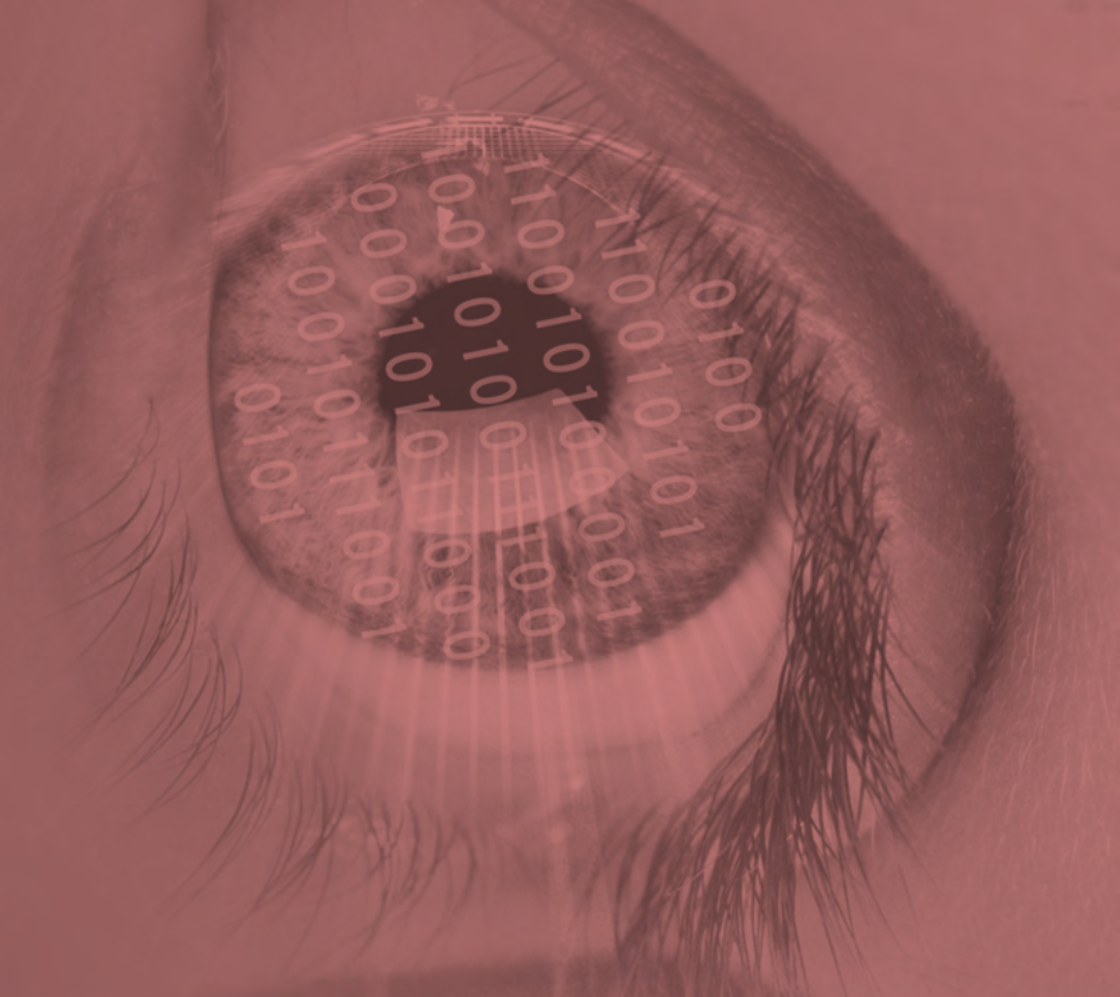


FROM PICTURE TO REALITY, FROM OBSERVER TO AGENT

VISION RESEARCH CONFERENCE
SECOND STUDENT CENTRE
YORK UNIVERSITY
JUNE 6-9
2022

YORK UNIVERSITÉ UNIVERSITY

VISTA VISION: SCIENCE TO APPLICATIONS
YORK UNIVERSITY

CANADA FIRST RESEARCH EXCELLENCE FUND

APOGEE CANADA FONDS D'EXCELLENCE EN RECHERCHE

Centre for Vision Research

VISTA VISION: SCIENCE TO APPLICATIONS

YORK UNIVERSITY

WHAT IS VISTA?

Vision: Science to Applications is a research program based at York University in Toronto. VISTA provides important graduate, post-doctoral and researcher funding opportunities to enable cutting-edge research. Funded in part by the Canada First Research Excellence Fund (CFREF), VISTA has created numerous projects and technologies that will help people live healthier, safer and more productive lives.

WHAT ARE THE BENEFITS OF BEING PART OF VISTA?

- Become a member of an interdisciplinary research community across the research areas of health, science, engineering, humanities and arts.
- Work with over 50 industry partners offering opportunities for collaboration, internship and career development.
- Join an international network pushing the boundaries of vision research in North America, Europe and Asia.

MASTER'S SCHOLARSHIPS

$10,000 per year for a max. of 2 years

2 competitions per year

PHD SCHOLARSHIPS

$10,000 per year for a max. of 4 Years

2 competitions per year

POST-DOCTORAL FELLOWS

$55,000 per year for a max. of 2 Years

$70,000 per year for Distinguished Post-Docs

2 competitions per year

Research allowance $7,500 per year
Networking allowance $1,000 per year

TRAVEL AWARDS

VISTA will support flight and accommodation for visiting scholars, students and fellows to/from academic institutions.

www.yorku.ca/vista | More to vision than meets the eye.
www.yorku.ca/cvr

The CVR conference 2022 connects the long history of picture perception with new display technologies that blur the differences between pictorial representations and the reality of the world that they attempt to depict. Controlled, principled visual stimulation is at the core of vision research. Techniques have evolved from painted or printed material, to computer screens that provide control and enable rapid changes of graphical contents that soon included moving pictures and stereoscopic displays, and further to today's high-fidelity VR/AR systems and the prospect of holographic displays.

Picture perception is interesting for two reasons. On the one hand, the ability to abstract the depicted contents from the physical object of the medium, as humans do, is a trait that only few other animals are capable of. There is a rich tradition within vision science that focusses on picture perception and connects vision research with arts history, the philosophy of perception, and cognitive science.

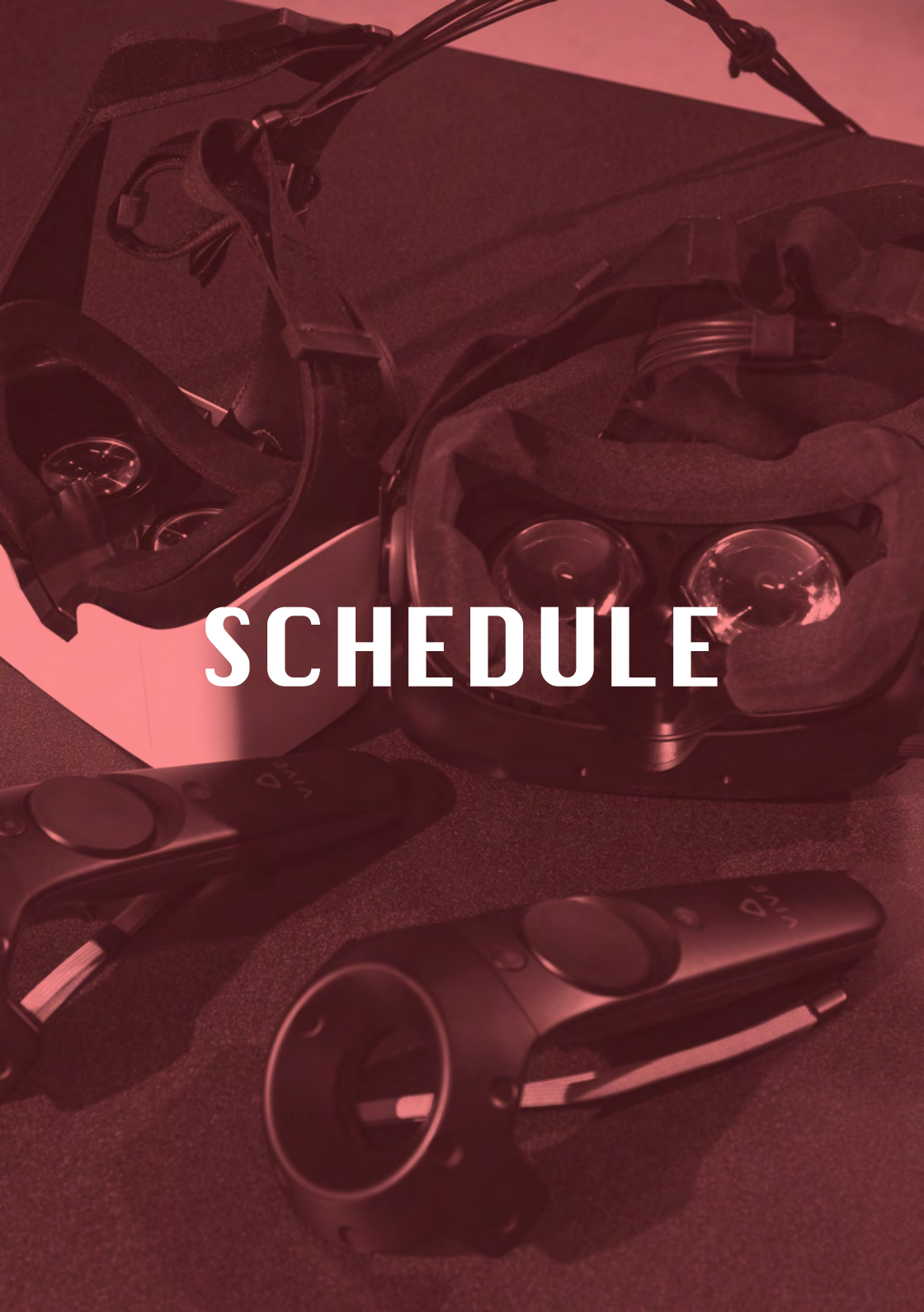
On the other hand, vision research has often used pictorial displays with the implicit assumption that they provide a valid surrogate for the visual stimulation experience in normal life. Consequently, the results obtained in the lab are expected to generalize into the real world. More recently, however, researchers began to challenge this assumption and started to compare picture perception and vision in the real world explicitly.

Virtual reality, augmented reality, and the prospect of light-field-based holographic displays provide excellent tools to that end. The defining feature, namely the ability to update the stereoscopic displays contingent with the user's movement, is more than an incremental increase in representational fidelity. It transports passive observers who look at pictorial space into active agents that become part of that space. It establishes visual presence and employs the user's visual system in ways much more similar to the ones in which we process the "real" visual environment. An industry exhibition during the meeting will showcase exciting new developments in display techniques and their applications.

The topic of CVR 2022 lends itself to celebrate the achievements of CVR's former director, Laurence Harris and recognize his decade of leadership of the centre. Reality constitutes itself in terms of consistencies between sensory modalities and the predictability of sensory stimulation in response to movement. Harris, while working on multisensory integration, space perception and self-perception in space, pioneered the emancipation from the picture domain. He started to work with real world stimuli, he adopted virtual reality for vision research long before others did, and he sends his participants into orbit to manipulate their sensory environment in ways that would not be possible anywhere on earth.

The Organization Committee:
Niko Troje, Joel Zylberberg, Denise Henriques, Michael Barnett-Cowan

Contents



SCHEDULE

MONDAY, JUNE 6

1:00 pm	**Registration**
4:00 pm	**Keynote Address** – **Kevin O'Regan**, Université Paris Cité What is it to perceive the world?
5:15 pm	**Poster Session 1**
6:30 pm	**Opening Reception**

Suggested locations for lunch on Tuesday–Thursday:

Several restaurants are located just across The Pond Road at **The Quad** residence building: https://quadatyork.ca/where-to-eat-on-campus/ [Google Maps]

There are also several restaurants at **York Lanes**, not too far from the conference centre: https://www.yorklanesmall.com/eateries [Google Maps]

Additional cafeterias are located at the **Dahdaleh building** [Google Maps] and **Central Square** [Google Maps]

TUESDAY, JUNE 7

9:10 am **Welcoming Remarks** – Conference Committee Chair, Niko Troje
CVR Director, Rob Allison; VISTA Director, Doug Crawford

9:40 am **Session 1 – Evolution and development of picture perception.**
Chairs: Patrick Cavanagh and Jennifer Steeves

9:40 am **Ludwig Huber**, University of Veterinary Medicine Vienna
Eye-tracking with dogs: Achievements and challenges

10:20 am **Jody Culham**, Western University
Closing the loop with Immersive Neuroscience: On the importance of real actions and feedback for brain and behavior

11:00 am **Coffee Break**

11:20 am **Jacqueline Snow**, University of Nevada, Reno
Comparing brain responses to real objects versus 2-D pictures: Emerging fMRI data from the adult brain and its implications for understanding the developing cortex

12:00 pm **Industry Presentations**
Host: James Elder

1–3 pm **Poster Session 2 / Lunch Break**

1–6 pm **Industry Exhibition**

3:00 pm **Session 2 – Looking at projections.**
Chairs: Marcus Brubaker and Erez Freud

3:00 pm **Heiko Hecht**, Johannes Gutenberg-Universität Mainz
The privatization of pictorial space

3:40 pm **Gabriel Greenberg**, University of California, Los Angeles
Picture perception and the radical variability of the visual system

4:20 pm **Coffee Break**

4:40 pm **John Kulvicki**, Dartmouth University
Non-spatial projections

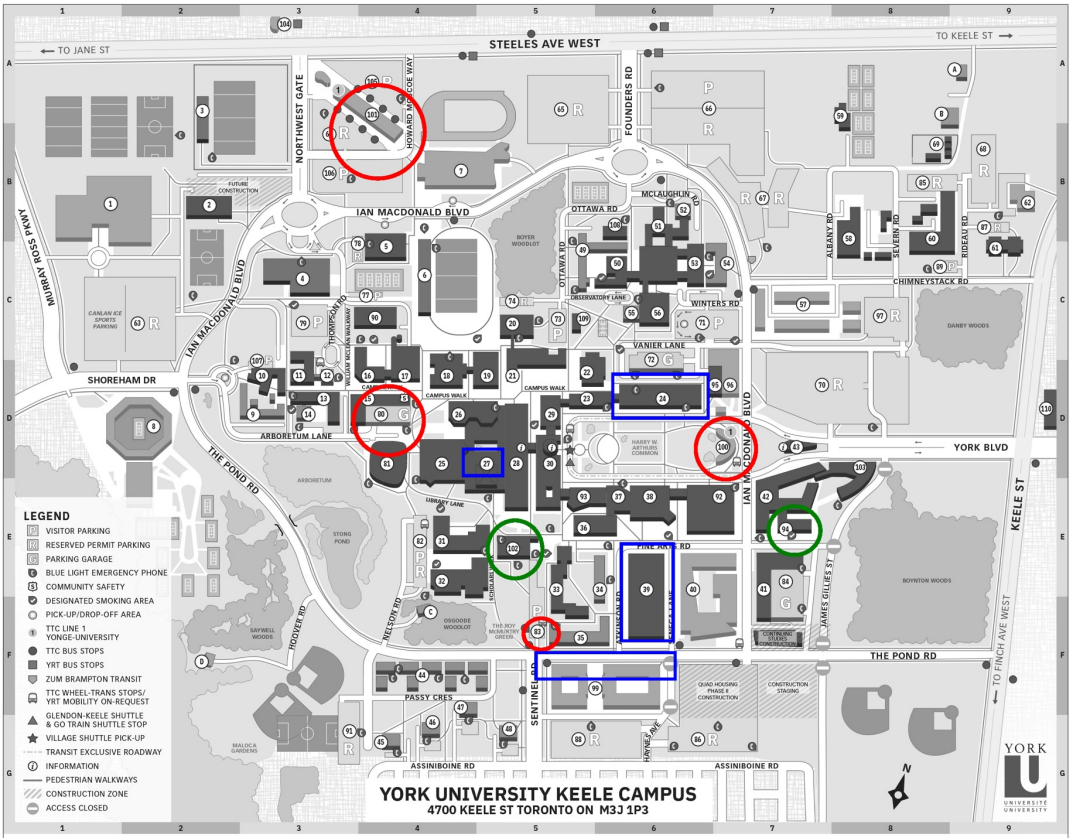
5:20 pm **James Elder**, York University
Single-view 3D perception in humans and machines

