494-E2503.Wong RK, Selvanayagam J, Johnston KD, Everling S (2022) Cerebral Cortex: bhac289.

Oluwaseyi Elizabeth Shodipe	oluwaseyi_dos@yahoo.com	Robert S. Allison	Modelling the relationship between the objective measures of car sickness	Car sickness is a potentially serious barrier to the adoption of self-driving vehicles, as autonomous driving could increase its incidence greatly. Thus, detecting and predicting the early onset of car sickness is important to prevent users' severe side effects and increase adoption. We investigated whether machine learning techniques based on objective physiological and behavioural measures can assist in the detection and prediction of car sickness. Participants (N=40) experienced two 20-minute rides in a motion-base car simulator while a variety of physiological and behavioural measurements were recorded. In one of the rides they read a passage on a tablet and subsequently answered reading comprehension questions; in the other ride, they performed no task and sat quietly in the car. The order of the two tasks was counterbalanced across participants. Car sickness incidence and severity were subjectively measured using the Fast Motion Sickness Scale (FMS) questionnaire at two-minute intervals during the rides and with the Simulator Sickness Questionnaire (SSQ) before and after each ride. Car sickness symptoms were reliably produced and the SSQ nausea subscale score correlated highly with the FMS ratings (up to r = 0.80). These car sickness symptoms were associated with a number of physiological and behavioural signs. Furthermore, results showed that the number of saccades participants made during reading reduced as car sickness increased, suggesting a reduction in the reading rate. While respiration measures showed no strong correlations with car sickness ratings, they were influential measures in the machine learning analysis. In addition, skin conductance levels generally increased over the experiment for several participants; however, this alone was not a reliable predictor of the severity of car sickness. Females had a higher and positive correlation between heart rate and FMS scores in comparison to the males. The suite of behavioural and physiological measures recorded was predictive of carsickness but only explained a relatively small amount of the variance for the overall severity of car sickness (up to 15%). These measures are ambiguous on their own and reflect many factors influencing the user's state. However, when combined in a Random Forest model they supported distinguishing between sick and non-sick participants with an acceptable accuracy score of 77%. Portions of this study were previously reported at IEEE CCECE 2023.
Onoise Gerald Kio	ogkio@yorku.ca	Robert Allison	Data-driven prediction of outcomes of a target search task	An indicator of the level of success of operators performing search tasks with remote-controlled drones is the rate of finding specific targets within operation time constraints. Can flight data obtained from such drones during search tasks be used to assess the performance of operators in such tasks? Thirty-seven operators aged 18 – 44 years were trained in using a remote-controlled drone to search for specific alphanumeric targets arranged on the laboratory floor. They performed the same task in three 3-minute runs, and then assessed the performance of other drone operators based on observing animated and non-animated simulations of the drone performing the same task. Statistical comparisons of subjective assessments showed less than 10% correlations between assessed and actual task performance scores. Principal Components Analysis of the drone's flight data was used to extract salient components and cluster operators based on two levels of performance namely, below average, and above average. These results reveal the possibility of developing data-driven systems having potential to predicting task outcomes, improve operator efficiency, and provide objective measures of performance-based trust for operational teams.

Parham Eftekhar	parham.eftekhar@gmail.com	Gene Cheung	Image satellite inpainting	Image inpainting is the process of restoring or completing an image by predicting and filling in missing or damaged regions. These regions can be missing due to various reasons, such as data corruption, object removal, or simply to enhance the visual appeal of an image. Image inpainting has various applications such as photo restoration, object removal, and medical imaging. Image inpainting can also be generalized to apply on 3D data such as point clouds. One interesting application of image inpainting is satellite images. Recent advances in satellite imaging technology have led to an abundance of satellite images. However, depending on location and time, some areas may be occluded by clouds. Thus, completing missing pixels in the occluded areas is of significant importance. There are many ways to address the problem of image inpainting. Patch-Based methods attempt to seek similar patches within known regions to fill the missing parts. Deep Learning-Based algorithms aim to learn the underlying distribution using a vast amount of data to fill in the gaps. Another stream of works use graphs to deal with the problem. Specifically, Graph Signal Processing (GSP) is an emerging and rapidly expanding field that extends traditional signal processing techniques, originally designed for regular data kernels, to encompass irregular kernels described by combinatorial graphs. One can treat an image as a graph with nodes representing the location of pixels and edges representing similarities of neighboring pixels. Therefore, the theory of GSP can be applied on images. Precisely, to solve the problem using GSP, one can formulate an optimization problem encompassing two terms. First is a fidelity term that encourages the reconstruction of the entire signal, which closely matches the known part of the original signal. The second one is a prior term utilizing the Graph Laplacian Regularizer (GLR) to promote signal smoothness in the reconstructed signal. Under this framework, the unique challenge arises from the extensive area of absent pixels, which implies a lack of information about the inherent similarity relationships required to establish a smoothness prior such as GLR. Therefore, an alternative approach needs to be introduced initially to create a graph that can be used to estimate a signal. Subsequently, this signal can be employed to iteratively update a suitable graph through alternating optimization. This alternating graph/signal optimization is currently under investigation.

Parsa Balalaie	p.balalaie@queensu.ca	Kayne Park Stephen H. Scott Jolande Fooken	Investigation of eye-hand coordination during rapid interception task	To catch a falling ball or avoid a collision with traffic, humans must be able to quickly respond to sudden changes in their environment. For example, when a visual target suddenly jumps to a different location, humans have been shown to rapidly generate a corrective response towards the new target location within ~150 ms. These rapid corrections are thought to be an automatic response mediated by a subcortical network. Here, we asked whether and how rapid eye and upper limb movements are linked behaviourally. We investigated eye-hand coordination in 14 participants on a fast feedback interception task (FFIT) in which participants viewed and intercepted a target that moved vertically down a visual scene. In some trials, the target could jump laterally to the left or right, requiring rapid eye and upper limb movement to intercept the target successfully. To understand whether eye and upper limb movements were similarly affected by changes in movement constraints, we manipulated the percentage that the target could jump (certainty) as well as the time available to respond to the target jumps (urgency). We manipulated interception certainty by introducing blocks in which the target jumped either in 100% of the trials (high jump-certainty) or in only 60% of the trials (low jump-certainty). We further manipulated interception urgency by changing the location at which the target jumped (easy/early jump, medium/middle jump, or hard/late jump). In the low certainty condition, the target continued to move down the middle in 40% of trials. Although eye movements were unconstrained, we commonly observed that participants tracked the moving target before initiating a saccade to the anticipated interception location. Results show that saccade latencies systematically preceded upper limb movement onsets by ~60 ms. On a trial-by-trial basis, we found a positive, but weak correlation between eye and upper limb movement responses, and earlier movement responses (short latencies) tended to yield less accurate movement responses. On a group level, we found that upper limb and eye responses were similarly affected by changes in certainty and urgency. We found that reaction times were faster when a target jump was certain (100% jump trials) and when the target jump occurred later, resulting in higher urgency. Overall, our results highlight that although early visuomotor responses are relatively automatic, the speed and accuracy of the movement responses are modulated intelligently depending on movement constraints in the environment. Overall, our results show that eye and upper limb movements are similarly modulated by context-dependent expectations, suggesting shared mechanisms in the neural control of eye and limb movement responses.

Patrick Cavanagh	patcav1@yorku.ca	Stuart Anstis	The frame effect is suppressed for stationary probes.	Targets flashed within a frame that moves back and forth are offset by as much as the frame's displacement (Özkan et al, 2021; Cavanagh et al, 2022), a far larger offset than that seen on stationary or continuously moving targets (induced motion, Duncker, 1929; Wallach et al, 1978). To understand the difference between continuous and flashed targets we varied the frequency from a single flash per transit of the frame up to continuously present (30 Hz). We find a monotonic decrease in the illusory offset as the flash frequency increases. However, the setting for a stationary target, where there is little or no effect, does not lie on this function, suggesting that the stationary target may group with the steady background instead of the frame. To test this, we presented continuous targets that were steady or jiggled vertically while the frame moved horizontally. When the continuous target had any movement, the illusory horizontal movement was visible (induced motion). It did not matter if the frame itself also jiggled vertical, synchronously or asynchronously with probe. In our original frame effect where the target flashes once per transit, the apparent separation between the two flashes is maximum and there is no sense of motion between the two flashes. When there are multiple flashes or continuous movement, there is illusory movement but the apparent distance travelled is reduced compared to the separation seen when there is just the one flash at each end of the frame's travel. When the target is stationary, no illusory movement is seen at the frame speeds used here, most likely because it groups with the stationary background.
Pranavan Thirunavukkarasu	thirunap@yorku.ca	Steven P. Errington, Amirsaman Sajad, Jeffrey D. Schall	Laminar architecture of visual responses in supplementary eye field of macaques	Previously, we have described the laminar organization of neurons in the supplementary eye field (SEF) that signal error, reward gain and loss, conflict, event timing, and goal maintenance. Here we describe the laminar organization of visually responsive neurons that were active during performance of a saccade stop-signal task. Nearly 40% of isolated neurons exhibited enhanced or suppressed responses to a visual target for a potential saccade, with the majority exhibiting enhanced activity and three-quarters with broad spikes. Visually responsive neurons were observed in all layers but were less common in layers 5 and 6. Response latencies were comparable to those reported previously, which are significantly later than those measured in occipital and temporal visual areas but overlapping those measured in cingulate cortex. Task-related visual response latency varied across cortical layers. Response latency was significantly earlier for neurons with narrow spikes. Neurons with task-related visual responses discharged until after saccade production. Around three-fifths of visually responsive neurons were most sensitive to the visual target appearing in one hemifield. Many neurons in layer 2 had ipsilateral receptive fields. Laminar current-source density aligned on visual target presentation revealed the earliest sink in layers 3 followed by a prolonged strong sink more superficially coupled with a weaker prolonged sink in layer 5 and a transient sink in layer 6. The current sink in layers 2 and 3 was stronger for ipsilateral stimuli. These findings reveal new details about visual processing in medial frontal cortex and complete the first catalogue of laminar organization of functional signals in a frontal lobe area.

Raphael Gastrock	gastrock@yorku.ca	Edward Ody, Denise Y. P. Henriques, Bernard Marius 't Hart	Motor adaptation versus de novo learning: Comparing neural markers of movement preparation and outcome error processing between two distinct visuomotor tasks	When people encounter movement errors, they process these errors to correct ensuing movements. This error processing contributes to motor learning, either when we are adapting well-known movements or acquiring new motor skills (de novo learning). While the distinct behavioral mechanisms underlying these two motor learning types have been investigated before, we understand less about the neural correlates of each type. Here, we investigated event-related potentials (ERPs) during movement preparation and outcome as participants performed reaching movements. We distinguished the two motor learning types by having participants (N = 32) train with two perturbations in counterbalanced order: a 30° visuomotor rotation to investigate adaptation and a mirror reversed cursor feedback for de novo learning. Before training with each perturbation type, participants completed a control condition, where they experienced a random rotation with magnitudes of ±15°, ±25°, or ±35°. Participants learned to compensate for both the fixed rotation and the mirror reversal, but not the random perturbations. For movement outcome, we time-locked to feedback onset at the end of the movement. We found a negative-going ERP before feedback onset in fronto-central and parietal electrodes, with perturbed reaches showing more negativity compared to aligned baseline reaches. However, the ERP amplitude did not scale with perturbation type nor error magnitude, suggesting that it was only processing the presence of an error. Furthermore, we found a larger P3 component after feedback onset in the perturbed conditions compared to aligned baseline reaches, suggesting more attention allocation for perturbed reaches. For movement preparation, we time-locked to the go signal onset before the movement. We found a Readiness Potential (RP) that depended on whether participants moved to the right or left side of the workspace. We then quantified changes in movement preparation across learning, using a Lateralized Readiness Potential [LRP = (right C3 – right C4) – (left C3 – left C4)] for different blocks of trials during training in each of the perturbation types. We found less pronounced LRPs for the random perturbation than in aligned reaches, suggesting weaker preparatory activity for such unpredictable trials. However, we found no LRP differences between the fixed rotation and mirror reversed reaches. Thus, although the ERPs we investigated represent movement preparation and outcome error processing, these markers are unable to distinguish between motor adaptation and de novo learning. We will further investigate other markers in relation to these two motor learning types.

Remy Cohan	Rcohan@yorku.ca	Jennifer Steeves	The effect of continuous theta burst stimulation to primary visual cortex on binocular rivalry	Promoting plasticity in neural networks is one of the hallmarks of neuromodulation techniques. Continuous theta burst stimulation (cTBS) is a technique that has been shown to induce inhibitory responses when applied to the primary motor cortex. To evaluate the efficacy of cTBS in the visual brain we applied 600-pulses of cTBS at 80% individual phosphene threshold (PT) to primary visual cortex (V1) before and after performing a binocular rivalry (BR) task. BR is a perceptual phenomenon that occurs when the two eyes are each presented with different images simultaneously. This leads to alternating dominance and suppression periods of encoding of visual information as the brain attempts to reconcile the conflicting information. BR involves an interplay between multiple levels of the visual system, from the eyes to V1 and higher-level cortical areas. However, research has identified cortical columns in V1 that respond preferentially to one eye or the other. These columns play a role in the alternating perception observed in BR. Our study explored the effect cTBS to V1 on BR in right-handed, right eye-dominant subjects. Participants engaged in a BR task where they viewed orthogonal grey-scale gratings of fixed orientation (+/-45°) through a stereoscope. Subjects were instructed to indicate perceived visual dominance (leftward tilting grating, rightward tilting grating, or mixed percept) by pressing and holding designated buttons upon perception change. cTBS was applied to left V1 and BR was again measured. Our preliminary results indicate an increase in the alternation rate of BR post-cTBS. This suggests that cTBS to V1 induces an alteration in visual perceptual dominance and furthers our understanding of the neuromodulatory effects of cTBS on visual perception. Overall, these results shed light on the potential of cTBS as a research and therapeutic intervention for visual disorders with neural origins and contribute to the growing body of literature on neuromodulation, visual perception, and brain plasticity.
Rezaul Karim	karimr31@yorku.ca	Mennatullah Siam, Richard P. Wildes	MED-VT: Multiscale Encoder-Decoder Video Transformer for Video Segmentation	Multiscale video transformers have been explored in a wide variety of vision tasks. To date, however, the multiscale processing has been confined to the encoder or decoder alone. We present a unified multiscale encoder-decoder transformer that is focused on video segmentation. Multiscale representation at both encoder and decoder yields key benefits of implicit extraction of spatiotemporal features (i.e. without reliance on input optical flow) as well as temporal consistency at encoding and coarse-to-fine detection for high-level (e.g. object) semantics to guide precise localization at decoding. Moreover, we propose a transductive learning scheme through many-to-many label propagation to provide temporally consistent predictions. We showcase our Multiscale Encoder-Decoder Video Transformer (MED-VT) as a general solution to multiple video segmentation tasks, in particular Automatic Video Object Segmentation (AVOS), Video Semantic Segmentation (VSS) and actor/action segmentation. We outperform state-of-the-art approaches on multiple benchmarks using only raw images, without using optical flow.

Ricky Chow	chowrk@yorku.ca	Stevenson Baker, Deena Herman, Shimin Mo, Jennifer A. Bugos, Claude Alain, R. Shayna Rosenbaum	Mismatch negativity as an index of auditory pattern separation in aging	The mismatch negativity (MMN) is an event-related potential associated with perceptual change-detection. Whether the MMN extends beyond an index of perceptual discrimination and correlates with higher-order cognition, such as episodic memory, is unclear. To what extent does perceptual discrimination, indexed by the MMN, contribute to subsequent memory for those same stimuli? Furthermore, how does this relationship change in aging? We hypothesized that the MMN, generated by incidentally encoded auditory stimuli, would be correlated with younger and older adult participants' ability to discriminate those stimuli (targets) from highly similar lures and from dissimilar foils. We measured the MMN in 30 younger adults and 26 older adults using a passive auditory oddball task with standard and deviant pure-tone sequences differing in pitch contour. After exposure, all participants completed an incidental memory test for targets, lures, and foils. As expected, participants at test exhibited high sensitivity at recognizing target items relative to foils. Within our younger adult sample, we found a significant correlation between MMN amplitude and lure discrimination, but not with foil discrimination. Preliminary analyses with older adults show evidence for null correlations with memory performance. Findings suggest that the MMN of younger adults relates to recognition memory, particularly mnemonic discrimination. Our investigation shows that our capacity to discriminate sensory inputs, as measured by the MMN, translates into precision in memory.

Romesa Khan	romesa.khan@mail.utoronto.ca	Matthias Niemeier, Hongsheng Zhong, Jack Cai	Predictive coding dynamics in a grasping neural network increase robustness to noise	The brain is widely assumed to be a “prediction machine”, according to which perception makes inferences about the causes of noisy sensory inputs relying on generative models. Top-down predictions from these generative models are conveyed through cortical layer-specific feedback connections during visual perceptual tasks. However, there is a dearth of understanding of the contribution of feedback when sensory input is used for action planning, such as during prehension behaviours. I.e., when grasping objects to move and manipulate them the brain transforms visual information about the object into representations of movement intention, as stimulus information ascends (feeds forward) from visual to motor cortex. Crucially however, recent evidence shows that object shape representations during grasping also involve the reactivation of earlier visual areas. Additionally, advanced grasping movement representations appear to be reflected in early visual cortical activity. Together these observations indicate that recurrent connections carry information from downstream stages of visuomotor processing to the earlier stages in the visual stream and are crucial for core object recognition behaviour. We investigated the contribution of such neural feedback to the visuomotor control of grasping by using convolutional neural networks (CNNs), trained to compute grasp predictions from visual images with colour and depth channels of real-world objects, as a modelling framework. The architecture of CNNs originally relies only on a feedforward flow of information. To make these models computationally and structurally more similar to the human cortex, we added generative feedback loops to a VGG-16 backbone (a 16-layer CNN), carrying advanced representations to early layers of the network. We compared the performance of a baseline feedforward VGG-16 to one augmented with predictive coding using visual images with and without additive Gaussian noise. During inference, at the first timestep the feedback-enhanced network was, by design, precisely identical in accuracy to its conventional feedforward counterpart. However, after several forward and backward passes through the network, the impact of predictive coding iterations emerged. We observed an improvement in network performance under adverse conditions, in comparison to the feedforward baseline. To conclude, our simulations show that introducing biologically plausible predictive coding dynamics to a feedforward neural network improves model robustness to noisy visual stimuli in a neural network model optimised for grasp prediction.