WEDNESDAY, JUNE 8

9:00 am	**Session 3 – Display systems: From drawings to holograms.**
	Chairs: Michael Brown and Laurie Wilcox
9:00 am	**Frank Steinicke**, Universität Hamburg B(l)ending realities
9:40am	**Austin Roorda**, University of California, Berkeley The sharpest, brightest, highest-contrast, most-impractical, gaze-contingent display ever built
10:20 am	**Coffee Break**
10:40 am	**Martin Banks**, University of California, Berkeley Binocular vision and oculomotor behavior in natural and VR-gaming environments
11:20 am	**Hong Hua**, University of Arizona Head-mounted light field displays for virtual and mixed reality
12–2 pm	**Poster Session 3 / Lunch Break**
2:00 pm	**Session 4 – Virtual reality in vision science.**
	Chairs: Michael Jenkin and Graham Wakefield
2:00 pm	**Mel Slater**, University of Barcelona Illusions and applications of virtual reality
2:40 pm	**Gabriel Diaz**, Rochester Institute of Technology The present and near future of mobile and XR-integrated eye tracking
3:20 pm	**Coffee Break**
3:40 pm	**Jennifer Campos**, KITE - Toronto Rehabilitation Institute - University Health Network Seeing, hearing, moving: Age-related changes to multisensory integration processes supporting mobility
4:20 pm	**Craig Chapman**, University of Alberta Gaze and Movement Analysis (GaMA) in real and virtual worlds

THURSDAY, JUNE 9

9:00 am	**Session 5 – Space perception.**
	Chairs: Rob Allison and Shayna Rosenbaum
9:00 am	**Elisa Rafaella Ferrè**, Birkbeck, University of London Vestibular system: From signal encoding to perception
9:40am	**Markus Lappe**, University of Münster Walking with others
10:20 am	**Coffee Break**
10:40 am	**Paul MacNeilage**, University of Nevada, Reno Sensory and motor signals mediating stationarity perception
11:20 am	**Dora Angelaki**, New York University Eye movements for active sensing and memory during visually-guided navigation
12:00 pm	**Awards & Announcements**
12:30 pm	**Lunch Break**
2:00 pm	**Session 6 – Multisensory perception.**
	Chairs: Doug Crawford and Denise Henriques
2:00 pm	**Michael Barnett-Cowan**, University of Waterloo Seas the day: Developing a VR exergame during COVID-19
2:40 pm	**Marc Ernst**, Ulm University Development of multisensory perception and displays
3:20 pm	**Coffee Break**
3:40 pm	**Charles Spence**, University of Oxford Coloured hearing, colour music, colour organs, and the search for perceptually meaningful correspondences between colour and sound
4:20 pm	**Michael Jenkin**, York University Growing old and the effects of multi-cue integration
5:15 pm	**Former CVR Director – Laurence Harris**, York University How I learned to stop worrying and love the pull of gravity
7:00 pm	**Conference Banquet**



CAMPUS DIRECTORY

places to eat | **parking lots and subway stations** | **conference space and hotel**

ACADEMIC, ADMINISTRATIVE & COMMERCIAL BUILDINGS

59	190 Albany Road	ALB	A8
110	4747 & 4751 Keele Street		D9
92	Accolade East	ACE	E7
93	Accolade West	ACW	E5
33	Atkinson	ATK	E5
17	Allan I. Carswell Astronomical Observatory, Petrie	PSE	D4
29	Behavioural Science	BSB	D5
41	**Bennett Centre for Student Services, Admissions**	BCS	E7
81	Bergeron Centre for Engineering Excellence	BRG	D4
24	Bookstore, York Lanes	YL	D6
69	Bus Compound		B8
10	Calumet College	CC	D3
27	Central Square	CSQ	D5
58	Central Utilities Building	CUB	B8
38	Centre for Film & Theatre	CFT	E6
16	Chemistry	CB	D4
26	Curtis Lecture Halls	CLH	D4
94	Executive Learning Centre	ELC	E7
21	Farquharson Life Sciences	FRQ	D5
23	First Student Centre	STC	D5
108	Founders Annex North	FAN	B6
109	Founders Annex South	FAS	C5
50	Founders College	FC	C6
31	Health, Nursing & Environmental Studies	HNE	E4
32	Ignat Kaneff Building, Osgoode Hall Law School	OSG	E4
36	Joan & Martin Goldfarb Centre for Fine Arts	CFA	E5
95	Kaneff Tower	KT	D7
61	Kinsmen	K	C9
19	Lassonde Building	LAS	D5
90	Life Sciences Building	LSB	C4
43	Lorna R. Marsden Honour Court & Welcome Centre	HC	D7
20	Lumbers	LUM	C5
51	McLaughlin College	MC	B6
11	Norman Bethune College	BC	D3
17	Petrie Science & Engineering	PSE	D4
60	Physical Resources Building	PRB	B8
103	Rob and Cheryl McEwen Graduate Study & Research Building	MB	D8
28	Ross Building	R	D5
25	Scott Library	SCL	D4
27	Scott Religious Centre, CSQ	SRC	D5
102	Second Student Centre	SSC	E5
42	Seymour Schulich Building	SSB	E7
2	Sherman Health Science Research Centre	SHR	B2
18	Steacie Science & Engineering Library	STL	D4
22	Stedman Lecture Halls	SLH	D5
13	Stong College	SC	D3
4	Tait McKenzie Centre	TM	C3
56	Vanier College	VC	C6
30	Vari Hall	VH	D5
39	Victor Phillip Dahdaleh Building	DB	E6
5	West Office Building	WOB	C4
15	William Small Centre	WSC	D4
53	Winters College	WC	C4
24	York Lanes	YL	D6

RESIDENCES & APARTMENTS

48	320 Assiniboine Road	AS2	G5
47	340 Assiniboine Road	AS4	F4
46	360 Assiniboine Road	AS6	G4
45	380 Assiniboine Road	AS8	G4
34	Atkinson Residence	AR	E6
9	Calumet Residence	CR	D3
49	Founders Residence	FR	C5
12	Norman Bethune Residence	BR	D3
44	Passy Gardens, 2–18 Passy Cres.	PASSY	F4
14	Stong Residence	SR	D3
52	Tatham Hall	TH	B6
35	The Pond Road Residence	PON	F5
55	Vanier Residence	VR	C6
54	Winters Residence	WR	C7

PARKING GARAGES – VISITORS

80	Arboretum Lane Parking Garage	APG	D4
84	Student Services Parking Garage	SPG	E7
72	York Lanes Parking Garage	YPG	D6

PARKING LOTS – VISITORS

83	Atkinson Lot	F5
107	Calumet Lot	D3
66	Founders Road East Lot	A6
73	Lumbers Lot	C5
82	Nelson Road Lot	E4
64	Northwest Gate Lot	B3
105	Northwest Gate North Lot	A4
106	Northwest Gate South Lot	B3
89	Physical Resources South Lot	C8
77	Thompson Road East Lot	C3
79	Thompson Road Lot	C3
71	Vanier Lot	C6

PARKING LOTS – RESERVED PERMIT

67	Albany Road Lot	B7
97	Chimneystack Lot	C8
66	Founders Road East Lot	A6
65	Founders Road West Lot	A5
88	Haynes Avenue Lot	G5
87	Kinsmen Lot	B9
74	Lumbers North Lot	C5
82	Nelson Road Lot	E4
64	Northwest Gate Lot	B3
91	Passy Crescent Lot	G3
85	Physical Resources North Lot	B8
68	Physical Resources East Lot	B9
63	Shoreham Drive Lot	C2
86	The Pond Road Lot	G6
78	West Office Building Lot	C4
70	York Boulevard Lot	D7

VISUAL PERFORMANCE ARTS FACILITIES

92	Art Gallery of York University	ACE	E7
37	Burton Auditorium *	BU	E6
93	Gales Gallery	ACW	E6
38	Joseph G. Green Studio Theatre	CFT	E6
92	McLean Performance Studio	ACE	E7
92	Price Family Cinema	ACE	E7
92	Sandra Faire & Ivan Fecan Theatre	ACE	E7
92	Tribute Communities Recital Hall	ACE	E7

ATHLETICS & RECREATION FACILITIES

3	Alumni Field	STA	A2
4	Tait McKenzie Centre	TM	C3
6	York Lions Stadium	YLS	C4

HISTORICAL HOUSES

D	Hoover House	HOH	F2
C	Skennen'kó:wa Gamig ᐢᑲᓇᑲᐊᑯ ᑲᒥᐢ	SKG	F4
B	Stong Barn	SB	A9
A	Stong House	SH	A9

NON-YORK U BUILDINGS

96	Archives of Ontario *	AO	D7
1	Canlan Ice Sports *	ICE	B1
62	CMIC Building	CMB	B9
57	Harry Sherman Crowe Housing Co-op	HCC	C7
101	Pioneer Village TTC subway station	PVS	A3
40	Seneca @ York, Stephen E. Quinlan Building	SAY	E7
8	Tennis Canada - Aviva Centre *	TC	D2
7	Track & Field Centre *	TFC	B4
99	The Quad Student Housing *(managed by Forum Campus Suites)*	QSH	F6
104	York Region Transit Terminal		A3
100	York University TTC subway station	YUS	D6

* Shared use
* Temporarily Closed

Please visit yorku.ca/mapskeele for the most up to date version of the campus map

August 2021

TALKS

Keynote Speaker

J. Kevin O'Regan

Integrative Neuroscience & Cognition Center,
Université Paris Cité
www.kevin-oregan.net

TALKS

Kevin O'Regan is emeritus ex-director of the Laboratoire Psychologie de la Perception, CNRS, Université Paris Descartes. After early work on eye movements in reading, he was led to question established notions of the nature of visual perception, and to discover, with collaborators, the phenomenon of "change blindness". In 2011 he published a book with Oxford University Press: "Why red doesn't sound like a bell: Understanding the feel of consciousness". In 2013 he obtained a five year Advanced ERC grant to explore his "sensorimotor" approach to consciousness in relation to sensory substitution, pain, color, space perception, developmental psychology and robotics. Currently, within other European projects he is exploring how young infants learn the structure of their bodies.

WHAT IS IT TO PERCEIVE THE WORLD?

Our visual systems are full of defects and distortions, yet we have the impression of seeing a continuously present, richly populated, colorful, detailed visual world. To explain this, it is usually assumed that the brain uses compensation mechanisms to remove the defects and create a perfected internal representation. Seeing, according to this standard view, consists in activating this perfected internal representation of the outside world.

I will question this idea and suggest an alternative, "sensorimotor" view, according to which seeing does not consist in activating an internal representation, but in being engaged in actively exploring the outside world. The brain guides the exploration process, but does not generate the experience of seeing. The experience of seeing lies in the process of actively exploring the outside world.

I shall give examples of how this approach deals with the apparent defects of the visual system, explains change blindness and inattentional blindness and has applications to color and space perception, as well as to camera calibration. It also suggests a pathway towards understanding the "hard" problem of consciousness.

Speakers & Abstracts

Ludwig Huber

Messerli Research Institute,
University of Veterinary Medicine Vienna

Ludwig Huber is Professor of Comparative Cognition at the Messerli Research Institute, an Interdisciplinary Institute of the University of Veterinary Medicine Vienna, the Medical University of Vienna, and the University of Vienna. There he is the head of the institute and of the Division of Comparative Cognition, which includes the Clever Dog Lab Vienna and the Research Station on Cognition and Communication (Haidlhof). His research focuses on animal cognition in a broad, comparative manner, including such diverse species as humans, marmosets, dogs, kea, pigeons, tortoise, lizards, poison frogs and archer fish. He has written ca. 160 research articles and 40 book chapters about these topics, and is the co-editor of several books including The Evolution of Cognition (MIT Press 2000).

EYE-TRACKING WITH DOGS: ACHIEVEMENTS AND CHALLENGES

In this paper I will review eye-tracking studies with dogs (Canis familiaris) with a twofold goal; I will highlight a few studies that exemplify the achievements in the field of canine perception and cognition using eye-tracking methods, but also discuss the challenges that arise in the application of a technology that has been developed in human psychophysics. For the first goal I will present studies that investigated dogs' perception of humans, mainly faces, but also hands, gaze, emotions, communicative signals, goal-directed movements and social interactions, and also the perception of animations representing possible and impossible physical processes and animacy cues. In the second part of my talk I will discuss the present challenges of eye-tracking with dogs, like doubtful picture-object equivalence, extensive training, small sample sizes, difficult calibration and artificial dog behavior. Finally, I will suggest possible improvements and solutions for the mentioned problems, among them better stimulus quality, better data quality, mobile and calibration-free eye-tracking, and combinations with behavioral and neuroimaging methods.

Jody Culham

Western University

Dr. Jody Culham is a Professor and Canada Research Chair in Immersive Neuroscience in the Department of Psychology at Western University (aka University of Western Ontario). Her research investigates how vision is used for perception and to guide actions in human adults. She uses a combination of cognitive neuroscience techniques -- including functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), as well as behavioral psychophysics and kinematics. One theme of her work is bringing cognitive neuroscience closer to everyday life by studying natural behavior and brain processing in real-world situations and compelling virtual-reality simulations. Dr. Culham received a Bachelor's degree from the University of Calgary and a PhD from Harvard University before joining Western as a postdoctoral fellow and then faculty member.

TALKS

CLOSING THE LOOP WITH IMMERSIVE NEUROSCIENCE: ON THE IMPORTANCE OF REAL ACTIONS AND FEEDBACK FOR BRAIN AND BEHAVIOR

Evolution and development have shaped brains that survive and flourish by optimizing actions over the short and long term based on the consequences of those actions. However, common proxies for studying brain and behavior in the laboratory often lack the potential for genuine physical interactions with meaningful consequences. I will review evidence from cognitive neuroscience suggesting that the potential for genuine actions modifies behavior and brain responses. I will also present data from recent neuroimaging studies of video games to show that active control of game outcomes affects activation levels and inter-regional correlations. I propose a new approach called Immersive Neuroscience in which we study active "agents" rather than passive "observers" or "subjects" to better understand the full gamut of cognitive functions using realistic scenarios with closed-loop feedback.

Jacqueline Snow

University of Nevada, Reno

Jacqueline Snow is an Associate Professor of Cognitive Neuroscience in the Department of Psychology at the University of Nevada, Reno. Dr. Snow completed her graduate training in clinical neuropsychology and cognitive neuroscience at the University of Melbourne, Australia, followed by postdoctoral fellowships in the United Kingdom and Canada. Dr. Snow's research examines how humans perceive and behave towards objects in naturalistic real-world environments, with an emphasis on studying whether real-world objects are processed and represented in the brain differently to 2-D images, 3-D stereoscopic images, and 'graspable' 3-D objects presented using augmented reality displays. Dr. Snow's research leverages multipronged methods, including behavioral psychophysics, neuropsychology, fMRI, EEG, eye-tracking and motion kinematics. Dr. Snow's research has wide-ranging translational applications, which include identifying effective target sites for the placement of brain-computer interfaces, improving our understanding of visual perceptual disorders, bolstering learning and memory, and understanding the causes of obesity.

COMPARING BRAIN RESPONSES TO REAL OBJECTS VERSUS 2-D PICTURES: EMERGING FMRI DATA FROM THE ADULT BRAIN AND ITS IMPLICATIONS FOR UNDERSTANDING THE DEVELOPING CORTEX

Artificial stimuli, such as two-dimensional (2-D) images presented on a computer screen, are thought to be adequate proxies for real-world solid objects to understand the cognitive and neural basis of object recognition. However, there is mounting evidence that real objects elicit responses that differ from those elicited by pictures – both in the adult and the developing brain. In this talk, I will review emerging fMRI data from my lab, collected in collaboration with Dr. Erez Freud and Dr. Marlene Behrmann, that reveals surprising differences in the way real objects and pictures are represented in healthy adult observers. Next, I will relate these fMRI findings to recent behavioral data that my students and I have collected in young children, and I will highlight future directions that we are taking to understand the developmental trajectory of object coding in human cortex.

Heiko Hecht

Johannes Gutenberg-Universität Mainz

Heiko Hecht studied psychology and philosophy at the universities of Trier, Germany, and Virginia, USA. He has held positions at the University of Munich, NASA Ames Research Center, Universität Bielefeld, and the Massachusetts Institute of Technology. He currently holds the chair of Experimental Psychology at the Johannes Gutenberg-Universität Mainz, Germany. He has published over 150 journal articles and edited several books in the domains of picture perception and virtual reality, artificial gravity, time-to-contact estimation, and intuitive physics. He is currently a fellow at the Wissenschaftskolleg zu Berlin.

THE PRIVATIZATION OF PICTORIAL SPACE

Traditionally, the observer of a painting or a photograph has been in a stable dual relation with the picture. On the one hand, the observer is in a well-defined relation with the physical surface. Distance and size of the canvas are perceived accurately and exist in the observer's physical space. On the other hand, the observer is in an ill-defined relation with the content of the picture. The objects she sees-in the picture are disembodied in the sense that their orientation and position remain strangely stable relative to the observer but not relative to her physical space. I will argue that this disembodiment gives rise to many aesthetic and divine properties of pictorial space. With the recent spread of virtual reality technologies, this quality of disembodiment is challenged. The observer acquires a new level of situatedness with respect to pictorial space. At the same time, being well-defined is no longer a privilege of the physical world. Well-defined virtual objects conquer pictorial space. In this process, pictorial space loses its ubiquity and is privatized. This change poses a number of questions. Will the exposure to these novel pictorial spaces alter our relationship with real space? Do novel virtual spaces call for a novel theory of seeing-in? How are artists reacting to the re-embodiment of pictorial space?

Gabriel Greenberg

University of California, Los Angeles

Gabriel Greenberg is an Associate Professor in the Department of Philosophy at UCLA. His research interests include iconicity, pictures, maps, and diagrams, and their role in a computational theory of mind.

PICTURE PERCEPTION AND THE RADICAL VARIABILITY OF THE VISUAL SYSTEM

It is widely thought that picture interpretation is carried out by the visual system using computational machinery that evolved for perception. A persistent problem for this idea is the variety of pictorial systems which appear to deviate from normal perception: black and white photography, line drawing, stick figures, and axonometric projection, to name only a few. In this talk, I develop a conjecture about how the visual system accommodates this variety: sub-modules within the visual system are selectively turned off or parametrically modified during picture perception. For example, color detection might be turned off while contour detection remained intact; I discuss this and a range of other candidate cases. The result is a view of the visual system as a powerful computational engine which, unlike the relatively fixed operation of perception proper, is capable of radical variability.

John Kulvicki

Dartmouth College

John Kulvicki is Professor of Philosophy at Dartmouth College, and he works primarily in the philosophy of perception and the philosophy of art, with a focus on non-linguistic representation.

NON-SPATIAL PROJECTIONS

We typically think of projections in primarily spatial terms. This talk shows that the same tools we have for understanding spatial projections can help us understand the perception of color. There are chromatic projections, too, and I will unpack what I mean by this and show why it is important.

James Elder

York University

James Elder is Professor and York Research Chair in Human and Computer Vision at York University. He is jointly appointed to the Department of Psychology and the Department of Electrical Engineering & Computer Science at York and is a member of York's Centre for Vision Research (CVR) and Vision: Science to Applications (VISTA) program. His research employs psychophysical and computational methods to understand the principles that underly both biological and machine vision systems. Dr. Elder's current research is focused on natural scene statistics, perceptual organization, contour processing, shape perception, single-view 3D reconstruction, attentive vision systems and machine vision systems for dynamic 3D urban awareness.

Dr. Elder has led numerous large-scale multi-sector collaborative research projects and currently directs the multi-institution ORF-RE project Intelligent Systems for Sustainable Urban Mobility. He holds three patents on attentive vision technologies and is the co-founder of the AI start-up AttentiveVision.

SINGLE-VIEW 3D PERCEPTION IN HUMANS AND MACHINES

While multi-view cues such as stereopsis and motion parallax are important to 3D perception, humans still perceive objects and scenes in 3D when viewed in a single image or at distances where stereopsis and motion cues are weak. How do our brains do this? While deep networks can be trained to recover 3D range from a single image, these systems fail to capture the crisp structures that humans perceive, and with millions of free parameters provide little insight into the geometry and computational principles that underpin monocular 3D perception.

For a deeper insight into these principles, recall the transformation in picture creation that emerged in the early Renaissance with the discovery of the geometric rules of perspective projection. These rules connect simple 2D cues in the picture plane (e.g., vanishing points), to 3D properties of surfaces (e.g., 3D orientation) based on relatively general regularities in the 3D scene (e.g., parallelism). Unlike stereoscopic disparity and motion parallax, these cues do not weaken with distance, depending only on visual angle, and thus are particularly important in the far field.

In this talk I will explore the use of these perspective projection cues by both humans and machines for single-view 3D perception. In lieu of high-dimensional black-box deep networks, I will focus on more explainable low-dimensional models that can be explicitly related to the underlying geometry and statistics of our visual world and can be applied to a range of problems in 3D urban scene understanding.

Frank Steinicke

Universität Hamburg

Frank Steinicke is professor for Human-Computer Interaction at the Department of Informatics at the Universität Hamburg. His research is driven by understanding the human perceptual, cognitive and motor abilities and limitations in order to reform the interaction as well as the experience in computer-mediated realities. He studied Mathematics with a minor in Computer Science at the University of Münster, from which he received his Ph.D. in 2006, and the Venia Legendi in 2010, both in Computer Science. He published about 300 peer-reviewed scientific publications and served as program chair for several XR and HCI-related conferences. Furthermore, he is chair of the steering committee of the ACM SUI Symposium, and member of the steering committee of GI SIG VR/AR. Furthermore, he is a member of the editorial boards of EEE Transactions on Visualization and Computer Graphics (TVCG) as well as Frontiers Section on Virtual Reality and Human Behaviour.

TALKS

B(L)ENDING REALITIES

The fusion of mixed reality (MR) and artificial intelligence (AI) will revolutionize human-computer interaction. MR/AI technologies and methods will enable scenarios with seamless transitions, interactions and transformations between real and virtual objects along the reality-virtuality continuum indistinguishable from corresponding real-world interactions. Yet, todays immersive technology is still decades away from the ultimate display. However, imperfections of the human perceptual, cognitive and motor system can be exploited to bend reality in such a way that compelling immersive experiences can be achieved. In this talk, we will review some mixed reality illusions, which bring us closer to the ultimate blended reality.

Austin Roorda

University of California, Berkeley

Austin Roorda received his Ph.D. from the University of Waterloo in 1996 in Vision Science & Physics. Since then, he's pioneered applications of AO and ophthalmoscopy, including mapping of the human trichromatic cone mosaic, designing the first adaptive optics scanning laser ophthalmoscope, tracking and targeting light delivery with cone-level accuracy in the human eye, and being part of the first team to use AO imaging to monitor efficacy of a treatment to slow retinal degeneration. Since 2005, he has been at Berkeley where he is a member of the Vision Science, Bioengineering and Neuroscience graduate programs. Notable awards are the Distinguished Alumni Award (University of Waterloo School of Optometry, 2007), the Glenn Fry Award (American Academy of Optometry, 2009), a John S. Guggenheim Fellowship (2014), an Alcon Research Institute Award (2016) and a Leverhulme Visiting Professorship (2020) from the University of Oxford.

THE SHARPEST, BRIGHTEST, HIGHEST-CONTRAST, MOST-IMPRACTICAL, GAZE-CONTINGENT DISPLAY EVER BUILT

The adaptive optics scanning laser ophthalmoscope (AOSLO) was originally designed to record movies of living human retina with cellular-level resolution. But the same raster scanning platform, when equipped with high-speed light modulators and high-speed image processers, can be used as a unique type of display. The display is ultra-sharp: Aberration correction can be used over a large pupil to display images sharper than anything humans have ever experienced. The display is ultra-bright: Direct projection of laser light sources can achieve equivalent luminances of over 5 x 105 cd/m2, only limited by retinal exposure safety limits. The display has ultra-high-contrast: The use of acousto-optic modulators can achieve contrast ratios of over 1010:1 for each 'primary', enabling measurements down to rod-threshold levels. The display is gaze-contingent: Real time eye tracking enabled through processing of the scanned retinal image allows for low-latency, targeted projection of the display with cone-level accuracy. Alas, the display is ultra-impractical: It fills an entire room, it costs upwards of $300K, it operates with four computers, the display is typically 1 to 2 degrees in size, and the subject needs to use a bite bar to maintain alignment with the system. Nevertheless, our display enables unprecedented control for vision experiments, from revealing the beneficial role of fixational eye motion for human visual acuity to testing the expansion of human color experience via direct control of cone excitations.

Martin Banks

University of California, Berkeley

Martin Banks received Bachelor's in Psychology at Occidental College (1970). He entered graduate school at UC San Diego receiving Master's (1973). Moved to University of Minnesota receiving Ph.D. (1976). Assistant and Associate Professor of Psychology at University of Texas (1976-1985). Associate and Full Professor of Optometry at UC Berkeley (1985-present). Chairman of Vision Science (1995-2002; 2012).

Awards include Young Investigator Award from National Research Council, McCandless Award from American Psychological Association, Koffka Medal from Giessen University, Prentice Award from American Academy of Optometry, Schade Prize from Society for Information Display, and Tillyer Award from Optical Society of America. He has been Fellow of the Center for Advanced Study of the Behavioral Sciences, Fellow of American Association for the Advancement of Science, Fellow of American Psychological Society, Holgate Fellow of Durham University, WICN Fellow of University of Wales, Borish Scholar of Indiana University. Elected to National Academy of Sciences.

TALKS

BINOCULAR VISION AND OCULOMOTOR BEHAVIOR IN NATURAL AND VR-GAMING ENVIRONMENTS

The human visual system evolved in an environment with statistical regularities. Binocular vision is adapted to these such that depth perception and eye movements are precise, fast, and performed comfortably in environments consistent with the regularities. We measured the statistics of binocular disparities and eye movements in natural and VR-gaming environments. In the natural environment, there is a clear pattern of binocular disparity across the visual field. And clear biases in fixation directions and distances. The disparity pattern is disrupted in the VR-gaming environment. Furthermore, fixation directions and distances are more restricted, and fixation distance is generally farther. We conducted a VR user study to investigate how deviation from the natural pattern of disparity affects comfort and performance. We found that content that is more consistent with the statistics of the natural world yields less discomfort than content that is not. Furthermore, consistent content yields slightly, but consistently, better performance than inconsistent content. Our results can inform the design of HMDs and VR games.

Hong Hua

University of Arizona

Hong Hua, Fellow of SPIE, OSA and NAI, is currently a Professor with the James C. Wyant College of Optical Sciences (OSC), The University of Arizona. She has over 25 years of experiences in designing and developing near-to-the-eye display technologies and developing virtual reality and augmented reality applications. She has published 250+ technical papers and presentations in these fields, including 100+ peer-reviewed journal and proceeding papers and numerous plenary speeches and invited colloquial addresses at academic institutions and major scientific conferences world-wide. Hua holds more than 50 issued U.S. and foreign patents that have been licensed to 7 companies and universities. Her current research interests include various head-worn displays and 3-D displays, optical engineering, collaborative virtual and augmented environments, and human-computer interaction.

HEAD-MOUNTED LIGHT FIELD DISPLAYS FOR VIRTUAL AND MIXED REALITY

A light field display aims to render the perception of a 3D scene by reproducing the geometric light rays apparently emitted by the 3D scene in different directions. This talk will review the recent development of head-mounted light field displays (LF-HMD).

Mel Slater

University of Barcelona

Mel Slater is a Distinguished Investigator at the University of Barcelona in the Institute of Neurosciences, and co-Director of the Event Lab (Experimental Virtual Environments for Neuroscience and Technology). He was previously Professor of Virtual Environments at University College London in the Department of Computer Science. He has been involved in research in virtual reality since the early 1990s, and has been first supervisor of 40 PhDs in graphics and virtual reality since 1989. He held a European Research Council Advanced Grant TRAVERSE 2009-2015 and has now a second Advanced Grant MoTIVE 2018-2023. He is a Research Award Winner of the Alexander von Humboldt Foundation in 2021. He is Field Editor of Frontiers in Virtual Reality, and Chief Editor of the Human Behaviour in Virtual Reality section. His publications can be seen on http://publicationslist.org/melslater.

ILLUSIONS AND APPLICATIONS OF VIRTUAL REALITY

Virtual Reality is paradoxical in the sense that it gives rise to perceptual illusions, that everyone knows for sure are illusions, but nevertheless people tend to behave realistically in response to virtual situations and events. In this talk I will describe the major illusions of virtual reality: the illusion of being in the virtual place, that events there are really happening, that the body has taken a different form, and that the space is shared with others. I will give examples of the practical uses of these illusions for social and personal enhancement.

Gabriel Diaz

Rochester Institute of Technology

Gabriel Diaz has a B.S. in Psychology from Skidmore College, and a Ph.D. in Cognitive Science from Rensselaer Polytechnic University, where he worked with Brett R. Fajen. From 2010-2013, he worked as a post-doc with Mary Hayhoe at UT Austin, and in August of 2013, Gabriel joined the faculty at the Rochester Institute of Technology's Center for Imaging Science, where he is now an Associate Professor and Director of the PerForM Lab (Perception for Movement). PerForM draws on Dr. Diaz's prior experience in the use of virtual reality, motion capture, eye tracking, and machine learning to study the role of eye and head movements in guiding behavior in natural and simulated environments. Basic research is conducted in parallel with industry-funded and applied work on the development of new technologies to improve the accuracy and precision of mobile and VR/AR-integrated eye tracking technology.

THE PRESENT AND NEAR FUTURE OF MOBILE AND XR-INTEGRATED EYE TRACKING

We are now amid a period of breakneck development in the area of mobile eye tracking. Although this development has largely been motivated by heavy investment from the XR industry, the ability for laboratory-quality eye tracking away from the desk will have a revolutionary impact on the ability for vision researchers to investigate the role of gaze behavior in natural contexts. The PerForM Lab at RIT has been motivated by this potential impact to work with industry in the development of new methods to train neural networks for the segmentation of features present in near-infrared eye imagery - a critical step in many mobile eye tracking pipelines that has great influence on the accuracy, precision, and robustness of the final estimate of gaze direction. These efforts draw heavily upon a our related work on RIT-Eyes, an automated and parametric Blender pipeline for the generation of computer-generated eye images with the pixel-level ground-truth labels of eye part (e.g. of the pupil, iris, and sclera) necessary for training segmentation networks. In combination, these tools allow us to test hypotheses for how best to train networks when using a combination of synthetic and real-world imagery, and how to leverage temporal dynamics of eye movements in the process of eye-image segmentation. Although the work is still in progress, it will eventually share our efforts with the broader vision research community through open-source licensing.