Rupsha Mutsuddi	rupsham@yorku.ca	Melanie Baljko, Shital Desai	Daily Rituals in Everyday Activities for Community Dwelling Adults with Dementia for Assistive Prompting Technology Design – Understanding the Current Research Landscape	The study presents the findings of a scoping review that aims to investigate the role of integrating daily rituals into assistive prompting technologies for people living with dementia (PwD). The scoping review involved a systematic search of electronic databases which yielded 19 articles. Thematic analysis was used to find recurring patterns and themes across the literature. Three main themes emerged from the analysis. First, multimodal, personalized, and context-specific prompts which are tailored to the unique needs, preferences, and abilities of PwD improve engagement, autonomy, and independence. This is especially true when these prompts also integrate cultural, personal, and social aspects of daily rituals. Second, when components of daily rituals like routines and habits are supported by assistive technologies, it creates familiarity, structure, and a sense of continuity. This has a positive impact on emotional wellbeing and in some cases, improves cognitive functioning. Lastly, daily rituals facilitate the development of relationships between PwD and their caregivers as well as their environment. This promotes a sense of belonging and leads to a more supportive care environment. The integration of daily rituals also leads to better habituation and adoption of assistive prompting technologies. Insights from this review are helpful for considering how daily rituals can be integrated in the design of prompting assistive technologies for people with dementia.
Sanjana Kapisthalam	skapisth@ur.rochester.edu	Martina Poletti	Temporal modulations of extrafoveal sensitivity to changes during fixation	Being able to correctly identify sudden changes in the environment is crucial for survival. Here we examine how sensitivity to brief changes is modulated over time during the course of fixation and whether it is impacted by the salience of surrounding stimuli. Subjects (n=7) maintained fixation on a central marker, either at rest or right after a saccade. An 8 cpd gabor patch, 1 deg in size, was presented 8 degrees away from the center of gaze. The orientation of the gabor changed briefly (50 ms) at a variable time (0-450 ms) either from saccade landing or from stimulus onset. Subjects were instructed to determine the direction of the orientation change. The gabor could be flanked either by salient or non-salient circular blobs of the same size (1.4 degrees center-to-center distance). To prevent visual fading stimuli were presented at high contrast and were jittered throughout the presentation time. Our results show that subjects' ability to discriminate the direction of sudden orientation changes was best right after the stimulus onset and upon saccade landing (0-150 ms) and it decreased over time during the course of fixation. On average, performance dropped by 17% around 350 ms. Importantly, the perceptual salience of the surrounding did not influence performance, suggesting that in this task the visual system is capable of actively suppressing salient distractors at no cost. These findings show that the ability to discriminate sudden changes in the visual surrounding varies drastically during the short periods of fixation in between saccades.

Sara Chaparian	Sarach@yorku.ca	Jeffery Schall Peter J.Kohler	Relating Variability in Scalp EEG to Variability in Cortical Morphology	Electroencephalography (EEG) is a widely used brain imaging modality that makes it possible to measure brain signals with minimal cost and invasiveness. A fundamental limitation to the usefulness of EEG is the challenge of relating signals measured at the scalp to the underlying cortical generators. Here we use previously collected EEG that was acquired under a Steady-State Visual Evoked Potentials (SSVEP) paradigm as participants (n=12) were viewing a stimulus set of regular textures. These data were acquired on two separate sessions, and the experimental design makes it possible to separately measure early visual responses related to image-level changes in the stimulus and more higher-level responses driven by symmetries within textures. We see substantial variability among individuals in early and especially in higher-level responses which clearly exceeds the variability between sessions. Because the underlying set of cortical areas that respond to symmetry is well-documented, it is possible to model the cortical sources of both low-level and higher-level symmetry-driven responses in individual participants. We do this based on structural MRI data for each of our participants and aim to quantify the extent to which variability in underlying cortical morphology determine variability in low-level and higher-level responses measured at the scalp. Specifically, we test whether features such as size of visual regions of interest, cortical thickness, and orientation of cortical surface can effectively predict the strength of responses.

Sara Djambazovska	sara.djambazovska@gmail.com	Gabriel Kreiman, Kohitij Kar	Visual angle and image context alter the alignment between deep convolutional neural networks and the macaque ventral stream	A family of deep convolutional neural networks (DCNNs) currently best explains primate ventral stream activity that supports object recognition. Such models are often evaluated with neurobehavioral datasets where the stimuli are presented in the subjects' central field of view (FOV). However, the exact visual angle often varies widely across studies (e.g., 8 degrees for Yamins et al., 2014; 2.9 degrees for Khaligh-Razavi et al., 2014; catered to V1 neuronal receptive field, 2 degrees for Cadena et al., 2019). A unified model of the primate visual system cannot have a varying FOV. In addition, the type of images used for model evaluation varies across studies, ranging from objects embedded in randomized contexts (Yamins et al., 2014) to objects with no contexts (Khaligh-Razavi et al., 2014). Thus, the inference made about the "goodness" of the alignment of these models with the primate brain becomes fairly incomparable across studies. To probe the severity of this concern, here we systematically tested how the predictivity of macaque inferior temporal (IT) neurons by DCNNs depends both on the FOV and the image-context. We used 150 images (with varying image-context) from the Microsoft COCO imageset. We performed large-scale recordings in one macaque (~78 reliable IT sites) while the monkey passively fixated images presented at 20 degrees. To estimate the optimal FOV for the DCNNs, we compared the DCNN IT predictivity at varying image crop sizes. We observed that ~ 7-11 visual degree crops (centered) produced the strongest DCNN IT predictions. Next, to test the effect of image-context, we generated nine additional versions of 60 full-context images: e.g. object only ("no-context"), swapped backgrounds ("incongruent-context"), removing the object ("no-object"), blurred context, among other manipulations. Preliminary results show that DCNN's (n=9) IT predictivity was significantly lower for "incongruent-context/full context" compared to the "no context" images. This was consistent across early (80-120ms) and late (140 - 180ms) neural responses. However, consistent with earlier studies, early IT responses were better predicted compared to late neural responses, across these context categories. In sum, our results provide evidence that controlling for image parameters like FOV and image-context is critical for accurate evaluation of DCNNs as models of the primate brain. Our results establish further constraints to guide the development of more brain-aligned DCNN models of primate vision.

Seohee Han	seohee.han@mail.utoronto.ca	Morteza Rezanejad, Dirk B. Walther	Memorability of line drawings of scenes: the role of contour properties	Why are some images more likely to be remembered than others? Previous work focused on the influence of global, low-level visual features as well as image content on memorability. To better understand the role of local, shape-based contours, we here investigate the memorability of photographs and line drawings of scenes. We find that the memorability of photographs and line drawings of the same scenes is correlated. We quantitatively measure the role of contour properties and their spatial relationships for scene memorability using a Random Forest analysis. To determine whether this relationship is merely correlational or if manipulating these contour properties causes images to be remembered better or worse, we split each line drawing into two half-images, one with high and the other with low predicted memorability according to the trained Random Forest model. In a new memorability experiment, we find that the half-images predicted to be more memorable were indeed remembered better, confirming a causal role of shape-based contour features, and, in particular, T junctions in scene memorability. We performed a categorization experiment on half-images to test for differential access to scene content. We found that half-images predicted to be more memorable were categorized more accurately. However, categorization accuracy for individual images was not correlated with their memorability. These results demonstrate that we can measure the contributions of individual contour properties to scene memorability and verify their causal involvement with targeted image manipulations, thereby bridging the gap between low-level features and scene semantics in our understanding of memorability.
Shanaathanan Modchalingam	s.modcha@gmail.com	Andrew King, Denise Y. P. Henriques	Effects of sensory prediction errors and visual environment cues on internal model updating and switching during motor learning.	When performing motor tasks, we improve future movements by detecting and correcting for motor errors. Errors are determined by comparing predicted outcomes of internal models of the motor interaction with the observed outcomes. When errors are detected, we correct for them by either updating existing internal models of the movement interaction, or by creating, switching to, and switching from new internal models. Assignment of the error's source, termed error attribution, can impact whether we update existing or create new internal models. Since the cause of an error is often ambiguous, sensory cues about the effector, the object being interacted with, and the environment in which the interaction occurs are used to estimate the likely source. In this study participants attempted to hit targets by rolling a ball along a surface. We tested whether informative visual cues about environment changes, represented by the horizontal slant of the surface, could successfully facilitate model creation and switching - characterized by fast, one-trial decay of learning when environment changes are detected - when adapting to two types of errors. We induced errors by either modifying the mapping between the arm movement and the initial movement of the ball (a visuomotor rotation), or by applying a constant acceleration to the ball's travel path only after the initial release of the ball. The surface-slant visual cues were informative of the direction of the visuomotor rotation, and the direction and magnitude of the acceleration perturbation. The error induction method alone, and not the visual cues, determined whether errors led to model creation and switching, or model updating. Visuomotor rotations led to model updating and acceleration perturbations led to model creation and switching. Additionally, in follow-up experiments, we found internal models that are updated consider both the hand used in the movement, and the physical properties of the environment on which the movement occurs. That is, the internal model being updated is not purely a model for the control of limb movement, but an interaction model.