Jennifer Campos

KITE - Toronto Rehabilitation Institute - University Health Network

Dr. Jennifer Campos is a Canada Research Chair (II) in Multisensory Integration and Aging. She is a Senior Scientist, Chief Scientist of the Challenging Environments Assessment Laboratory, and the Associate Director–Academics at KITE, Toronto Rehab–UHN. Jenny is also the Associate Scientific Director of AGE-WELL and an Associate Professor of Psychology (University of Toronto). Jenny's research focuses on multisensory self-motion perception and mobility (e.g., walking, driving) under realistic and challenging conditions. This includes understanding how age-related sensory loss (e.g., vision, hearing) and cognitive declines can increase the risk of falls and vehicle collisions. She uses VR and simulation technologies to a) systematically manipulate aspects of multisensory experiences (sights, sounds, motions) to examine how sensory inputs are integrated in the brain; b) optimize simulation technologies for use as novel tools for research and application and c) help bridge the gap between highly-controlled lab studies and real world impact.

SEEING, HEARING, MOVING: AGE-RELATED CHANGES TO MULTISENSORY INTEGRATION PROCESSES SUPPORTING MOBILITY

In research and in clinical applications, sensory abilities including hearing, vision, and vestibular (inner ear balance) functions are most commonly studied and assessed independently of each other. However, during most everyday behaviours, information from sights, sounds, and movements occur simultaneously and jointly inform our perceptions and actions. For example, as we navigate through our environment, our brain must effectively and dynamically integrate information from visual inputs, auditory inputs, our muscles and joints (proprioception), and the acceleration detectors in our inner ear (vestibular). Combining and optimally integrating across multiple sensory inputs allows us to estimate important movement parameters such as speed and heading direction with greater certainty.

Emerging evidence also indicates that the processes of multisensory integration change with older age. For example, multisensory integration may be heightened in older adults compared to younger adults and sensory weightings may be non-optimal in older adults for some conditions. Little is known about how these age-related effects are observed for multisensory self-motion perception specifically or for complex behavioural tasks generally. Problems integrating across sensory inputs during self-motion may lead to problems such as poor mobility, driving collisions, or increased falls risk.

In this talk I will describe studies that use state-of-the-art virtual reality simulators and motion platforms to investigate the integration of visual, auditory, and vestibular inputs on tasks ranging from psychophysical heading discrimination tasks to simple driving tasks in healthy younger and older adults. Unique effects of sensory declines, such as age-related hearing loss on self-motion perception will also be highlighted.

Craig Chapman

University of Alberta

Dr. Chapman has a BSc from the University of British Columbia (UBC) in Cognitive Science and an MSc and PhD in Psychology from the University of Western Ontario where he worked with Dr. Melvyn Goodale. After finishing a Killam postdoctoral fellowship with Drs. Jim Enns, Alan Kingstone and Todd Handy back at UBC he joined the Faculty of Kinesiology, Sport, and Recreation and the Neuroscience and Mental Health Institute at the University of Alberta (U of A). He has been running the Action in Complex Environments (ACELab) Laboratory at the U of A since 2013. Dr. Chapman is a Canadian Institute for Advanced Research (CIFAR) Global Scholar in Brain, Mind and Consciousness and most recently the CEO of a spin-off company called Gaze and Movement Analysis Inc (GaMA for short). At the heart of both his research and entrepreneurial aspirations is the fundamental belief that we can learn much about what is going on inside someone's head by doing careful measurement of what's going on outside their head, namely how they move their body and where they look. With raw gaze and movement data becoming increasingly available via advances in technology, now more than ever we need better tools that can unlock the powerful insights this data holds.

GAZE AND MOVEMENT ANALYSIS (GAMA) IN REAL AND VIRTUAL WORLDS

In the first half of this talk, I will show how much we can learn about cognition by studying eye and motion tracking data. I call this the "Moving is Thinking" hypothesis and I will highlight studies of reach tracking during decision making and reach / eye tracking during object interaction tasks. These examples will include a comparison of eye-hand coordination between the same task conducted in the real world and replicated in Virtual Reality (VR). Given the power of these analysis tools, it begs the question - why don't more people measure these aspects of behaviour, especially in augmented and virtual reality (AR/VR) where the data is natively available? I think part of the answer is that this data is hard to collect and harder to analyze. Therefore, I'll spend the second half of my talk outlining some new tools we've developed that make human movement analysis easier. Specifically, I'll introduce our Gaze and Movement Analysis (GaMA) software and describe some exciting developments about how we use these tools to provide automated metrics of human performance in AR/VR and in concert with other data streams like EEG.

Elisa Rafaella Ferrè

Birkbeck, University of London

Dr Elisa R. Ferrè completed a BSc and a MSc in Psychology at the Università degli Studi di Pavia (Italy). She obtained a PhD in Psychology at the Università degli Studi di Pavia in 2012 for the investigation of multisensory integration between vestibular, somatosensory and visual function in humans. She was a Postdoctoral Fellow at the Institute of Cognitive Neuroscience at University College London (UK) from 2012 to 2015. She took up her first faculty position at Royal Holloway University of London (UK) in 2015. Since September 2021, Elisa is Senior Lecturer and Director of the Vestibular Neuroscience laboratory at Birkbeck University of London (UK). Her research combines experimental psychology, cognitive neuroscience, neuroimaging, vestibular physiology and space science methods to understand how vestibular information shapes human behaviour.

VESTIBULAR SYSTEM: FROM SIGNAL ENCODING TO PERCEPTION

The vestibular system plays an essential role in everyday life, contributing to a surprising range of functions from reflexes to the highest levels of perception and cognition. Three orthogonal semicircular canals detect rotational movements of the head in space and the otolith organs sense translational acceleration, including the gravitational vertical. But, how vestibular signals are encoded by the human brain? We have recently combined innovative methods for eliciting virtual rotation and translation sensations with fMRI to identify brain areas representing pure vestibular rotation and translation movements. We have identified a bilateral inferior parietal, ventral premotor/anterior insula and prefrontal network and confirmed that these areas reliably possess information about vestibular rotation and translation movements. We have also investigated how vestibular signals are integrated with other sensory information to generate our perception of the external environment.

Markus Lappe

University of Münster

Markus Lappe received his original training in physics and graduated from the University of Tuebingen, Germany, in 1992. After research positions at the Max-Plank-Institute for Biological Cybernetics in Tübingen, the National Institute of Mental Health, USA, and the Ruhr University Bochum he became Professor for Experimental Psychology and Cognitive Neuroscience at the University of Münster in 2002. His lab forms an interdisciplinary research community focussing on computational modelling and psychophysical and electrophysiological experiments on motion perception, eye movements, spatial vision and virtual reality. Markus Lappe was awarded the Biofuture Prize by the German Ministry for Research and Education in 1999. He served as Director of the Institute for Psychology II from 2002 to 2010 and as Dean of the Department from 2010 to 2012. He is a founding member of the Otto Creutzfeldt Institute for Cognitive and Behavioral Neuroscience in Münster.

WALKING WITH OTHERS

The movement of oneself and the movements of others are the two main generators of retinal motion. The first, self-motion, gives rise to optic flow, a motion pattern that that is important for visual navigation and motion control. The second, the motion of other living beings, is known as biological motion and conveys information about their actions and movements. Past research has shown fundamental differences between the computational requirements and perceptual mechanisms of these two types of retinal motion, suggesting that they are analysed in different cortical pathways. Yet, ecologically relevant situations containing both types of motion, as when one walks amongst others, have seldom been investigated. Computational considerations suggest, on the one hand, that the two types of motion will inflict problems onto each other since they violate respective computational prerequisites. On the other hand, since both are supported by distinct perceptual mechanisms there may be synergies and cross talk between the systems. I will present experiments on self-motion perception from stimuli combining optic flow and biological motion. They investigate how self-motion is perceived when a point-light walker is encountered in a scene, when self-motion is towards a point light walker in isolation, or when moving amongst a crowd of walkers.

Paul MacNeilage

University of Nevada, Reno

Dr. MacNeilage completed a PhD in Vision Science at UC Berkeley and post-doctoral training in vestibular physiology and perception at Washington University School of Medicine. Following a research group leader position at the German Center for Vertigo and Balance in Munich, Germany, he took up a professor position in the Psychology Department at the University of Nevada, Reno. His research is focused on psychophysical measurement and probabilistic modeling of human spatial orientation perception. This is supplemented by work characterizing the natural statistics of human head, eye, and visual motion to understand how these shape spatial orientation processes.

TALKS

SENSORY AND MOTOR SIGNALS MEDIATING STATIONARITY PERCEPTION

Stationarity perception refers to the ability to accurately perceive the surrounding visual environment as world-fixed during self-motion. Perception of stationarity depends on mechanisms that evaluate the congruence between retinal/oculomotor signals and head movement signals. In a series of psychophysical experiments, we systematically varied the congruence between retinal/oculomotor and head movement signals to find the range of visual gains that is compatible with perception of a stationary environment. On each trial, human subjects wearing an HMD experience a yaw head movement and report whether the visual gain was perceived to be too slow or fast. A psychometric fit to the data across trials reveals the visual gain most compatible with stationarity (PSE) and the sensitivity to visual gain manipulation (JND). Across experiments, we varied 1) the spatial frequency and retinal stimulus location (central versus peripheral) of the visual stimulus, 2) the scene-fixed versus head-fixed nature of the oculomotor signal, and 3) the active versus passive nature of the head movement. Stationarity perception is most precise (lowest JND) and accurate (PSE~1) during active head movement, with scene-fixed fixation, and central retinal stimulus location. Passive head movement, head-fixed fixation, and peripheral stimulus location all lead to a reduced gain perceived as stationary (PSE<1, meaning slower scene motion is more compatible with stationarity) and impaired sensitivity (increased JND, presumably due to increased noise on the contributing signals). Measures of VR sickness appear to covary with PSEs and JNDs. We interpret these results in the context of known differences in physiological processing across these scenarios.

Dora Angelaki

New York University

Dora Angelaki is a Professor at the Center for Neural Science and the Tandon school of Engineering at New York University. She holds a diploma and Ph.D. degrees in electrical and biomedical engineering from the National Technical University of Athens, Greece, and the University of Minnesota. She is interested in understanding the principles that make our brain so much better than man-made machines and AI. She uses naturalistic foraging tasks that combine uncertainty, spatial navigation, decision-making and episodic memory to understand inference in the brain. She explores how task-relevant latent variables and multisensory signals flow dynamically across brain areas to generate perception and cognition, how hierarchical causal inference is implemented in the brain, how beliefs propagate through the network, and how internal states modulate this information flow.

EYE MOVEMENTS FOR ACTIVE SENSING AND MEMORY DURING VISUALLY-GUIDED NAVIGATION

We will summarize recent findings on the role of active sensing (gaze) in planning and memory. By analyzing the spatial distribution of human gaze to transiently visible goals in simple and complex virtual mazes we found that environmental complexity mediated a striking trade-off in the extent to which attention was directed towards two complimentary aspects of the world model: the reward location and task-relevant transitions. The temporal evolution of gaze revealed rapid, sequential prospection of the future path, evocative of hippocampal neural replay. These findings suggest that the spatiotemporal characteristics of gaze during navigation are significantly shaped by the unique cognitive computations underlying real-world, sequential decision making.

Michael Barnett-Cowan

University of Waterloo

Dr. Michael Barnett-Cowan is an Associate Professor of Neuroscience in the Department of Kinesiology at the University of Waterloo where he is the Director of the Multisensory Brain & Cognition laboratory. Michael received his PhD in Experimental Psychology in 2009 at York University with Laurence Harris at the Centre for Vision Research. He then took up a postdoctoral fellowship at the Max Planck Institute for Biological Cybernetics in Tübingen, Germany with Heinrich Bülthoff where he led the Cybernetics Approach to Perception and Action (CAPA) research group and was project leader for the Simulation of Upset Recovery in Aviation (SUPRA) F7 EU Research Grant. In 2012 he returned to Canada to work with Jody Culham at Western University's Brain and Mind Institute where he held appointments as an adjunct research professor and a Banting fellow. Michael's research program uses psychophysical, computational modelling, genomic as well as neural imaging and stimulation techniques to assess how the normal, damaged, diseased, and older human brain integrates multisensory information that is ultimately used to guide perception, cognition and action. More information about Michael's research can be found here:

https://uwaterloo.ca/multisensory-brain-and-cognition-lab/
https://scholar.google.com/citations?user=WqcCbxoAAAAJ&hl=en

SEAS THE DAY: DEVELOPING A VR EXERGAME DURING COVID-19

Physical activity is associated with physical and cognitive benefits among people living with dementia or mild cognitive impairment (PLWD/MCI) and is a meaningful activity that can improve their confidence in everyday life. Virtual reality (VR) has potential to help older adults remain physically and mentally active, but content is lacking and while co-design can significantly improve the design of technology, it is rarely done with PLWD/MCI. This study used participatory design methods and collaborative approaches involving key stakeholders to develop and test an immersive multisensory VR Exergame "Seas the Day", a novel solution targeting PLWD/MCI well-being that is freely available (Oculus Quest 1, 2). The multidisciplinary and collaborative design process occurred over 15 months (overlapping with COVID-19 pandemic). We involved persons living with dementia/MCI, exercise professionals, community-dwelling older adults, a VR company for content creation, and a multidisciplinary research team with game designers, engineers, and kinesiology experts. The game was designed to target movements identified by exercise professionals and researchers (aerobic exercises, range of motion, seated-balance, quick response to stimuli) and is structured in three exercise stages (warm-up, conditioning, cool-down). To ensure safety of participants while using VR headsets, only seated upper-limb exercises were targeted. Stakeholder feedback regarding game mechanics, aesthetics, and visual/auditory cues were gathered during brainstorming and playtesting sessions and implemented into specific game-related scenarios (tai-chi, rowing, fishing). Results from the study and a model for the triadic interaction (health care institution, industry partner, academia) will be presented to illustrate how different stakeholders can contribute to the design of VR exergames that consider/complement complex needs, preferences, and motivators of an underrepresented group of end users. The insights and lessons learned from this research can be used by others to co-design games, including remote engagement techniques that were used during the COVID-19 pandemic.

Marc Ernst

Ulm University

Marc Ernst is professor for Applied Cognitive Psychology at Ulm University, Germany. He studied physics in Heidelberg and Frankfurt/Main. In 2000, he received his PhD for work conducted at the Max Planck Institute for Biological Cybernetics in Tübingen, Germany. Between 2000/2001, he worked with Marty Banks at the University of Berkeley at California, before returning to the Max Planck Institute. In 2007 he became independent research group leader for Multisensory Perception and Action at the Max Planck Institute. In 2011 he became full-professor in Biology at the University of Bielefeld, Germany heading the cognitive neuroscience group. In Bielefeld he was also scientific director of the Centre for Interdisciplinary Research (ZiF) and co-coordinator of Centre of Cognitive Interaction Technology (CITEC). 2016 he then moved to Ulm. His research interests include multisensory perception and action, sensorimotor control and decision making, perceptual development and learning, and Virtual/Mixed Reality and HCI.

DEVELOPMENT OF MULTISENSORY PERCEPTION AND DISPLAYS

We recently had the chance to investigate several Ethiopian children and adolescent who were born with dense bilateral cataracts and were thus legally blind or with severely low-vision during the first years of their lives. In collaboration with an Israeli doctoral team and the Hawassa Referral Hospital these patients were surgically treated for their cataracts and we were able to follow their development after surgery sometimes for three or more years. In this talk I will report on several studies focusing on the development of multisensory perception and action. Many of the multisensory abilities, such as optimal multisensory integration or visuomotor recalibration, seem to develop within the first months to years after sight recovery. However, other tasks such as for example the visually guided grasping behavior did not significantly improve over a period of more than three years. In the final part of my talk, I will report on some of our recent developments towards interactive multisensory VR displays which enable us to study the development, learning, and plasticity of multisensory perception and action in the healthy population as well as some groups of people with particular deficits, such as the blind or deaf.

Charles Spence

University of Oxford

Professor Charles Spence is a world-famous experimental psychologist with a specialization in neuroscience-inspired multisensory design. He has worked with many of the world's largest companies across the globe since establishing the Crossmodal Research Laboratory (CRL) at the Department of Experimental Psychology, Oxford University in 1997. Prof. Spence has published over 1,000 academic articles and edited or authored, 15 books including, in 2014, the Prose prize-winning "The perfect meal", and the international bestseller "Gastrophysics: The new science of eating" (2017; Penguin Viking) – winner of the 2019 Le Grand Prix de la Culture Gastronomique from Académie Internationale de la Gastronomie. His latest book Sensehacking was published in 2021. Much of Prof. Spence's work focuses on the design of enhanced multisensory food and drink experiences, through collaborations with chefs, baristas, mixologists, chocolatiers, perfumiers, and the food and beverage, and flavour and fragrance industries. See: https://vimeo.com/170509976; & http://www.newyorker.com/magazine/2015/11/02/accounting-for-taste.

COLOURED HEARING, COLOUR MUSIC, COLOUR ORGANS, AND THE SEARCH FOR PERCEPTUALLY MEANINGFUL CORRESPONDENCES BETWEEN COLOUR AND SOUND

There has long been interest in the nature of the relationship(s) between hue and pitch or, in other words, between colour and musical/pure tones. In this talk, I want to take a closer look at the motivations that have lain behind the various assertions that have been made in the literature concerning the analogies, and possible perceptual similarities, between colour and sound. During the last century, experimental psychologists have investigated the nature of the correspondence between these two primary dimensions of perceptual experience. The multitude of different crossmodal mappings that have been put forward over the centuries are summarized, and a distinction drawn between physical/structural and psychological correspondences. The latter being further sub-divided into perceptual and affective categories. Interest in physical correspondences has typically been motivated by the structural similarities (analogous mappings) between the organization of perceptible dimensions of auditory and visual experience. Emphasis has been placed both on the similarity in terms of the number of basic categories into which pitch and colour can be arranged and also on the fact that both can be conceptualized as circular dimensions. A distinction is drawn between a dimensional alignment of pitch and hue (based on structural mapping), and the existence of specific correspondences between particular pairs of auditory and visual stimuli (often based on the idiosyncratic correspondences that have been reported by synaesthetes). Ultimately, though, the emotional-mediation account would currently appear to provide the most parsimonious account for whatever affinity the majority of people experience between musical sounds and colour.

Michael Jenkin

York University

Currently at Visiting Professor at Samsung AI Center in Montreal, I am a Professor of Electrical Engineering and Computer Science at York University and hold an Honorary Professorship at Shizuoka University in Japan. My research interests include virtual reality, human performance, computer vision and visually-guided autonomous systems.

GROWING OLD AND THE EFFECTS OF MULTI-CUE INTEGRATION

For the past 30 years or so, Laurence and I have been engaged in collaborative research exploring multi-cue integration. Leveraging a number of gifted collaborators, virtual reality technology, and a range of different props, we have explored how humans integrate perceptual cues to build a coherent model of self-orientation and self-motion. This talk reviews some of the highlights of this collaboration with particular emphasis on the effect of multi-cue integration on the visual appearance of the collaborators. The talk will also highlight some of the research results and following the style of the late Béla Julesz and others, explain how we did it better with cardboard and string 30 years ago.

Former CVR Director:

Laurence Harris

York University

PhD Cambridge University, 1979; Addison Wheeler Fellow, University of Durham, UK 1981-1983; lecturer in Physiology at Cardiff University, Wales 1983-1990; Professor in Psychology at York University, Toronto 1990-present. Associate Director of the Centre for Vision Research 1999-2003; Director from 2010-2020. Chair of Psychology Department 2006-2009. Fellow of the Canadian Psychological Association. Director of the Multisensory Integration Laboratory at York University. York Research Chair in Multisensory Integration 2017-2022. Founding editor of the journal Multisensory Research. Over 150 peer reviewed papers including articles in Nature, Science, Vision Research, Journal of Physiology, Experimental Brain Research, Scientific Reports, Microgravity, the Journal of Vision, and Multisensory Research. Supervised 35 graduate students and 9 post-doctoral fellows. Edited eleven books on vision-related topics. Research concentrates on how information from multiple senses is combined especially in unusual environments such as in space. Funded by NSERC, the Canadian Space Agency, the Canadian Foundation for Innovation, VISTA.

TALK: HOW I LEARNED TO STOP WORRYING AND LOVE THE PULL OF GRAVITY

POSTERS

Poster Session 1

Monday, June 6 (5:00–7:00pm)

Contribution of individual differences in imagery to episodic memory and spatial navigation	Adrienne Li	1
Differences in resting-state functional connectivity underlie visuomotor task performance declines in older adults with a genetic (APOE e4) risk for Alzheimer's disease	Alica Rogojin	2
The direction of optic flow affects perceived travel distance	Ambika Tara Bansal	3
Minimizing pedestrian falls on outdoor walkways: Minimum foot clearance estimation (MFCE) system	Anchana Kuganesan	4
Video-based traffic analytics at intersection	Sajjad Pakdamansavoji	5
Consumer-level facial tracking prototype for non-immersive virtual reality reminiscence therapy	Daniel Presas	6
Lightness and brightness characterized via decision spaces, in real and rendered scenes	Jaykishan Patel	7
Perception of self- and externally-generated visual stimuli: Evidence from EEG and behaviour	Edward Ody	8
Cortical integration of multimodal cues for reach and grasp planning	Gaelle Nsamba Luabeya	9
Evidence for a temporal cortex attention network in humans and monkeys	Hamid Ramezanpour	10
Automated usability assessment of individuals through outdoor street crossings and walkways	Jakson Paterson	11
A method for improving regression and correlation coefficient estimates in the presence of noise	Jason Pina	12
Influence of an allocentric cue shift on reach accuracy in monkeys	Jennifer Lin	13
Cognitive factors may affect the updating of an object's position during linear lateral translation.	John Jong-Jin Kim	14
A scoping review of the fear of falling research: Conceptual framework and empirical gaps	Kamila Kolpashnikova	15

Poster Session 2

Tuesday, June 7 (1:00–3:00pm)

The role of attention in acquiring frames	Mohammad Shams-Ahmar	1
Reconsidering the impact of disparity gradients on diplopia perception in natural stimuli	Arleen Aksay	2
The effect of frequent cannabis use on executive functioning components	Assel Al-Bayati	3
Visual hallucinations caused by Charles Bonnet Syndrome are not impacted by the psychosocial factors resulting from COVID-19 restrictions	Aysha Kinakool	4
Is the self-motion-induced bias in time-to-contact estimation attenuated when object motion is consistent with gravity?	Björn Jörges	5
The effect of perceived distance on cognitive performance	Carmel Camilleri	6
Role of the ventromedial prefrontal cortex in mnemonic discrimination	Claire Lauzon	7
Investigating facial emotion recognition in Parkinson's disease using COVID-19 face masks	Geneva Mariotti	8
Modularity and saccade influences in the cortical vision network	George Tomou	9
fMRI responses to real objects and pictures are differentially modulated by stimulus reachability	Grant Fairchild	10
The age-related effect of face masks on face identity and emotion perception	Jamie Cochrane	11
Adaptive eye-hand coordination when manipulating and monitoring the environment in parallel	Jolande Fooken	12
Stereoscopic distortions when viewing geometry does not match inter-pupillary distance	Jonathan Tong	13
Sex and sexual orientation differences in perceptual and cognitive processing	Katerina Andrinopoulos	14
The influence of sensory and cognitive abilities on perceptions of audio-visual impact events across the lifespan	Katherine Bak	15

Poster Session 3

Wednesday, June 8 (12:00–2:00pm)

Affective priming as a behavioural measure of implicit COVID-19 attitudes in people who are vaccine hesitant	Stefania S. Moro	1
Hybrid model-based / data-driven graph transform for image coding	Tam Thuc Do	2
Enhanced visual resolution for objects in the close space is mediated by affordance	Tasfia Ahsan	3
Extending the range of depth cameras using linear perspective	Tasneem Naheyan	4
Local & non-local factors in perceptual shape completion	Tenzin Chosang	5
Efficient single-view 3D vehicle ground-truthing for visual traffic analytics	Trong Thao Tran	6
Spatial coding in prefrontal visual responses during a reach task	Veronica Nacher	7
Spatial coding of visual targets in the frontal and supplementary eye fields	Vishal Bharmauria	8
Face-race modulates the perceived proximity of faces in adults, but not in children	Wei Fang	9
Binocular depth and distance cues enhance tolerance to virtual motion gain	Xue Teng	10
The influence of navigational experience on boundary effects in spatial memory	Xuehui Lei	11
ChartQA: A benchmark for question answering about charts with visual and logical reasoning	Ahmed Masry	12
Exploring the impact of immersion on situational awareness and trust in teleoperated maritime autonomous surface ships	Alex Gregor	13
Noise2NoiseFlow: Realistic camera noise modeling without clean images	Ali Maleky	14
Exploring volumetric video for increased realism in nurse debriefing	Colin Orian	15

Abstracts

IMPLEMENTING THE WATER LEVEL TASK IN AUGMENTED REALITY [3.26]

Romina Abadi & Robert S. Allison, York University

The "Virtual Water-Level Task" is developed based on Jean Piaget's experiments on spatial cognition in children. In the conventional Water Level Task, participants see a two-dimensional outline drawing of a vertically oriented bottle containing water and are asked to draw the water edge in pictures of the same container tilted to different degrees. Piaget proposed that as the spatial cognition abilities in children grow, they form a horizontal and vertical frame of reference with which they perceive the world. Therefore, children can understand that the water level remains horizontal and thus depict it correctly by about age nine. However, future studies showed some older children, adolescents, and adults had difficulties with the task and failed to draw a horizontal water edge. Up to now, most studies have been based on simple, schematic 2D presentations. Some researchers believe that most people understand how water behaves, but the pictorial nature of the task confuses this. My main goal is to determine if the misconception about the water orientation happens in a more natural environment, where the participants can interact with a simulated container and observe the water surface movements. I designed a "Virtual Water Level Task" using a Hololens2 device and a simple simulation of virtual fluid in a container to study the task in a more natural setting. The virtual fluid is created using a simple water effect, a surface inside the container that oscillates when the container moves or rotates. This method is used in games and creates the illusion of the liquid in the container. The surface oscillation is proportional to the container's acceleration and rotation speed. When the container does not move, the surface oscillation decreases until it becomes still. The simulation allows altering the surface direction when the container tilts so that the surface does not remain horizontal. The participants can see and interact with the containers, in some of which the water does not remain horizontal. They are then asked to determine if the simulation is realistic or not. I will compare the participants' performances in the conventional and virtual WLTs. I am running an online version of the traditional task and will recruit observers that are either accurate or prone to error to participate in the virtual WLT tasks. The results will help understand the reason some people fail the WLT.

REPRESENTATION AND INTEGRATION OF ALLOCENTRIC AND EGOCENTRIC VISUAL INFORMATION FOR GOAL-DIRECTED MOVEMENTS: A CONVOLUTIONAL / MULTILAYER PERCEPTRON NETWORK APPROACH [1.24]

Parisa Abedi Khoozani, Vishal Bharmauria, Adrian Schütz, Richard P. Wildes, & J. Douglas Crawford, York University

Allocentric (landmark-centered) and egocentric (eye-centered) visual information are optimally integrated for goal-directed movements. This process has been observed within the supplementary and frontal eye fields, but the underlying processes for this combination remain a puzzle, mainly due to inadequacy of current theoretical models to explain data at different levels (i.e., behavior, single neuron, and distributed network). The purpose of this study was to create and validate a theoretical framework for this process using physiologically constrained inputs and outputs. To implement a general framework, we modelled the visual system as a Convolutional Neural Network (CNN) and the sensorimotor transformation as a Multilayer Perceptron (MLP). The network was trained on a task where a landmark shifted relative to the saccade target. These parameters were input to the CNN, the CNN output and initial gaze position to the MLP, and a decoder transformed MLP output into saccade vectors. The network was trained on both idealized and actual monkey gaze behavior. Decoded saccade

output replicated idealized training sets with various allocentric weightings, and actual monkey data (Bharmauria et al. Cerebral Cortex 2020) where the landmark shift had a partial influence (R2 = 0.80). Furthermore, MLP output units accurately simulated motor response field shifts recorded from monkeys (including open-ended response fields that shifted partially with the landmark) during the same paradigm. These results suggest that our framework works and provides a suitable tool to study the underlying mechanisms of allocentric-egocentric integration and other complex visuomotor behaviors.

VISUAL ILLUSIONS MODULATE PERCEPTION AND ACTION IN AUTISM [3.28]

Zoha Ahmad, Noam Karsh, Orly Aziza, Tzvi Ganel, Bat-Sheva Hadad, & Erez Freud, York University

Perceptual changes are a widely acknowledged but poorly understood feature of autism spectrum disorder (ASD). One hypothesis is that these changes are the result of reduced specialization of the different cortical systems. The goal of the current study is to examine this hypothesis in the context of the functional dissociation between perception and action. Past studies have demonstrated that unlike their robust effect on perception, visual illusions have little or no effect on grasping movements. Here, we examine whether this dissociation is observed in individuals with ASD using a version of the Ponzo illusion. We used two objects that differ in size (42 mm, 40 mm), and placed on an illusory background such that they were perceived as "close" or "far" from the observer. We examined the participants' sensitivity to object size differences and their susceptibility to the illusion. Both groups were equally affected by the illusion in the manual estimation (perceptual) task, such that their estimations were larger for objects placed on the far surface. In contrast, in the grasping task, the two groups exhibited differential susceptibility to the illusion. Particularly, controls were not affected by the illusion, such that their maximum grip apertures (MGA) were scaled to the size of the object regardless of the illusory context. In contrast, the MGA of individuals with ASD were modulated by the illusion with larger apertures for objects placed on the "far" surface. Importantly, the observed susceptibility to the illusion could not be attributed to a deficit in motor control because sensitivity to size was similar for the two groups. Taken together, these results provide evidence for reduced functional dissociation between perception and action in ASD.

ENHANCED VISUAL RESOLUTION FOR OBJECTS IN THE CLOSE SPACE IS MEDIATED BY AFFORDANCE [3.3]

Tasfia Ahsan, Laurie M. Wilcox, & Erez Freud, York University

Depth perception is a critical aspect of the human visual system which promotes localization and interactions with objects in space. Over the years, psychophysical and neuropsychological research have revealed that sensory processing is modulated by the location of stimuli relative to the body. In particular, objects closer to the observer has shown an enhanced processing benefit, appropriately termed the "close advantage" effect. Previously the close advantage was limited only to multisensory research, with little focus on a purely visual advantage. However, recent investigations show that depth information modulates visuo-perceptual resolution, even when other visual attributes (such as retinal size) are matched. Despite these advancements, the underlying mechanisms that give rise to the close advantage effect are still unexplored. One hypothesis is that objects that perceived as closer to the observer might be more behaviorally relevant (e.g., one can grasp a hammer only if it is within arm's reach), and therefore benefit from enhanced processing. To evaluate this hypothesis, we quantify the close advantage effect for shapes that either offer a potential for action (affordable – e.g., elongated) or do not potentiate action (non-affordable – e.g., non-elongated). The stimuli were presented on either the 'close' and 'far' versions of the Ponzo Illusion. We used a method of constant stimuli to measure

the Just Noticeable Differences (JND) and Reaction Times (RT). Our results replicate previous findings by demonstrating an enhanced discrimination for both affordable and non-affordable shapes in the "close" space. Critically, and consistent with our hypothesis, affordable shapes had a stronger "close advantage" effect than non-affordable shapes. Together, these results provide novel evidence for the role of affordance in mediating the close advantage effect.