Shaya Samet	shayasam@yorku.ca	Baker, N, Freud, E, Elder, J, Kohler, PJ	EEG Studies of Configural Shape Perception	The perception of object shape is an important goal of vision that enables adaptive and goal-directed behaviours and interaction with the world around us. Objects have both local and configural shape properties. Local shape properties manifests in a confined region of the object and can be interpreted without reference to other shape features on the object, while configural shape properties is a function of one or more specific local features, and spatial relationships among those features. We used EEG to investigate the cortical mechanisms that enable configural shape perception. We collected high-density EEG data from 32 participants using 128-sensor HydroCell Sensor Nets (Magstim EGI) while presenting object shape silhouettes. We presented the stimuli to participants for passive viewing using a Steady-State Visual Evoked Potentials paradigm. We recorded EEG of participants' neural responses to 1) Curvature-Matched Controls, a family of synthetic, maximum-entropy shape stimuli that progressively match the local curvature statistics of natural shapes but lack global regularities; 2) Unmanipulated animal-shape silhouettes; as well as 3) Upside-down inverted animals, and compared the responses to the different stimuli. The inverted animals were added to measure brain-based correlates of behavioural findings indicating attenuated configural processing for inverted stimuli. We found that curvature matched-controls vary in brain responses, with moderately constrained objects evoking a stronger brain response than objects with no constraint, and animal-shape objects evoke the strongest response among our stimuli. We observed an inversion affect as the neural response to animal silhouettes was removed when the animal stimuli were inverted.

Shenoa Ragavaloo	shenoa01@yorku.ca	Peter J. Kohler	Brain Responses to Symmetries in Naturalistic Novel Three-Dimensional Objects	Symmetries are prevalent in natural and man made objects and scenes. During natural vision, symmetries in the world are subject to perspective and rarely produce symmetrical images on the retina. In this exploratory study, we investigated the ability of the human visual system to overcome distortions, by using high density EEG to measure responses to naturalistic, novel three dimensional (3D) objects. We used a Steady State Visual Evoked Potentials (SSVEPs) paradigm to isolate EEG signals specifically associated with symmetry processing. We presented images of symmetrical objects paired with images of asymmetrical objects under two conditions: One where images were rendered to produce symmetries in the image plane, and another where the objects were rotated relative to viewing direction such that symmetries in the object would be distorted in the resulting image. We selected image pairs that were matched on non symmetry related features based on activations in an artificial neural network. The two images in a pair were shown one after the other, each for 500ms, to form a stimulus cycle (stimulation frequency=1 Hz). For both conditions, we created control conditions where the objects never had symmetry. Participants (n=30) passively viewed 10 such cycles per trial, with 10 unique image pairs. A follow up experiment (n=30) followed the same design but used images in which the shading cues to 3D object shape had been removed to create 2D object images. Previous studies have found that distorted symmetries in dot patterns elicits symmetry responses only when participants are engaged in symmetry related tasks. We find that image level symmetry elicits strong and broad responses suggesting occipital and temporal cortical sources. Perspective distorted symmetry elicits strong responses, but they are weaker in posterior locations and right lateralized. The 2D versions of the image level symmetry stimuli elicit similar responses to 3D. Removing cues to 3D shape eliminates symmetry responses from perspective distorted stimuli, serving as a manipulation check. Our results show that during passive viewing, perspective distorted symmetry can elicit SSVEPs that are comparable to those elicited by image level symmetry, but strongly right lateralized and only in regions likely driven by activity in higher level visual cortex. Future work will determine how task and viewing conditions influence responses.

Shirin Taghian Alamooti	staghian@yorku.ca	Hamidreza Ramezanpour, Na Yeon Kim, Ralph Adolphs, Kohitij Kar	Developing a non-human primate model to dissect the neural mechanisms of facial emotion processing relevant in autism spectrum disorder	Facial expressions are fundamental to human social interaction, allowing us to comprehend emotions, moods, and intentions. Accurate recognition and interpretation of these expressions are vital for effective communication, conflict resolution, and empathy. Therefore, differences in facial emotion judgments can pose significant challenges. Prior work has demonstrated that autistic adults significantly differ from neurotypically developed adults in comprehending social cues, particularly those conveyed through facial expressions. A mechanistic understanding of the underlying neural correlates of such behavioral mismatches is key to designing efficient cognitive therapies and developing other approaches to help autistic individuals. One major roadblock toward achieving this is that heterogeneity and idiosyncrasies are commonplace across behavioral reports in autism, including facial affect processing. The inability to parsimoniously explain such heterogeneous findings prevents us from designing more efficient follow-up experiments to probe the underlying mechanisms further. In addition, the differences in facial emotion judgments between neurotypical (NT) and autistic (ASD) adults are often interpreted with inferential models (e.g., psychometric functions) that base their predictions on high-level categorical descriptors of the stimuli (e.g., overall facial expression levels of “happiness”, “fear” and other primary emotions). Such modeling efforts ignore an important source of variance produced by the individual image-level sensory representations of each stimulus being tested. Kar (2022) demonstrated that differences in behavior between NT and ASD groups can be more prominently identified at the image-level. Therefore, in this study, we develop an image-level behavioral framework to study facial emotion recognition (across 6 basic emotions). We curated a selection of images (n=360) from the Montreal Set of Facial Displays of Emotion (MSFDE), which showcases emotional facial expressions (6 emotions) across various ethnicities (n=3) and genders. We measured behavior across human subjects (ASD and NT) while they performed a binary match to sample emotion discrimination tasks. Preliminary results show a weak but significant difference in the correlation between the image-level behavioral error patterns between NT and ASD. These data provide us with a critical behavioral benchmark. We have addressed the heterogeneity issue by specifically identifying the shared variances in the image-level behavioral metrics that we developed. Therefore, in our future efforts to identify neural markers, we will specifically focus on this shared variance. To further facilitate neurally mechanistic investigations, we have performed large-scale neural recordings across the inferior temporal (IT) cortex of rhesus macaques (n=3, ~200 neural sites). Preliminary results show that consistent with the prediction from Kar (2022), macaque IT-based decodes of facial emotion responses are more correlated with NT than ASD behavior. In sum, our study aims to synthesize a novel framework for combining non-human primate neural investigations with the behavioral phenotype of autism.

Shouyu Ling	shl329@pitt.edu	Max Kramer, Sophia Robert, Christina Patterson, Michael Granovetter, William Welch, Arish Alreja, Avniel Ghumann & Marlene Behrmann	Neural Dynamics of Category-Specific Visual Perception in Adolescents: Insights from sEEG	Stereoelectroencephalography (sEEG) has emerged as an invaluable tool in neuroscientific research, shedding light on the spatial and temporal complexities of neural activity. While its primary use has been to map epileptic networks (David et al., 2011; Englot et al., 2016), the role of sEEG in cognitive neuroscience, especially in understanding visual processing, is gaining increased recognition. sEEG's capacity to probe deep cortical and subcortical regions offers unprecedented insights into the processing and discrimination of visual stimuli. Such revelations have deepened our understanding of neural hierarchies, the flow of visual information, and the temporal intricacies of human cognition (Lachaux et al., 2012). Despite the emphasis on adult populations in previous studies, the neural representations of visual perception during adolescence remain largely unexplored. Given the dynamic neural changes occurring during teenage years, understanding visual perception during this phase is crucial. This research seeks to address this gap by introducing an sEEG dataset from a teenage cohort. We collected sEEG data from four teenagers while they viewed images from four distinct categories: faces, houses, words, and objects. This approach allowed for an evaluation of category-specific neural activity and its temporal dynamics. We employed MVPA to identify and characterize areas of the brain that are sensitive to each category. Furthermore, we delved into the broadband profile of these signals to understand the underlying spectral properties associated with each category. In future endeavors, a key aim is to contrast these teenage data with corresponding adult datasets. Such a comparison promises to highlight potential developmental shifts in neural category processing, offering a deeper understanding of the maturation processes that shape human cognition.