RECONSIDERING THE IMPACT OF DISPARITY GRADIENTS ON DIPLOPIA PERCEPTION IN NATURAL STIMULI [2.2]

Arleen Aksay & Laurie M. Wilcox, York University

The limiting disparities for binocular fusion, which define the boundaries of Panum's fusional area, are constrained by the 2D distance (lateral separation) between two disparate points. The ratio of disparity to separation defines the disparity gradient, and the disparity gradient limit is the ratio at which fusion of both elements is not possible simultaneously. This limit has been found to be approximately 1, though it can vary as a function of viewing duration and stimulus size and context. Importantly, studies of the disparity gradient limit have used simple isolated lines or dot patterns, which tend not to occur in real-world viewing environments. In the real world objects overlap and intersect, creating large disparity gradients that are well beyond the reported limits. Consequently, diplopia should often be apparent in natural viewing conditions; clearly this is not the case. The aim of this project is to understand how the visual system copes with large disparity gradients and with the resultant loss of fusion. As a starting point we measured diplopia thresholds in complex 3D scenes presented in a virtual reality head-mounted display. The scene consisted of realistic tree structures with a range of disparities modelled using Blender, presented in the headset using Unity. A central triad of branches comprised the target framed by two reference branches. Multiple other branches with uncrossed disparities surrounded the triad forming a volume. The relative disparity of the central target branch was varied across trials using the method of constant stimuli. The stimuli were displayed with a limited viewing time. Observers were instructed to maintain fixation on a continuously visible small fixation square, which appeared randomly just in front of the left or right reference branch and indicate a 'single' or 'double' percept of the target. The diplopia threshold was the point of subjective equality computed from a psychometric function fit to each observer's data. In a series of experiments, we measured 1) baseline thresholds with no intersecting branches 2) the effect of an undefined disparity gradient or 'disparity gradient limit violation' and 3) performance for isolated lines with the same size and layout. We found that diplopia thresholds were stable across all conditions when the tree structure was viewed; there was no effect of introducing disparity gradient limit violations. Comparison with isolated lines showed that thresholds were higher and more variable for the complex branching structure. Control experiments showed that these results were not due to the use of co-varying stimulus properties such as the length of the target branch. Our results challenge the long- held belief that large disparity gradients result in loss of fusion. To the contrary, we find that thresholds remain stable, even when stimuli overlap, and the disparity gradient limit is violated. Our working hypothesis is that in natural complex environments the visual system can capitalize on the spatial continuity of objects and surfaces to reduce the perception of diplopia.

THE EFFECT OF FREQUENT CANNABIS USE ON EXECUTIVE FUNCTIONING COMPONENTS [2.3]

Assel Al-Bayati, Holly Clayton, Ryan Cortez, Bernard Marius 't Hart, & Denise Y. P. Henriques, York University

The legalization of recreational cannabis use in Canada has raised many questions regarding its immediate and sustained effect on the performance of various critical daily tasks (e.g., driving). To

investigate the sustained effect, we created an online battery of tasks that assess the main components of executive functioning that are involved in all aspects of our daily activities. The performance of healthy, young (age= 17-64; M= 21.7, SD= 6.0) frequent cannabis users (at least once a week; n > 93), infrequent users (at least once in the last 3 months; n > 58), and nonusers (never used cannabis before; n > 253) was compared. Here, selective visual attention (using a serial visual search task), response inhibition (using a speeded Go/No-Go task), visuospatial working memory (using a visuospatial N-back task), and cognitive flexibility and set shifting abilities (using Trials B of the Trail Making Test) was analyzed. No significant or meaningful differences in performance were found on any of the measures of executive functioning components between frequent users, infrequent users, and nonusers. Overall, while there may be an impairment associated with the effects of acute cannabis use, our findings suggest that frequent cannabis use in healthy, young adults did not impair the main components of executive functioning.

EFFECTS OF THE BRAIN DERIVED NEUROTROPHIC FACTOR VAL66MET POLYMORPHISM ON CORTICAL STRUCTURE OF THE HUMAN BRAIN [2.31]

Vicente Alba Suarez, Katherine M. Newman, Georg Zoidl, & W. Dale Stevens, York University

Genetic variance among individuals and populations can influence the typical structure and function of the brain. One of the most important genes in the early development, maintenance, and plasticity of neural pathways is brain derived neurotrophic factor (BDNF) gene, which codes for the BDNF protein. The Val66Met single nucleotide polymorphism (SNP) of the BDNF gene has been studied extensively across various disciplines of neuroscience, but the influence this SNP has on structural differences across the brain and on functional brain networks has been largely unexplored. The few studies investigating this relationship are limited in sample size, and findings have been narrow in scope. Given the varied distribution of the Met allele across different ethnic groups, its impact on the structure and function of visual brain regions, and the differences in severity and prognosis for various disorders across different populations, this SNP could be impacting the reliability of fMRI studies. The proposed research aims to fill this knowledge gap using the Human Connectome Project (HCP), a large, open-source neuroimaging dataset that includes genetic, behavioural, and lifestyle data, in addition to a variety of neuroimaging protocols. With over 1000 participants, the HCP provides a comparatively large sample of neuroimaging and genetic data that can be studied to improve our fundamental understanding of the relationship between the Val66Met SNP of the BDNF gene and structural and functional differences in visual regions of the human brain. Here, we will provide an overview of how genotype differences for this SNP impact cortical thickness (CT) differences across the human brain. Analyses of CT provide insightful measurements of brain health at the level of the gray matter layer, whereby reduced CT in different regions may suggest reduced structural integrity of the brain. Using FreeSurfer, structural images were segmented into tissue types, separated by hemisphere, and adjusted to remove brain stem and cerebellum. From this newly generated brain surface, cortical thickness was measured as the shortest distance from the pial and white matter surfaces. This was done for each participant in our sample (n = 392). We present our findings on CT differences between genotype groups for Val66Met and outline future plans to analyze whole-brain functional connectome differences within and between functional networks for these genetically different groups.

SEX AND SEXUAL ORIENTATION DIFFERENCES IN PERCEPTUAL AND COGNITIVE PROCESSING [2.14]

Katerina Andrinopoulos & Jennifer K. E. Steeves, York University

Sex differences have been found in visual perception and cognition for some specific abilities. These abilities include mental rotation, face recognition and face detection. One way to measure mental

rotation is by using the mental rotation task (MRT) and is one of the largest tasks to demonstrate cognitive sex differences, with males routinely outperforming females (Voyer, 2011). Face recognition and detection show robust differences favouring females (McBain et al., 2009; Brewster et al., 2012). Few studies have considered the effect sexual orientation may have on these male and female advantages in visual cognition and perception. Same-sex attracted males tend to perform at the level of females in face recognition ability (Brewster et al., 2012). However, little research has focussed on same-sex attracted females and whether they also show cross-sex advantage of same-sex attracted males. It is possible that same-sex attracted females have a male typical advantage on mental rotation ability. This thesis seeks to elucidate the effect sexual orientation has on these visual and perceptual abilities that show sex differences. The typical male advantage was found for mental rotation ability, with heterosexual males outperforming heterosexual females. However, within the same-sex attracted groups, this difference was not found, with same-sex attracted females performing at the level of same-sex attracted males. No significant differences were found for face recognition, however there was a trend for same-sex attracted males and females as well as heterosexual females to outperform heterosexual males. Surprisingly, for face detection, all groups outperformed heterosexual females, bringing into question whether this group truly has an advantage or whether this specific sample of heterosexual females in particularly unskilled at face detection.

INCOMPATIBLE OCCLUSION AND BINOCULAR DISPARITY CAUSE SYSTEMATIC LOCALIZATION ERRORS IN AUGMENTED REALITY [2.30]

Domenic Au, Jonathan Tong, Robert S. Allison, & Laurie M. Wilcox, York University

Under natural viewing conditions binocular disparity can provide metric depth information; many monocular depth cues, such as occlusion, provide depth order only. Nonetheless, there is evidence that occlusion can influence both the direction and magnitude of perceived depth from stereopsis. In two experiments, we explored the integration of depth information from occlusion and binocular disparity in complex real-world environments using a distance matching paradigm. We used augmented reality (AR) to simultaneously present virtual and physical objects and varied 1) a virtual target's distance defined by binocular disparity and 2) the strength of an occlusion cue characterized by a physical surface's opacity. In both experiments, the virtual stimulus was a green letter 'A' presented using a Microsoft HoloLens 2 and superimposed on a real-world frontoparallel surface at 1.2 m. The letter was placed at one of eight positions – between 0.9 and 1.6 m, including the surface location. The letter was presented for 1 second and observers then moved a green spherical virtual probe to the remembered distance of the letter. For comparison, observers performed the same task without the physical surface. In Experiment 1, distances beyond the surface location produced an occlusion-disparity incompatibility where the letter should have been occluded under natural viewing, but instead remained superimposed in AR. In Experiment 2, the same conflict occurred with an opaque surface, but the strength of the conflict was altered by varying the opacity of the surface. Our results show that when the surface was absent, or the letter was rendered in front of the surface the letter was accurately localized. However, when the letter was rendered beyond the surface, observers progressively underestimated the letter's distance, even though the relative disparity between the probe and the target should have been equally informative at all locations. All surface opacities produced this underestimation, but overall, estimates were more accurate with a highly transparent surface. Our results suggest that 1) observers are unable to ignore contradictory occlusion when judging depth from binocular disparity and 2) such conflicts bias the perceived position of the target in the direction of the occluder even when highly transparent. Our results were modelled using a Bayesian ideal observer with an asymmetric distribution for an occlusion cue representing letter positions in front of vs beyond the surface for different surface

opacities. The goodness of model fits varied across observers, but it captured the overall pattern of results. In addition to providing insight into the integration of ordinal and metric depth information, these results speak to the impact of such errors on user interaction in AR.

SPECTRAL GRAPH LEARNING WITH FIRST EIGENVECTORS PRIOR VIA ITERATIVE GLASSO AND PROJECTION [3.27]

Saghar Bagheri, Tam Thuc Do, Gene Cheung, & Antonio Ortega, York University

Learning a suitable graph is an important precursor to many graph signal processing (GSP) modules, such as graph signal compression and denoising. Previous graph learning algorithms make assumptions in the nodal domain on i) graph connectivity (e.g., graph sparsity), and/or ii) edge weights (e.g., positive edges only). In this paper, given an empirical covariance matrix C computed from limited observable data, we consider a structural assumption in the spectral domain on the graph Laplacian matrix L: the first K eigenvectors of L are pre-selected, e.g., based on domain-specific knowledge, and the remaining eigenvectors are then learned from C. One example use case is image coding, where the first eigenvector is chosen to be constant a priori. We first prove that, inside a defined Hilbert space, the subspace of symmetric positive semi-definite (PSD) matrices with a common set of pre-chosen first K eigenvectors is a convex cone. We then construct an operator to project a given positive definite (PD) matrix L to this subspace, inspired by the Gram-Schmidt procedure. Finally, we design an efficient hybrid graphical lasso/projection algorithm to compute the most suitable graph Laplacian matrix L^* in the subspace given C. Experimental results on Gaussian random Markov field (GMRF) synthetic data, political voting data with negative correlations, and prediction residuals in image coding show that our algorithm with first K eigenvector prior outperformed competing graph learning schemes in the literature.

THE INFLUENCE OF SENSORY AND COGNITIVE ABILITIES ON PERCEPTIONS OF AUDIO-VISUAL IMPACT EVENTS ACROSS THE LIFESPAN [2.15]

Katherine Bak, George S. W. Chan, Michael Schutz, & Jennifer L. Campos, University of Toronto and The KITE Research Institute, University Health Network

Perceiving everyday events often requires the integration of multiple sensory inputs. With older age, declines to sensory (i.e., hearing, vision) and cognitive abilities may influence the way sensory inputs are integrated in the brain. Our previous work has examined age- related effects on audio-visual integration using an impact event paradigm called the Schutz-Lipscomb Illusion (SLI). The SLI paradigm involves the participant observing a marimbaist striking their instrument, producing a percussive sound. The illusion occurs when perceived tone duration is influenced by the length of a dynamic visual striking gesture (longer gesture = longer perceived tone). Previous results showed that both older and younger adults had similar perceived illusion strength, suggesting that audio-visual integration of impact events was not influenced by age in healthy older adults. However, sensory and cognitive impairments were strictly controlled (i.e., all participants had normal hearing and vision). Therefore, it remains unknown whether illusion strength is affected by age-related vision and hearing loss and across a range of cognitive abilities. Further, sensory loss often begins in middle age and yet illusion strength has not been evaluated in middle-aged adults. Therefore, the objective of this study was to examine whether subjective hearing, vision, and cognitive abilities affect the strength of the SLI in younger, middle-aged, and older adults. In an online study using Qualtrics, forty-six older (65+), 45 middle-aged (36-64), and 46 younger (18-35) adults viewed a single dot representing a striking gesture of a marimbaist (short or long) combined with a percussive tone and were asked to estimate tone duration. Participants also self-rated their hearing, vision, and cognitive abilities (poor to

excellent). Results demonstrated that illusion strength was similar across age groups and was not influenced by subjective sensory or cognitive abilities. Findings provide new insights toward how unisensory abilities of adults across the lifespan influence integration of audio-visual inputs to perceive realistic impact events.

THE DIRECTION OF OPTIC FLOW AFFECTS PERCEIVED TRAVEL DISTANCE [1.3]

Ambika Tara Bansal, Meaghan McManus, & Laurence R. Harris, York University

Although estimating travel distance is essential to our ability to move and navigate through the world, our distance estimates can be imprecise and inaccurate. When moving to the location of a previously seen target, the further the intended target distance, the more people tend to undershoot its location (Redlick et al., 2001). This phenomenon has been modelled as resulting from a leaky spatial integrator, meaning that these mis-estimations occur because (1) the integration "leaks" the further you move, and (2) there is a gain factor involved in transforming visual motion to travel distance (Lappe et al., 2007). The model has only been tested using forward translational movements, and postulates that perceived travel distance results from integration over distance and is independent of travel speed. Speed effects would imply integration over time as well as space. To test this, we measured participants' (n=15) perceived travel distance over a range of speeds (1, 3, 5 m/s) and distances in four different directions (up, down, forward, backward). Results show no effect of speed on either the gains or "leakage" (alphas). In terms of the effect of direction on gain, backward differed significantly from down ($p=0.01$), and up ($p<0.001$). In terms of the alphas, backwards differed significantly from forward ($p<0.001$), down ($p=0.01$), and up ($p=0.004$), and forward also differed significantly from down ($p<0.05$). These findings show that when making these distance estimates, we are likely integrating over space and not time, which means that Lappe's leaky spatial integrator model does not to be expanded to include a speed term. The effects of direction shows that transforming visual motion into travel distance differs depending on the direction of movement, and that the Lappe model can be applied to movement in these directions as well.

SPATIAL CODING OF VISUAL TARGETS IN THE FRONTAL AND SUPPLEMENTARY EYE FIELDS [3.8]

Vishal Bharmauria, Adrian Schütz, Xiaogang Yan, Hongying Wang, Frank Bremmer, & J. Douglas Crawford, York University

Previous neurophysiological studies have suggested that visual response fields in the gaze control system primarily utilize an eye-centered frame of reference at both the single unit and population level (Sajad et al. 2015; Sadeh et al. 2016). We recently confirmed that visual responses to saccade targets in the monkey frontal (FEF) and supplementary eye fields (SEF) showed a preference for eye-centered coding at the population level, even in the presence of a large visual landmark (Bharmauria et al. 2020, 2021). Specifically, visual activity showed less variance when plotted relative to target-in-eye coordinates (Te) as opposed to Th (head) or Ts (space) coordinates. Here, we re-examined this database at the single unit level, and used a more sensitive spatial continuum analysis in the spatially tuned visual responses of 101 FEF neurons and 43 SEF neurons. We found that, consistent with previous reports, approximately 80 % of tuned FEF neurons preferentially coded for Te. However, in the case of SEF, neurons almost equally preferred Te, Th and Ts. Our continuum analysis (of fits between these three models, i.e., Te-Ts, Te-Th and Ts-Th) showed distribution peaks near Te, Th and Ts with a sparse scattering of best fits in between these continua. These data suggest that whereas FEF retains a fairly simple egocentric visual code, SEF visual responses can show more complex levels of spatial preprocessing (at least in the presence of a visual landmark). This result seems to complement the finding where it was shown that SEF stimulation is capable of producing gaze shifts toward eye, head, or space-fixed goals (Martinez-Trujillo et al. 2004).

FROM EARTH TO SPACE: THE EFFECT OF GRAVITY AND SEX ON SELF-MOTION PERCEPTION [1.23]

Nils-Alexander Bury, Laurence R. Harris, Michael Jenkin, Robert S. Allison, Sandra Felsner, & Rainer Herpers, Hochschule Bonn-Rhein-Sieg

Humans are inherently multisensory, perceptual agents. We combine vision, audition, touch, vestibular, and proprioceptive cues to navigate and interact with the world which helps us distinguish self-generated acceleration from the acceleration of gravity. But what happens when there is no gravity? The effectiveness of visual motion in evoking the perception of self-motion has shown both over-/underestimation depending on the task, a contradiction that has been effectively resolved by the Lappe Model (Lappe et al., 2007) which predicts both phenomena using only two parameters: gain (output/input) and a spatial decay (perceived distance reduced by leak a factor over distance). Here, we used tasks associated with underestimation (the adjust-target task: AT) and overestimation (the move-to-target task: MTT) to look for changes associated with self-motion under multiple gravity states. 24 participants (12 males, 33.6±7.2yrs.; 12 female, 33.9±6.0yrs.) were tested during parabolic flights that created one period of microgravity (~22s) and two periods of hypergravity (~25s) per parabola. Participants were tested: twice on the ground before flight (1&2), during level flight (3), in the hypergravity (4), and microgravity (5) phases of the flight, and after the flight (6). For all sessions, participants were tested either lying supine or free-floating. They wore a virtual reality head-mounted display presenting optic flow that elicited perceived forward self-motion along a corridor. We measured their perceived "travel distance" in an egocentrically upright, visual simulation. Participants performed two tasks in separate blocks in each of the 6 sessions: (1) "MTT" in which they moved visually to the remembered position of a previously presented target, and (2) "AT" in which they were first moved visually through a given distance and then adjusted a target's position to indicate how far they felt they had just traveled. The data were fitted using the Leaky Spatial Integrator model (Lappe et al., 2007) to determine the gain and decay factors for each condition. For the MTT task, participants slightly underestimated traveled distance for most short distances (<18m) and overestimated it for the longer distances (>18m). The results for the AT task showed an unexpected overestimation for each distance. There was a significant difference between distance estimates of the MTT task and the AT task, and a tendency for both tasks to differ between the 1-g conditions (on ground and in level flight) and under both hypergravity and microgravity. In both altered g-levels participants' gains were higher than during level flight with the most convincing effects found in females. Our experiment so far reveals differences between the perceived travel distances in the MTT task and the AT task which are partly compatible with the predictions of the Leaky Spatial Integrator model of self-motion perception. We also reveal a potentiation of self-motion perception in hyper- and microgravity. This agrees with anecdotal reports of increased effectiveness of optic flow in inducing self-motion in space and reveals a role of gravity in the perception of linear self-motion. However, differences in gravity might change the perception of self-motion for females only.

THE EFFECT OF PERCEIVED DISTANCE ON COGNITIVE PERFORMANCE [2.6]

Carmel Camilleri, Tasfia Ahsan, Laurie M. Wilcox, & Erez Freud, York University

Depth information is a vital component that supports various perceptual processes—accumulating evidence point to enhanced visual processing of the close space when defined by either pictorial or binocular cues. However, an outstanding question is whether this enhancement reflects attentional modulations (that is – we dedicate more attentional resources to the close surface, which in turn increases the precision of visual processing), or whether this enhancement is dissociable from attention (that is attention is allocated similarly across the close and far spaces). Here we report on two experiments in which participants were asked to detect a target positioned on either the close or far

surface of the Ponzo illusion. The location of the task was cued (75% cue validity) by a flash of light on one of the surfaces (Experiment 1, exogenous cue) or by an arrow presented at the fixation point (Experiment 2, endogenous cue). In Experiment 1, shorter reaction times (RTs) were observed for valid trials, across target location (i.e., close vs far). Additionally, participants detected the closer targets faster, demonstrating the close advantage effect. In contrast, in Experiment 2, we found an interaction between cue validity and target location, such that valid cues facilitated target detection more for targets presented on the closer surface. Taken together, these results suggest that voluntary (endogenous), but not non-voluntary (exogenous), attentional processes might contribute to the privileged processing of closer objects.

LOCAL & NON-LOCAL FACTORS IN PERCEPTUAL SHAPE COMPLETION [3.5]

Tenzin Chosang, Keyi Liu, & James H. Elder, York University

Humans are known to rely profoundly on bounding contours for object detection, segmentation and recognition. This task is complicated by the interposition of other objects in the visual field that lead to occlusion: partial blocking of one object by another. However, the impact of occlusion on human perception is mitigated by the human ability to perceptually complete partially occluded bounding contours, i.e., to fill-in the missing shape information. Here we examine the degree to which the human brain uses local and non-local cues to solve this perceptual completion task. Each visual stimulus consisted of a sequence of dots regularly sampling the outline of a 2D shape. To simulate occlusion, a contiguous interval of 10-50% of the dot pattern was extinguished. Observers were asked to adjust a probe dot along a linear axis orthogonal to the gap until the dot appeared to lie where the contour would be, were it visible. Two classes of shape were employed: animal shapes, which afford both local and global cues to completion, and metamer shapes, which match the curvature statistics of the animal shapes but are otherwise random, thus affording local but not global cues to completion. Mean absolute error was lower for animal shapes than for metamer shapes and, while completions tended to be negatively (inward) biased for both, the bias was less for animals than metamers. Together these findings point to a contribution of non-local shape cues to perceptual contour completion.

AGE-RELATED EFFECT OF FACE MASKS ON FACE IDENTITY AND EMOTION PERCEPTION [2.11]

Jamie Cochrane, M. Eric Cui, Eugenie Roudaia, Björn Herrmann, Allison B. Sekuler, & Patrick J. Bennett, McMaster University

The current study aims to investigate the effects of masks on the discrimination of facial identity and emotional expressions in younger and older adults. Masks that occlude the lower half of the face should impair the discrimination of disgust and happiness, which rely largely on visual structure near the mouth. However, masks may affect identification differently in younger and older adults: young adults rely strongly on informative structure near the eyes for face identification, whereas older adults are less specifically reliant on the eye region for face identification. Using a Garner Interference Task, participants judged if two female faces presented sequentially showed different persons (identity task) or different emotional expressions (emotion task). Faces were randomly shown in three blocked conditions: unmasked faces, faces with the lower half covered by a surgical mask, or faces with features in the lower half removed. Obscuring the lower half of the face, by mask or removal, significantly impaired emotion perception in both age groups. Face identification was more accurate in younger adults than older adults in all conditions, but masks affected performance differently in two age groups. Specifically, obscuring the lower half improved identification in older adults, but did not significantly impact performance in younger observers. These results suggest that masks may affect social communication differently in younger and older adults.

THETA BURST STIMULATION DOES NOT MODULATE FUNCTIONAL CONNECTIVITY IN THE PRIMARY VISUAL CORTEX: A SHAM-CONTROLLED MULTI-ECHO FMRI STUDY [2.25]

Remy Cohan, Karlene S. Stoby, Diana J. Gorbet, Sara A. Rafique & Jennifer K. E. Steeves, York University

We previously examined the effects of a single 20 min session of low frequency (1 Hz) repetitive transcranial magnetic stimulation (rTMS) to primary visual cortex (V1) on resting-state functional connectivity (FC) and found widespread FC changes not immediately but 1 hr following stimulation (Rafique & Steeves, 2022). Theta burst stimulation (TBS) is a sub-protocol of rTMS which has the advantage of a short delivery time over traditional rTMS. TBS delivers a train of bursts of three magnetic pulses at 50 Hz every 200 ms (5 Hz) to targeted brain regions. When applied to motor cortex, intermittent TBS (iTBS; 2s trains of TBS repeated every 10 s for 190s) has been shown to yield excitatory aftereffects, whereas continuous TBS (cTBS; a continuous 40 s train of TBS) may lead to inhibitory aftereffects, both lasting from minutes to hours (Huang et al., 2005). The majority of TBS research has targeted motor, frontal and parietal regions, and to date very few studies have examined its efficacy at visual areas. We designed a sham-controlled study to investigate the short-term (immediately post-stimulation) and longer-term (1 hr post-stimulation) effects of iTBS and cTBS to V1. Using multi-echo functional magnetic resonance imaging, we compared resting state FC before and after stimulation in seeds from V1 (stimulation site) to neighbouring networks. There were no changes in FC between the iTBS, cTBS and sham stimulation groups, as well as no within-group changes in FC from baseline to post-stimulation timepoints. Our results indicate that unlike low frequency rTMS to V1, TBS to V1 does not produce widespread FC alterations at any timepoints.

THE EFFECT OF CHRONIC CANNABIS USE ON VISUOMOTOR AND COGNITIVE FUNCTION [1.25]

Ryan Cortez, Assel Al-Bayati, Raphael Q. Gastrock, Bernard Marius 't Hart, & Denise Y. P. Henriques, York University

Since the legalization of recreational cannabis use in 2018, there has been an increased interest in the potential detriments related to its usage. This extends to the visuomotor and cognitive abilities that we rely on to complete activities of daily living, where deficits in these abilities may increase the risk of harm or injury. Hence, our goal in this study is to determine if deficits exist in the visuomotor and cognitive performance of cannabis users. York University students (N~880) completed a browser-based battery of well-established tasks that assess visuomotor and cognitive performance. Three of the eight tasks will be analyzed: 1) a task switching task (cognitive flexibility, attention switching) where the stimulus-response mapping changes every other trial, a tunnelling task (visuomotor acuity) where participants move a cursor through a standard track at different scales, and a mirror reversal task (goal-directed planning) where participants move a cursor to a target with the cursor position flipped along the y-axis. We compared the performance of non-users (no prior usage; N>=160), infrequent users (>once per 3 months & <once per week; N>=71), and frequent users (once per week or more; N>=89) on all three tasks. We found no significant performance deficits associated with cannabis use, suggesting that cannabis use does not lead to persisting visuomotor and cognitive deficits.

THE FEEDBACK-DEPENDENT TIME-COURSE OF IMPLICIT LEARNING [1.28]

Sebastian D'Amario, Jennifer E. Ruttle, Bernard Marius 't Hart, & Denise Y. P. Henriques, York University

People constantly adapt their movements to their changing circumstances, which is mostly handled by our automatic, unaware, or implicit motor adaptation systems. While the time course of these implicit processes is thought to be slow, this is actually largely unknown. Motor adaptation is usually induced by

having people reach to targets with a cursor whose motion is misaligned with respect to their unseen hand. Here, I have tested the effects of various kinds of feedback of the unseen hand motion on the speed of implicit learning. One group, ("continuous"), received the typical visual feedback, i.e. the hand-cursor was continuously visible, but deviated 45° during the reach. The second group, ("terminal"), used impoverished visual feedback, i.e. they only saw a static cursor at the very end of each reach. For the third group, ("cursor jump"), the cursor was aligned with the unseen hand at the start of each reach, but jumped to the misaligned direction of 45° in order to show the source and nature of the perturbation on each trial. All groups completed the same rotation schedule beginning with 20 aligned trials, we then abruptly rotated the cursor 45° for 100 trials and then flipped it for 8 with a final error clamped phase where the cursor moved straight to the target regardless of hand movement direction. After every training trial participants completed a no-cursor trial, where the cursor was not visible and they were instructed to not use a strategy to probe implicit adaptation. By alternating between training and testing trials, we could measure the rate of implicit learning at a fine temporal resolution. For both the "terminal" and "cursor-jump" groups, we hypothesized a decrease in the magnitude of implicit adaptation (as measured by exclusion reach aftereffects) and possibly on the rate of implicit adaptation. For the training trials, the continuous group showed a higher learning rate and thus reached asymptote faster (16 trials to saturation vs 55 for terminal and 61 for cursor jump). Overall adaptation was surprisingly smaller for the continuous group with an asymptote of 26° vs 37° for terminal and 33° for cursor jump. Testing trials, however, showed more implicit learning for the continuous group at 22° and the slowest rate of change requiring 22 trials to hit asymptote vs 6 for terminal and 11 for cursor jump. Our results indicate that despite a lower final amount of adaptation, implicit learning saturates faster with changed visual feedback. The data here also confirms our earlier finding that implicit adaptation is much faster than usually assumed. The observation that implicit adaptation can be changed, despite being mostly automatic, suggests that there may also be useful feedback-dependent mechanisms which can increase the amount as well as the rate of implicit learning.

HYBRID MODEL-BASED / DATA-DRIVEN GRAPH TRANSFORM FOR IMAGE CODING [3.2]

Tam Thuc Do, Saghar Bagheri, Gene Cheung, & Antonio Ortega, York University

Transform coding to sparsify signal representations remains crucial in an image compression pipeline. While the Karhunen-Loève transform (KLT) computed from an empirical covariance matrix C is theoretically optimal for a stationary process, in practice, collecting sufficient statistics from a non-stationary image to reliably estimate C can be difficult. In this paper, to encode an intra- prediction residual block, we pursue a hybrid model-based / data-driven approach: the first K eigenvectors of a transform matrix are derived from a statistical model, e.g., the asymmetric discrete sine transform (ADST), for stability, while the remaining N-K are computed from C for performance. The transform computation is posed as a graph learning problem, where we seek a graph Laplacian matrix minimizing a graphical lasso objective inside a convex cone sharing the first K eigenvectors in a Hilbert space of real symmetric matrices. We efficiently solve the problem via augmented Lagrangian relaxation and proximal gradient (PG). Using WebP as a baseline image codec, experimental results show that our hybrid graph transform achieved better energy compaction than default discrete cosine transform (DCT) and better stability than KLT.

MULTI-USER VR SIMULATOR FOR ROBOT DEVELOPMENT [3.31]

Silas Franco dos Reis Alves, Alvaro Uribe-Quevedo, & Jon Morris, Ontario Tech University

Access restrictions to long-term care facilities (LTCFs) due to the COVID-19 pandemic disrupted service robot research and development. The lack of access to LTCFs ignited the interest in using digital twins to

overcome the restrictions associated with data collection and testing. We propose a multi-user Virtual Reality (VR) Simulator that allows remote teams to co-design social robot applications and fine-tune robot navigation and fall detection employing virtual avatars. We employed the Aether robot developed by JDQ Systems to assist caregivers in LTCFs by providing companionship while monitoring falls, water spills, or tripping objects. Exit Game's Photon was used to add multi-user and voice-chat capabilities to our previous simulator. The digital twin environment the virtual version of Aether to be transferable to the real world. The Robot Operating System (ROS) was integrated into Unity and run on a Ubuntu server. Once inside the VR simulation, the user is able to interact with the virtual Aether by employing upper limb gestures and voice commands. We used the digital twin of JDQ's office space for testing the Simultaneous Localization and Mapping (SLAM) method provided by Cartographer. The map built with real sensor data had an area of 27.7 m2, while the map built with simulated sensor data had 29.99 m2. Both virtual and real robots used the same set of parameters for navigation and completed a set of 10 trips to 4 different locations. The average distance from the actual goal measured with Augmented Reality tag was of 0.18 m (σ=0.11) for the real robot, and 0.07 m (σ=0.03) for the virtual robot. To evaluate the fall detection, we used the official YOLOv4 classifier trained on COCO and retrained it with the Fallen Person Data Set (FPDS). When tested against the FPDS Test Set, the retrained weights improved the Average Precision, Precision, and Recall at an IoU (intersection over union) of 0.75 and 0.5. YOLOv4 was also able to detect virtual avatars, which will allow us to leverage the simulator for improving Aether's fall detection behaviours. The simulator can produce similar navigation and voice-interaction results as the real robot, expediting robot development even with restrictions to real-world testing or when working with remote teams. The simulator also saved us time and costs associated with real-world testing, such as test planning, transportation, and staff hours. Future work will focus on improving the error model of the virtual sensors and actuators so they can accurately reproduce the noise observed in their real counterparts, and adding different avatars for users to impersonate and produce synthetic data sets.