Simran Rooprai	srooprai@yorku.ca	Harsimran Dogra, Joseph FX DeSouza, Jenna Smith-Turchyn, Emily D'Alessandro, Nicole Anderson, Karolina Bearss	Assessing Gait and Cognitive Outcomes in Individuals with Parkinson's Disease After a Six-Year Dance Intervention	Introduction: Parkinson's disease (PD) is a neurodegenerative disorder primarily characterized by several distinctive motor symptoms, such as tremors, bradykinesia and rigidity. Gait disturbances are also common in the later stages of this disease. Although there is a greater emphasis on the motor manifestations of PD, the presence of mild cognitive impairment during the early stages may pose a major risk factor for developing dementia as the disease progresses as well. Research has shown that physical activity may delay the progression of PD by invoking neuroprotective effects. In a recent study, Bearss and DeSouza (2021) provided evidence that extended periods of high intensity exercise may induce neuroprotective effects, which can help to improve motor and non-motor symptoms among individuals with PD. As gait disturbances and impaired cognition are primarily exhibited in the later stages of the disease, examining their relationship would provide better insight on the neuroprotective effects of dance therapy for PD. Examining these neuroprotective benefits would provide significant advancement in the field of brain disorders and management. Methods: Data were collected from 2014 to 2017 during a longitudinal study conducted by Bearss and DeSouza (2021). The initial study only reported on data collected during this time period, although data collection was ongoing for a total of six-years from 2014 to 2019. Informed consent for participation and secondary analysis was obtained from all participants. Participants with PD (PwPD; n=44; 62% male; Mage = 69.43) and controls (n=18; 46% male; Mage = 61) attended a minimum of one dance class, lasting 1.25 hours, between 2014 to 2019 at the National Ballet School in Toronto, Ontario, Canada (see Bearss & DeSouza, 2017 for dance intervention procedure). Measures: The Mini-Mental State Examination (MMSE) and the Movement Disorders Society–Unified Parkinson's Disease Rating Scale (MDS–UPDRS III) were used to collect cognitive and motor data from all participants before each dance session. Electroencephalogram (EEG) was performed pre- and post-dance intervention as well. Proposed Results: We expect to see preserved or improved cognitive outcomes in PwPD who underwent the six-year dance intervention. By comparing the MDS-UPDRS III scoring to the MMSE scores collected, we expect to see a positive correlation between gait and cognition. Future Directions: This study provides preliminary results for future studies to be done. Based on the literature, we plan on continuing with a dance intervention with different measures. The Montreal Cognitive Assessment (MoCA) will be used going forward to better assess cognitive changes. Additionally, to measure cortical thickness, MRI data will be collected at various time points. Neuroimaging data should exhibit an increase in cortical thickness among the temporal and parietal regions as such areas are substantially thinner in PwPD.

Takao Fukui	takao-fukui@tmu.ac.jp	Mingze Zhang, Akira Hasegawa	Effects of short-term chopsticks usage on kinematics of subsequent reach-to-grasp and uplift movements: Relation to object familiarity and autistic traits	Tool-use is a fundamental ability of human kind, and chopstick manipulation is one of the quotidian motor behaviors for East Asians. Previous studies (e.g., Cardinali et al., 2009) demonstrate that tool use induces morphological updating of the body schema, which is a dynamic representation of the relative positions of body parts for planning and executing movements (e.g., Schwoebel & Coslett, 2005). It is also known that individuals with Autism Spectrum Disorder (ASD) show different patterns from typically developing (TD) peers (e.g., Bhat et al., 2011). It remains unclear (1) whether short-term chopsticks usage by Japanese people, who are familiar with chopsticks, affects the kinematics of subsequent manual reach-to-grasp movements and (2) whether and how the performance is modulated by autistic traits. The present study investigated these two questions by manipulating object familiarity in a sample of TD university students using the Japanese version (Wakabayashi et al., 2004) of Autism Spectrum Quotient (AQ) test (Baron-Cohen et al., 2001). 40 right-handed Japanese males participated in the experiment. The participants were divided into two groups (20 participants each): SUSHI group and GRAY group. While the participants in the SUSHI group reached for and grasped three types (small, medium, large) of soft mockups of SUSHI samples which are assumed to be familiar to Japanese people, those in the GRAY group reached for and grasped the gray objects that only reproduced their geometrical structures without the visual features. The experiment was comprised of two sessions, each of which consisted of three blocks. The first session consisted of three hand-grasping blocks (i.e., PRE, HAND, POST). After at least one day apart the participants did the second session, which consisted of hand-grasping tasks in the first and third blocks and chopsticks usage block in the second block (i.e., PRE, TOOL, POST). Peak grip aperture (PGA) and the difference between the grasp-end time and the time of lifting initiation (transition time) were calculated. We found a significant decrease of the PGA from PRE to POST in the TOOL session of GRAY group (but not SUSHI group), and no significant differences of the transition time between conditions. The Pearson correlation coefficient between transition time and AQ score was also calculated. A marginally significant positive correlation between transition time in the POST phase and AQ scores was found in the GRAY condition while no significant correlation between them was found in the SUSHI condition. The present study suggests that the extent of the modification of body schema after chopsticks usage in Japanese participants is dependent on the object familiarity and the participant’s autistic traits.

Tasfia Ahsan	ahsant@my.yorku.ca	Erez Freud	From 2D to 3D: importance of depth information to high level vision	In our three-dimensional world, perceiving and interacting with objects and people at varying distances is essential for our survival and success. Despite the significance of depth perception in our daily lives, vision research has predominantly focused on two-dimensional (2D) stimuli, such as static images. This project sought to elucidate how the human brain processes three-dimensional (3D) information in healthy adults. Functional magnetic resonance imaging (fMRI) was employed to investigate whether the addition of binocular depth cues enhances neural responses to established categories of visual stimuli. Participants laid in the MRI and viewed images representing categories such as faces, scenes, objects, bodies, and tools, presented in both 2D and 3D conditions. In the 2D condition, the same image was presented to both eyes, eliminating binocular disparity. In contrast, the 3D condition involved presenting stereoscopic images to each eye from viewpoints aligned with the average interpupillary distance, providing a more realistic representation of the objects. Data from five participants were analyzed using fMRI pipelines and software (fMRIPrep, TEDANA, BrainVoyager), followed by representation similarity analysis (RSA) iteratively applied to all possible runs of the data. Our findings revealed a significant 3D-2D effect, with a closer examination of the time course indicating that 3D presentation enhanced neural activation for some categories more than others. Notably, several brain regions exhibited robust responses to 3D representations of human bodies, including the body-selective cortex and superior parietal regions associated with embodiment. The variability in 3D-specific activation patterns across categories and regions suggests that these outcomes cannot be attributed solely to low-level properties such as binocular disparity.

Teodora Neagu	tadelen@my.yorku.ca	Rebecca L Hornsey, Arleen Aksay, Laurie M. Wilcox	Integration of motion parallax and stereopsis for surface segmentation	Motion parallax can support depth percepts comparable to those generated by stereopsis. However, recent work shows that when motion parallax and stereopsis are presented in combination in depth estimation tasks, observers tend to veto depth information from motion parallax, relying on stereoscopic information alone. One potential reason for this is that the depth estimation task biases the observers to use stereopsis. Here, we used a depth segmentation task to explore the integration of these two cues. Observers were presented with two curves occupying frontoparallel planes; one curve was a duplicate of the other rotated at 180°. To assess the role of complexity, we created two types of curves of low and high amplitude to manipulate the number of intersections between the curve pair. Two indicator probes were shown on either side of the curves, and were aligned with the same or different curves; the observer's task was to judge which was the case (same or different). The probes and stimuli were positioned 100cm from the observers. In the first experiment, each observer was run in 3 counterbalanced blocks, corresponding to the three cue conditions: 1. stereopsis alone, 2. motion parallax alone, 3. stereopsis and motion parallax combined. Our results show that in the high complexity condition, motion parallax was less effective than the stereopsis only and combination conditions. Additionally, we found that the JNDs for the stereopsis condition were similar to the combined condition. In Experiment 2, we evaluated whether the range of motion was responsible for the poorer performance in the motion parallax condition by testing a larger range of self-motion distances. We found no effect of this manipulation. In subsequent studies we evaluated if having to move disrupted performance by keeping the observer stationary and adding object motion. We found that performance was poorer in all conditions compared to Experiment 1 and stereoscopic thresholds were elevated substantially more than in the combined or motion only conditions. We propose that these outcomes are mediated by both uncertainty (noise) and the level of experience (experts vs. novices). In sum, our results show that models of cue integration must consider individual differences in expertise as this plays a significant role how cues are weighted. Further the results consistently show that performance in the combined cue condition is better than predicted by simple weighted averaging of motion and stereopsis.

Tim Henley	henleyt@yorku.ca	Timothy M. Henley, Jason Pina, Josh Tindell, Jérôme Lecoq, Natalia Orlova, Shiella Caldejon, Blake Richards, Joel Zylberberg	Reconstructing Pyramidal Neurons Recorded From Awake Behaving Subjects	Mounting evidence supports the idea that learning is driven by difference between predictions about upcoming sensory inputs and the incoming sensory inputs (the unexpected or unpredictable aspects of such inputs). Predictions about sensory inputs (e.g., signals from associative regions to sensory regions) arrive primarily in the superficial layers of the cortical sheet, where they are received by the distal apical dendrites of pyramidal neurons. On the other hand, incoming sensory inputs primarily target the region closest to the cell body (or soma). However, apical inputs can strongly drive neuronal firing when they are coincident with inputs near the cell bodies. *Hypothesis 1: When predictions match sensory inputs, there should be increased coupling between the apical dendritic and somatic activity. Alternative theoretical work suggests that the difference between predicted and actual somatic activity may be computed in the apical dendrites, triggering disinhibition at the apical trunk. *Hypothesis 2: When predictions do not match sensory input, there should be increased somato-dendritic coupling and increased plasticity. To our knowledge, no studies have examined somato-dendritic coupling during predictive learning. To fill this critical knowledge gap, my colleagues and I have been collaborating with the Allen Institute for Brain Science (AIBS) to conduct near-simultaneous recordings from the distal apical dendrites and somata of GCaMP-expressing pyramidal neurons using 2 photon calcium imaging. We are recording from pyramidal neurons in both superficial (Layer 2/3) and deeper (Layer 5) cortical layers, as somato-dendritic coupling and the functional roles in predictive learning may differ between these layers. In each subject, we imaged near-simultaneously on 8 planes, from the distal apical dendrites and somata in 4 different regions of mouse visual cortex: V1, AM, LM, and PM. Observing somato-dendritic coupling in both lower- and higher-order regions of mouse cortex is critical for testing the generality of the two competing hypotheses across the visual hierarchy. Linking distal apical dendrites with their somata in the pyramidal neurons from our experimental recordings is essential to the analyses of somato-dendritic coupling during predictive learning. Common approaches for linking ROIs between imaging planes depend on correlations between ROIs, which would confound these analyses. Instead, we collaborated with technicians at the AIBS throughout 2022 to develop novel methods of acquiring high-resolution z-stacks in awake behaving mice (sets of cross section images spanning the volume of the neurons in our experimental recordings). I adapted various convolutional neural networks to remove shot noise and segment these images. By correcting for motion, combining multiple denoised images per plane and aligning across recording sessions I have been able to isolate the structure of large neuron populations in our z-stacks. I developed a novel contrast adjustment algorithm which proved critical to resolving the activity dependent neural structural signal. I also developed methods to match dendritic and somatic ROIs from our experimental data to our reconstructions. Based on the above experimental and data processing techniques, I made it possible to successfully trace distal apical dendritic compartments to their somata across hundreds of planes in our z-stacks.

Veronica Nacher	vnacherc@yorku.ca	Parisa Abedi-Khoozani, Harbandhan Arora, Xiaogang Yan, Hongying Wang, John Douglas Crawford	A cortical mechanism for eye-head-hand coordination: lPFC 'gaze' signals encode future head and hand motion during visually guided reach.	Most of what we know about sensorimotor neuroscience comes from the study of single effector control, such as eye or hand motion, or relatively stereotypical behaviours such as eye-head coordination for gaze shifts. However, real-world behaviour requires the adaptive coordination of multiple effectors for different circumstances. For example, eye-head coordination patterns change during reach, with gaze movement followed by enhanced head (and hand) motion toward the reach target (Arora et al., 2019). The high-level cortical mechanisms for such strategies (particularly for head control) are unknown. We investigated this by recording from the lateral prefrontal cortex (lPFC; spanning Brodman areas 45, 46, & 8a) while two Rhesus monkeys performed head-unrestrained reaches toward visual targets. Many (208 / 499) task-related neurons showed time-locked gaze- (and later reach) related responses, but surprisingly these 'gaze' responses disappeared (35 / 84 neurons) or diminished during gaze shifts toward the same targets without reach. Further, an in-depth spatial analysis (based on model fits to neural response fields) confirmed that these 'gaze' responses were not what they appeared to be. In direct contrast to the saccade system (Sajad et al. 2015, Sadeh et al. 2015, Bharmauria et al. 2021), gaze displacement models provided the worst fits to the data. Instead, lPFC 'gaze' neurons preferentially coded skeletomotor motion, either future head (49%) or hand (33%) motion, with reach codes predominating later in the task. This is an important demonstration that signal timing does not always reflect spatial tuning in the same neurons. We conclude that many lPFC 'gaze' responses are not involved in gaze control, but rather reflect gaze inputs that trigger complex head-hand repertoires: in other words, a high-level neural mechanism for eye-head-hand coordination.

Viswajit Vembukumar	vkumar8@yorku.ca	Nikolaus Troje	Evaluating gaze perception with simulated motion parallax	Recently, the frequency of human interaction mediated by screens has increased exponentially. The prevalence of screen-based communication makes it important to understand how a person's gaze is perceived when viewed on screen. A better understanding of screen-based gaze perception would facilitate the enhancement of communication tools, thus increasing communication efficiency. An abundance of literature demonstrates that motion parallax is an important depth cue that enables participants to perceive objects more accurately in real and virtual worlds. It allows us to better understand the perception of faces in screens and perhaps even enhance our perception of them. If adding motion parallax can increase the sensitivity to gaze perception, then cheaper and more widely accessible solutions could be developed to integrate the depth cue into standard video communication, enhancing communication efficiency. This study aims to examine whether the addition of simulated motion parallax increases participants' sensitivity to the gaze direction of faces on a screen. By using motion capture technology to track a user's head location, we can generate a virtual camera whose movement in a virtual environment corresponds with the user's own motions which allows the image on the screen to be dynamically rendered based on their position. The study examines two conditions: one using simulated motion parallax, and another using static images on a screen. In the simulated motion parallax condition, the avatar's head and eye gaze are adjusted to a set of angles and users are asked to move themselves into the line of sight of the avatar. Since the screen is dynamically rendered, they will be able to establish eye contact with the avatar. Once they reach a location at which they perceive eye contact their head location position is recorded and stored. The angular difference between the participant's head location and their expected position will then be analyzed. In the static condition, the avatar head will be rotated, and participants must indicate whether it is looking to their left or right. Both datasets are modelled by normal distributions such that means (accuracy) and variance (precision) of eye gaze perception can be assessed and compared. After analyzing the data from participants at York University, it was found that the addition of motion parallax made participants significantly more precise ($p=0.0001$). In the static condition, participants had an average mean of 0.87 degrees and an average standard deviation of 6.2 degrees. In the motion parallax condition, the average mean was 3.5 degrees and the average standard deviation was 3.1 degrees. The difference between the standard deviations in the two conditions was found to be statistically significant ($p=0.02$).

Wanyi Lyu	wanyilyu@yorku.ca	Jeffrey Schall	Resolving stages of processing in visual search: Investigation of behavioral errors and the underlying covert operations with frontal eye field neurophysiology double factorial design	Behavior is the outcome of covert perceptual, cognitive, and motor operations that can be described by mathematical models and are produced by brain systems comprised of diverse neurons. Behavioral errors can expose specific flaws of covert operations and the associated neural substrates. Using the logic of selective influence, we designed a task to separately modify two operations necessary for visual search—singleton localization and stimulus-response mapping. Macaque monkeys performed a GO-NOGO color singleton visual search task. Singleton localizability was manipulated by varying the similarity between singleton and distractor colors. Stimulus-response mapping was manipulated by varying the discriminability of the GO/NOGO cue, which was elongation of the search array elements. The response times of both monkeys were modified selectively by the 2x2 (High vs Low localizability) x (High vs Low discriminability) manipulations. Performance on correct trials was analyzed using System Factorial Technology (SFT) to diagnose the alternative architectures of the two operations (serial or parallel, self-terminating or exhaustive). However, the validity of this diagnosis is not necessarily secure when error rates are moderate. Additionally, mathematical models can theorize ways performance may fail, but have no information about which mechanism/operation made those errors. To inform mathematical model and to identify the neural source of the error, we examined monkeys' error performance and the corresponding neural mechanisms. Monkeys made two key errors: on GO trials, monkeys occasionally shifted gaze to a distractor due to unsuccessful localization, constituting a 'GO error'. Unlike in other visual search tasks, monkeys very rarely shifted gaze to the correct target after misdirection errors. On NOGO trials, they failed to inhibit their saccade towards either the singleton or the distractors, producing 'NOGO errors'. NOGO errors reflect failure in discrimination alone or both operations, respectively. We probed the neural sources of these error saccades using single-unit spiking in frontal eye field of both monkeys. Neurons representing stimulus salience were distinguished from neurons mediating saccade preparation. We found that GO errors happened when visual salience neurons selected a distractor as if it were the singleton. Saccade preparation neurons were active for GO and NOGO errors when monkeys looked into the respective movement field. The findings suggest that GO error may result from salience neurons misrepresenting distractor items as the singleton. NOGO error to singleton may arise from incorrect discrimination by saccade preparation neurons whereas NOGO error to distractor may arise from inaccurate response from both neuron type. In summary, single-unit activity during error allows us to make inference about the operation(s) each neuron participate in and their role in error performance. The convergence of performance and neural results on error trials provides evidence so that distinct operations during visual search can be resolved. Supported by NIH RO1-EY08890, F32-EY028846, T32-EY007135, and P30-EY008126.

Xianze Meng	mengxz@yorku.ca	Alexander Barnett, Joseph FX DeSouza	Structural & functional connectivity correlation between PMN and Hippocampus during segmentation and recall of naturalistic stimuli	Abstract Introduction. Investigating event segmentation in real time using naturalistic stimuli is an emerging field for vision and memory-related research to reveal how episodic memory organizes while encoding and reiterated during recall. Previous work by Barnett et al. unveiled that the high functional connectivity (FC) between the hippocampus (HC) and the posterior medial network (PMN) during an event offset is necessary for the naturalistic events to be recalled regardless of the delay period. This study aims to reveal the white matter (WM) pathway that promotes communication between HC and PMN with the hope of systematically correlating SC and FC with memory performance. We selected the fornix and bilateral posterior cingulum as the most potential candidates for their anatomical proximity. We hypothesized that high SC features, such as high fractional anisotropy (FA), high axial diffusivity (AD), and low radial diffusivity (RD) within the fornix and bilateral cingulum should reliability predict high FC between HC and PMN and subsequently better naturalistic memory performance. These two tracts' anatomical proximity rendered them the most possible candidate to support HC-PMN communication. Methods. Young, neurologically healthy participants (n=24) were split into four equal groups to encode and recall two movies in different delay orders. After encoding one movie, the participant freely recalled the movie before viewing another one on the same day and then recalled that movie two days later. We performed structural and diffusion-weighted MRI after the participant's initial visit at the first encoding session. During each verbal recall session, we collected the fMRI from the participants with their recall transcripts. We then extracted the FA, AD, and RD for the fornix and bilateral cingulum from probabilistic tractography on the diffusion-weighted MRI as the SC metrics. The event-specific HC-PMN FC was directly obtained from Barnett et al. for each participant. To examine if these two tracts were responsible for HC-PMN FC, we performed cross-subject linear correlations between FA/AD/RD versus FC. We should expect a reliable positive linear relationship between AD/FA and FC and a negative linear relationship between RD and FC if high SC predicts high FC. Results and Conclusions. We found a positive SC-FC relationship in the left cingulum, in which the increase in AD and decrease of RD reliably predicted an increase in FC at recalled events. A weaker SC-FC correlation was observed in the right cingulum; only AD increase predicted FC increase in the right cingulum, but the RD-FC correlation lacked significance. However, the FA-FC correlation was insignificant on either side of the cingulum. Individual differences in FA may not affect FC in healthy subjects as much as in neurodegeneration studies. Surprisingly, none of the fornix metrics displayed a linear correlation pattern with FC changes, which suggests the fornix may not directly participate in episodic recall despite the anatomical position. We can confidently conclude from our results that the bilateral cingulum actively facilitates HC-PMN communication for successful naturalistic stimuli recall. The fornix may support but does not directly carry episodic information between HC and PMN.

Xue Teng	xueteng@yorku.ca	Laurie M. Wilcox, Robert S. Allison	Increasing motion parallax gain compresses space and 3D object shape	Humans rely on visual, proprioceptive and vestibular cues to perceive depth and distance when moving about the world. In the physical world, these sources of information are consistent. However, what happens if we receive conflicting information about how far we have moved? A previous study reported that at distances of 1.3 to 1.5 m, portrayed binocular 3D shape was not affected by motion gain; however, apparent distance and monocular depth settings were influenced. In our study, we extended the range of distances to 1.5 to 6 m. A VR headset was used to display gain distortions binocularly and monocularly to the right eye. Observers swayed from side to side through 20 cm at 0.5 Hz to the beat of a metronome. The simulated virtual motion was varied by a gain of 0.5 to 2.0 times the physical motion. Observers first adjusted a vertical fold stimulus until its sides appeared to form a 90-degree angle. Upon each adjustment input, the fold then disappeared and they indicated its remembered distance by adjusting the position of a virtual pole. As gain increased, monocular observers provided increasingly compressed fold depth settings at 1.5 and 3 but not at 6 m. Under binocular viewing, increasing gain compressed distance but not object shape settings. To ensure that the weak binocular effects were not due to failure to perceive the gain, we separately assessed gain discrimination thresholds using the fold stimulus. We found that observers were sensitive to the manipulation over this range and tended to perceive a gain of 1.1 as having no motion distortion under both viewing conditions. It is clear from our data that monocular viewing of kinesthetic/visual mismatch results in significant variations in portrayed depth of the fold. These effects can be somewhat mitigated by increasing viewing distance, but even more so by viewing with both eyes.
Yael Goldstein Marcusohn	ygolds@gmail.com	Rahaf Asaad, Leen Asaad, and Erez Freud	Shape sensitivity in the dorsal pathway does not depend on attention	The two visual pathways model posits that visual information is processed through two distinct cortical systems: The ventral pathway promotes visual recognition, while the dorsal pathway supports visuomotor control. Recent evidence suggests the dorsal pathway is also involved in shape processing and may contribute to object perception, but it remains unclear whether this sensitivity is independent of attentional mechanisms that were localized to the same cortical pathway. To address this question, we conducted two fMRI experiments that utilized different parametric scrambling manipulations in which human participants viewed novel objects in different levels of scrambling, and were instructed to attend to either the object or to another aspect of the image (e.g., color of the background). The results showed similar large-scale organization for shape processing along both pathways regardless of the focus of attention. Particularly, in both pathways and across tasks, shape sensitivity increased from early visual cortex to extrastriate cortex but then decreased in anterior regions. These findings support the idea that shape processing relies on a distributed set of cortical regions across the visual pathways, independent of attentional processes.

Yara Mary Iskandar	misk22@my.yorku.ca	Christopher Lee, Sebastian Bosse, Peter J Kohler	Spatial Mechanisms Mediating Visual Responses to Symmetries in Textures	Symmetries are present at many scales in natural scenes. Humans and other animals are highly sensitive to visual symmetry, and symmetry has been shown to play a role in numerous domains of visual perception. Brain imaging studies have demonstrated that several regions in visual cortex exhibit robust and precise responses to symmetry. The current study explored the mechanisms underlying these responses, by measuring Steady-State Visual Evoked Potentials (SSVEPs) using high-density electroencephalography. Our stimuli were a class of regular textures, known as wallpaper groups: 17 unique combinations of symmetry types that represent the complete set of symmetries in 2D images. We focused on wallpaper groups PMM, which contains bilateral reflection symmetry, and P4, which contains four-fold rotation symmetry. Our SSVEP approach allows us to measure brain responses that are specific to the symmetries within each group. We measured these responses in two experiments, one (n=40) testing the influence of spatial frequency content and another (n=14) testing the influence of the repeating lattice structure that tiles the plane in all wallpaper groups. Exemplars for the spatial frequency experiment were generated based on log-domain band-limited random noise patches with center frequencies between 1 and 8 cycles-per-degree. For the lattice experiment, spatial frequency was kept constant at 2 cycles-per-degree and the ratio of the lattice to the overall wallpaper area varied between 1/12 and 1/2. Symmetry-specific responses were weaker overall for rotation compared to reflection, consistent with prior studies, but the manipulations had broadly similar effects for both: Responses were strongest at low spatial frequencies and weakened rapidly with increasing frequencies. The lattice manipulation had less dramatic effects, but results suggest that responses are stronger at lower ratios. Responses to reflection and rotation may thus depend on a similar mechanism that is highly dependent on spatial frequency and benefits from a repeating lattice structure.
Yeganeh Gharedaghi	ygn.gharedaghi@gmail.com	Gene Cheung, Xianming Liu	Retinex-based Image Denoising / Contrast Enhancement using Gradient Graph Laplacian Regularizer	Images captured in poorly lit conditions are often corrupted by acquisition noise. Leveraging recent advances in graph-based regularization, we propose a fast Retinex-based restoration scheme that denoises and contrast-enhances an image. Specifically, by Retinex theory we first assume that each image pixel is a multiplication of its reflectance and illumination components. We next assume that the reflectance and illumination components are piecewise constant (PWC) and continuous piecewise planar (PWP) signals, which can be recovered via graph Laplacian regularizer (GLR) and gradient graph Laplacian regularizer (GGLR) respectively. We formulate quadratic objectives regularized by GLR and GGLR, which are minimized alternately until convergence by solving linear systems---with improved condition numbers via proposed preconditioners---via conjugate gradient (CG) efficiently. Experimental results show that our algorithm achieves competitive visual image quality while reducing computation complexity noticeably.

Yousif Kashef Alghetaa	yousif95@yorku.ca	Simon Kornblith, Kohitij Kar	Quantifying Alignment between Human and Machine Explanations: A Novel Approach Using Explanation Masked Images	Understanding how artificial neural network models make decisions is becoming increasingly relevant, especially as demands from institutions like governments grow. While numerous Explainable Artificial Intelligence (XAI) methods exist, each offers a slightly different interpretation of a model's inner workings. To address this, Kar et al. (2022) proposed that one approach to benchmark the goodness of the machine explanations might be their alignment with human explanations. To achieve this, a key assumption is that there are tools to reliably estimate the primate visual system's explanations. However, prior research indicates that psychophysical methods like 'bubbles' and classification images used to probe human explanations suffer from multiple shortcomings. In this study, we focus on object discrimination behavior to propose and validate a novel method to approximate the alignment between human and machine explanations. We first define a target (whose explanation we seek, e.g., ResNet-50) and a reference model (a robust model, whose explanation serves as a gold standard for evaluating the goodness of explanations for the target model). While ultimately, we want to use humans as the reference model, we need an image-computable, fully differentiable model (e.g., AlexNet) to develop and validate our method. Explanation of a model's output typically refers to how different features (pixels) of the input image contribute to the model's output (feature attribution). It is available as a heat map. To test our method, we first estimated the ground truth in how similar the explanations of ResNet-50 (for 200 natural images,on object discrimination) are to AlexNet, by directly comparing the feature attribution maps produced by ten different explanations (e.g., saliency, integrated gradients) using a few distance metrics (e.g., L1-norm, L2-norm). This step produces the ground-truth rank order of how explanations compare across ResNet-50 and AlexNet. Once such a ground truth has been established, we now desire a human-compatible procedure (“proxy”) to revive the ground-truth rank order. So, we first generated filtered versions of the original images by only retaining the top percentiles (50, 90, etc.) of the highly informative pixels (given the feature attribution map of each explanation) for ResNet-50. We refer to them as explanation masked images (EMI). We hypothesized that the way such images impact the behavior of two systems might be symptomatic of how similar the underlying explanations used to generate those images are. Therefore, next, we measured the behavioral accuracies of both ResNet-50 and AlexNet on the ResNet EMIs. We computed a value per image that approximates the average object discrimination accuracy for that image against all possible distractor objects. The correlation (AlexNet vs. ResNet-50) between these values for the EMIs on specific explanations gave the rank order of how similarly the models behaved on these images. A a strong correlation between these rankings and the ground truth would validate our approach. Indeed, we observed a significant positive correlation (SpearmanR=0.8, p=0.003). Our results now allow us to perform human behavioral testing on the EMIs to estimate the true goodness of the ResNet-50 explanations.

Zainab Haseeb	zainab.haseeb@mail.utoronto.ca	Silvia Guidi, Anna Kosovicheva, Benjamin Wolfe	Stimuli for sentence readability – a new, validated corpus for studies of digital reading	Reading is a complex process in which visual factors affect how well and how quickly we understand written content. Understanding the impact of the visual appearance of text in reading is essential for optimizing reading efficiency. To support in-lab investigations of visual factors in reading, we've curated a human-validated corpus of sentences that can be used for true-false classification tasks. This corpus was derived from GenericksKB, a repository containing 3.5 million internet-derived sentences. We filtered the dataset based on word count to include sentences with 6 to 8 words, and then further restricted the dataset based on character count, taking sentences between the 33rd and 66th percentile of character lengths (35- 42 characters). We further filtered these based on BERT classification scores, which assessed sentence usefulness and truthfulness, retaining the top 70% of sentences. Using WorldLex, an English word frequency database, we retained the top 17% of terms based on word frequency metrics to ensure broad understanding. To generate a matched set of grammatically correct false sentences, we shuffled the sentence topics between sentences (e.g., 'Beer is the collective memory of a people') and manually filtered them for sensitive content. Sentence validity was ensured through crowdsourced assessments of truthfulness and grammatical correctness. Participants rated the truthfulness of each statement on a continuous slider from 0 (completely false) to 1 (completely true) and indicated whether each sentence was grammatically correct. Participants ratings of sentence truthfulness revealed high inter-rater agreement (mean ICC of 0.99; n=79). True sentences were classified as the top 25% of the mean value distribution across participants while false sentences were classified as the bottom 25% of the mean value distribution. Lastly, we used sentence grammatical error counts to flag potential issues for manual review and correction. This corpus allows us to investigate how visual elements impact sentence-level reading, simplifying the study of the impacts of digital text appearance on reading. Specifically, we show how we can use psychophysical techniques to measure duration thresholds for true-false classification using these sentences. Variable fonts, which offer customizable font appearance within a single file, are ideal for this research. We use Roboto Flex, Google’s variable font to explore which font axes (e.g., stroke thickness, thinness, weight, width, and slant) affect reading. Our approach systematically manipulates font characteristics to uncover their impact on reading speed and gaze behaviour. By leveraging our validated sentence corpus and variable font paradigm, we aim to reveal how visual factors shape reading experiences. This research will deepen our understanding of reading efficiency and comprehension, informing text design for screens and enhancing reading skills.

Zoe Stearns	zstearn2@ur.rochester.edu	Martina Poletti	Comparing the temporal dynamics of pre-microsaccadic and pre-saccadic vision	Saccade execution is known to modulate visual perception in distinctive ways; sensitivity is briefly enhanced at the saccade goal location before the eyes start to move, and a drop in sensitivity follows at the commencement of the saccade. Recent work has shown that microsaccades induce similar modulations at the scale of the fovea. Here we compare the temporal dynamics of pre-microsaccadic and saccadic perception. Eye movements were recorded with a high-resolution digital Dual-Purkinjie Image eye tracker while subjects performed a 2AFC discrimination task. Subjects (N=6) were required to move their gaze to one of two possible locations surrounding a central fixation marker based on the direction indicated by the cue. Stimuli were flashed at both locations at a variable time before the gaze shift, either 0.3 deg or 5 deg away from fixation, in the microsaccade and saccade task respectively. At the end of the trial, subjects reported the orientation of the stimuli previously presented at the location indicated by a response cue. Our findings show that the pre-microsaccadic and pre-saccadic perceptual enhancements at the goal location are characterized by a comparable duration (59 ± 16 ms vs. 60 ± 12 ms, $p = 0.8$). Yet, the enhancement associated with microsaccades begins approximately 30 ms earlier than for saccades ($p < 0.01$). At both scales the enhancement is followed by a suppression that begins before the gaze shift onset. Differently from perceptual enhancement, suppression lasts 34 ± 27 ms longer for microsaccades than saccades ($p < 0.05$). These findings suggest that although saccades and microsaccades induce similar perceptual modulations, the enhancement and suppression unfold following different dynamics likely as a result of the longer latencies characterizing microsaccades.

Zoha Ahmad	zohahmad@my.yorku.ca	Krista Kelly, Erez Freud	Reduced perception-action dissociation in children with amblyopia	Visual illusions differentially affect perception and action, providing insights into the functional dissociation between vision for perception and vision for action. One such illusion is the Ponzo illusion where perceptual behaviors are modulated by the monocular depth cues, while grasping trajectories resist the effect of the illusion. Children with amblyopia (poor vision in one eye) provide the opportunity to examine an outstanding question - to what extent normal visual experience is critical for the development of the perception-action functional dissociation. Here, we examined the effect of the Ponzo illusion on perception and action in a group of children with amblyopia (n=20, 9.1±2.2 years old) compared to typically developed children (n=20, 9.2±2.1 years old). In the incongruent trials, two objects that differ in their real size (40 & 42 mm) were placed on an illusory Ponzo background to create a conflict between the real and the perceived size of the object and vice versa in the congruent trials. Control and amblyopic participants were asked to grasp (action task) and manually estimate the length of the object (perceptual task). Across conditions, we recorded hand kinematics using the Optitrack motion capture system. Both groups showed equal effect of the illusion in the perceptual task, such that the object placed on the far surface was perceived as longer compared to when placed on the near surface. In contrast, a clear differentiation between the groups was observed for the grasping task. While amblyopic participants' Maximum Grip Apertures (MGAs) were modulated by the illusion, the MGAs of typically developed children evaded the illusion and were scaled to the real size of the object regardless of the object's placement on the background. Taken together, our results provide novel evidence that atypical visual experience in childhood from amblyopia impacts the development of the dissociation between perception and action.