WHAT CAUSES THE FACIAL ORIENTATION BIAS? [1.22]

Maxwell Esser, Anne Thaler, & Nikolaus F. Troje, York University

When looking at a picture of a human face in half-profile view, human observers tend to overestimate the orientation of the face. For instance, if oriented 30° with respect to the fronto-parallel view, observers asked to estimate the orientation indicate an angle of about 45°. What is the cause for this facial orientation bias? Is it specific to faces? Does it happen only in pictures? We tested three different hypotheses. (1) In pictures, the planar shape of the medium induces depth compression of the depicted content. (2) The optic array coming from an image is interpreted as if it came from a 3D object. If so, disparities between the picture's center-of-projection and the observer's vantage point cause depth distortions. (3) Unless additional information is provided, the visual system uses a shape prior that assumes a circular horizontal cross section, rather than matching the elliptical cross section of a typical head. Experiment 1 was conducted in virtual reality where we simulated both 3D heads and images of them. We used a method of adjustment ("orient this face into a 45° position"). We found the orientation bias was much smaller than expected and only marginally different between picture and 3D conditions. In Experiment 2 we presented static pictures and used a method of constant stimuli ("is the person oriented more or less than 45° from frontal view?"). We systematically varied the center-of-projection of the picture while keeping the observer's vantage point constant. We observed a pronounced bias, but that bias did not depend on the point of projection. We hypothesized that the differences in experienced orientation bias between Experiments 1 and 2 were due to participants having access to shape-from-motion in Experiment 1, but not in Experiment 2. In Experiment 3, we replicated the orientation bias with a non-facial object – a coffee mug with a handle that defined its

orientation. We systematically modified the shape of the mug between circular and elliptical horizontal cross sections. Mugs were then presented either as static images or as short animations with the mug rotating about its vertical axis. Participants displayed predictable errors when estimating the orientation of elliptical shapes. Additionally, the shape dependent orientation biases were much smaller for the animations compared to the pictures. We conclude that the visual system adopts the heuristic of a cylindrical head shape unless explicit information about its shape is provided, for example, through structure-from-motion which was available in the animations but not the static images.

FMRI RESPONSES TO REAL OBJECTS AND PICTURES ARE DIFFERENTIALLY MODULATED BY STIMULUS REACHABILITY [2.10]

Grant Fairchild, Desiree E. Holler, Sara Fabri, Michael A. Gomez, & Jacqueline C. Snow, University of Nevada, Reno

Most studies of human vision use two-dimensional (2-D) pictorial stimuli as a substitute for the real environment. Importantly, however, pictures lack numerous properties of real objects; in particular, real objects have the potential for physical interaction whereas pictures do not. Recent studies from our lab have shown that the actability of real objects both elicits different behavioral responses from 2-D pictures, and modulates differences between real objects and pictures in EEG signatures of motor preparation in dorsal cortex. Here, we used fMRI to localize the effect of the potential for interaction on neural responses towards meaningful everyday real objects (tools) and closely matched printed 2-D pictures of those objects. Because physical size has been shown to modulate object perception, and could influence actability, all stimuli were presented in their typical real-world size. In the scanner, the real objects and pictures were positioned either within reach of participants' hands, or beyond their reach. Participants performed a 1- back task in which they indicated using a button-press response whether the hand action they would perform to use each tool was similar or different to that of the tool shown on the previous trial (e.g., the hand action to use a hammer is similar to that of a cleaver, but different to that of a key). We found that real objects elicited stronger BOLD activation than pictures did in ventro-lateral occipito-temporal cortex and in posterior parietal cortex –both of which are typically associated with object form processing. Intriguingly, these same areas also showed a significant interaction effect, where stimulus reachability modulated BOLD activation more strongly for real objects than it did for pictures. The same interaction effect was also observed in ventro-medial temporal cortex in areas typically associated with texture processing. Few, if any, brain areas showed the reverse interaction effect (where reachability would modulate BOLD activation more strongly for pictures than for real objects). These results suggest that real objects are processed with respect to their egocentric distance more so than pictures are –perhaps because of their actability. Together, these data suggest that real objects elicit quantitatively and qualitatively different brain responses than do artificial stimuli, not only along the dorsal 'action' pathway, but also throughout the ventral 'perceptual' pathway.

FACE-RACE MODULATES THE PERCEIVED PROXIMITY OF FACES IN ADULTS, BUT NOT IN CHILDREN [3.9]

Wei Fang, Elysha Spring, Cristina I. Galusca, Zhe Wang, Yu-Hao Sun, Olivier Pascalis, & Naiqi G. Xiao, McMaster University

Recent face perception studies suggest that our representation of faces closely relates to spatial information. This spatial dependency can be revealed in a face-specific illusion, referred to as the "Fat Face" illusion: when two identical faces are aligned vertically and the bottom one is consistently perceived as larger. This illusion suggests that our visual system uses the vertical layout to implicitly

infer the proximity of faces. The face lower in the picture plane is perceived as closer, therefore being seen as larger, than the top face. While the "Fat Face" illusion is highly robust in adults, its developmental origins remain largely unknown. The face-specificity implies a developmental change driven by experiences with faces in the environment. Here, we probed the development of this illusion via the impact of face-race information on the illusion in children and adults. Adult participants (N=34, 17 Asian, 17 Caucasian) saw two identical faces vertically aligned and clicked the one they perceived as larger. The faces were morphed from Asian and Caucasian faces at 5 morphing proportions (0%, 25%, 50%, 75%, & 100% of Asian). Overall, adults misperceived the bottom face as larger in 68% of trials ($p<.001$). Moreover, adults exhibited a larger illusion for faces with higher morphing proportions of other-race faces ($p=.007$), suggesting that other-race faces were perceived as closer than own-race faces. Children (7.11yrs, 5 to 10yrs, N=49) also showed the illusion (70% of trials, $p<.001$), but their illusion was not affected by face- race ($p=.214$). The existence of the illusion in children suggests that our sensitivity to face locations might emerge early in development. The difference in the effect of face-race between children and adults indicated that the tuning to face location undergoes further development, likely driven by growing experience with face categories.

ADAPTIVE EYE-HAND COORDINATION WHEN MANIPULATING AND MONITORING THE ENVIRONMENT IN PARALLEL [2.12]

Jolande Fooken, Roland S. Johansson, & J. Randall Flanagan, Queen's University

Research on eye-hand coordination has focused on actions performed in isolation. Here we examined eye-hand coordination in a ball- drop task that was performed in parallel with a perceptual task. Participants repeatedly grasped and inserted a ball in a slot, using either their fingertips or tweezers, while simultaneously detecting letter changes that occurred every 1.5 to 6.5 seconds. Compared to the action task performed in isolation, participants adapted their eye and hand movements in two ways. First, gaze shifts from the text display to the action were temporally coupled to the sub-goals of the task, specifically to ball grasp and slot entry. Moreover, gaze was more frequently directed to the action when performing the task with tweezers compared to fingertips. Second, participants modulated the timing of their hand movements, such that the probability that a letter change would occur at the time of ball grasp was reduced. In addition, participants were more likely to allocate gaze to the action task during the 1.5 seconds between successive letter changes during which no letter change could occur. These results indicate that people can actively modulate their eye-hand coordination to take advantage of temporal regularities in the environment when acting and perceiving in synchrony.

COMMUNICATION EFFICIENCY IN VIDEO CALLS WITH MOTION PARALLAX [3.22]

Ashley Funkhouser & Nikolaus F. Troje, York University

Due to the recent pandemic, video conferencing services (Skype, Zoom, Teams, etc.) have boomed as a medium of communication. Compared to face-to-face communication, communication through video conferencing is less efficient and more exhausting. We hypothesize that this is caused by a lack of directionality, which is the ability to discern how someone is oriented relative to ourselves. The lack of directionality becomes particularly obvious in the context of gaze direction and other deictic behaviors. No matter where a person is in front of their screen, that person only feels looked at if the other person is looking directly into the camera. Physical cues crucial to communication efficiency, such as pointing and eye gaze, rely heavily on directionality. Consequently, communicating through screens is less desirable than in real-life settings. Studies demonstrate that motion parallax facilitates the accurate evaluation of directionality. We hypothesize that adding directionality to video calls through motion parallax improves communication efficiency. We developed an app that adds motion parallax to a video

chat with head avatars. The camera on the device tracks the user's head movements and a virtual camera inside the program follows the coordinates from the physical camera to simulate motion parallax as if the user is talking to the other person's avatar through a window. Communication efficiency is evaluated using a tangram game, where two people cooperate to finish 8 unique puzzles. Participants assume two roles: instructor and student. The instructors have 45 seconds to explain the solution to the puzzle while the student is unable to see it. Then, communication between the players is disabled while the student completes the puzzle. There are two viewing conditions, one with motion parallax and one without. Participants experience both during the experiment, presented in random order. We measure the puzzle-completion time of the student. We expect them to be quicker in the condition with motion parallax than in the condition without motion parallax, as the added directionality provides certain nonverbal cues otherwise unavailable in a non-motion parallax condition. To account for potential differences in puzzle difficulty between trials, we performed a full analysis of variance to compare conditions for each puzzle. There were two factors, a within-subjects variable labeled 'viewing condition' with two levels (motion parallax and non- motion parallax) and one between-subjects factor labeled 'puzzle' with eight levels (puzzles one through eight). We expected that regardless of puzzle difficulty, the motion parallax condition would be faster overall. Should results prove to be promising, the real-world applications are quite wide, as video communication services are used in most settings nowadays. Communication efficiency is critical to business, academia, and more. Improving video communication would overall bolster productivity for everyone.

LEARNING IN A MIRROR REVERSAL TASK PROVIDES DISTINCT MECHANISMS BETWEEN DE NOVO LEARNING AND MOTOR ADAPTATION [2.24]

Raphael Q. Gastrock, Bernard Marius 't Hart, & Denise Y. P. Henriques, York University

When people encounter movement errors, they process these errors to correct for subsequent movements. Such error processing contributes to learning when we are either adapting well-known movements or acquiring new motor skills (de novo learning). Previous studies have compared these two types of motor learning, however, several aspects of de novo learning, including its retention and generalization, still warrant investigation. Here, participants completed an online version of the mirror reversal task, a paradigm that captures de novo learning mechanisms, across two sessions. In session 1 (N = 63), participants reached to three targets located in the upper-right quadrant of the workspace (5, 45, 85 degrees in polar coordinates), with the mirror located along the vertical midline axis. Although targets farther from the mirror axis produced larger errors, we found that asymptotic learning did not differ across target locations. Moreover, we observed quick progression in learning and no reach aftereffects, which are persistent deviations in reaches after perturbation removal and are indicative that adaptation has occurred. Interestingly, participants make faster and shorter movements towards the 5 degree target, while taking the longest amount of time and path to reach the 45 degree target. Participants returned for a second session (N = 48; days apart: M = 4.77, SD = 2.52), and showed retention of learning upon re-experiencing the perturbation. They then reached for corresponding target locations within the lower-right and upper-left quadrants of the workspace, followed by reaches using their opposite and untrained hand. We observed almost complete and near immediate generalization of learning to targets across the workspace and the opposite hand. Taken together, these results provide further behavioral mechanisms that distinguish de novo learning from adaptation.

SIGV: A MINI-LANGUAGE FOR TRANSMODAL LIVE CODING [1.20]

Marcus A. Gordon, York University

This research project brings together command line based utilities in a workflow for live coding and composition with modular synthesizers. The minimalist approach taken in this project seeks to reduce

the energy, cognitive load, data complexity, often brought on by software/programs via poor graphical user interfaces, poor user experience flows, and the overall management, support and longevity of them. In real-time performance and research scenarios, these deficiencies can impact artists negatively, in preparation, performance, and studio activity. The approach taken here counters these narratives with lightweight programs that are fast and efficient. The results of the research creation, demonstrate how the making of a minimal language, is paving a path towards the making of a custom environment for live coding, algorithmic composition and live performance. The making of sigv was inspired by the need to make a custom environment for live coding sound, image, and physical instruments. Improvisation, the semantics behind live coding, and modular synthesis were the premise behind the making of sigv and its components. The modular approach taken with sigv was to make the adding of modules quick and easy during live performance, allowing the performer to start with the necessary modules they want or need, and alter that configuration as they perform.

EXPLORING THE IMPACT OF IMMERSION ON SITUATIONAL AWARENESS AND TRUST IN TELEOPERATED MARITIME AUTONOMOUS SURFACE SHIPS [3.13]

Alex Gregor, Robert S. Allison, Onoise Gerald Kio, & Kevin Heffner, York University

Maritime Autonomous Surface Shipping (MASS) has potential benefits for mitigated environmental impact, reduced risk for seafarers, and increased economic efficiency. MASS systems will require remote human monitoring which would occur at Shore Control Centres (SCCs). Key factors to consider when building interfaces for future MASS systems are increased Situational Awareness (SA), and calibrated Trust, which serve to mitigate risk for SCC Operators (SCCOs) when properly accounted for. A concept that might positively impact these cognitive constructs is one's sense of immersion - a design factor associated with mixed reality (MR) technologies. Immersion involves presenting, and constraining, a user's perceptual experience to a virtual or remote environment. As such, increased field-of-view, peripheral visual occlusion, naturalistic control, haptic feedback, and focused auditory stimuli correlate with a greater sense of immersion. Since immersion depends on varying degrees of multi-sensory engagement, it is not binary and can have differing levels. This research seeks to examine how SA and Trust are affected by different immersion-levels in SCC interfaces, and in turn how these cognitive constructs correlate to each other. A simulation was developed to emulate how data would be received by, and displayed to, an SCCO and includes built-in assessment methods for SA and Trust. A user study will be performed with participants using the simulation on three different immersion-levels: a virtual reality (VR) head-mounted display (HMD), a concave projector system, and a standard computer monitor. Having different immersion-levels will help explore if there is a statistically-supported benefit to an SCCO's SA and Trust by using immersive MR equipment in SCCs. This work is a requisite step for the construction of future SCCs, and will consequently hold value for MASS stakeholders.

LOAD-DEPENDANT NEURAL NETWORK UNDERPINS MENTAL ATTENTION FUNCTION WHEN CAPACITY RESOURCES ARE EXCEEDED [2.22]

Mylann Guevara, Marie Arsalidou, Juan Pascual-Leone, & W. Dale Stevens, York University

Mental attention capacity refers to the amount of information that one can maintain and manipulate in conscious awareness while completing a cognitive task. This capacity is domain general and varies across individuals. As a fundamental cognitive resource, it facilitates much of our conscious cognitive ability to flexibly navigate complex visual scenes and spontaneously synthesize solutions to real world problems. A functional magnetic resonance imaging study of the neural underpinnings of visual mental attention capacity employed a task where 30 healthy young adults made same/different judgements about serially presented visual images with parametrically varied colour and number stimuli. Univariate

analysis demonstrated that brain regions associated with the dorsal attention network and default mode network parametrically increased in activation and decreased in activation, respectively, with rising task demand. Once participants reached their expressed capacity, amplitude of brain activity reversed, such that brain regions that previously increased with rising task demand sharply decreased in activity and vice versa for previously decreasing regions. This widespread crossover effect suggests that a major neurocognitive shift takes place during mental attention function when mental attention demand of the task begins to exceed an individual's capacity limit. There remains uncertainty about whether these regions are functionally interrelated, which is imperative for understanding mental attentional function mechanistically under the condition of high mental demand. Spatiotemporal task-related Partial Least Squares (task PLS) is a multivariate technique that analyzes covariance of activity across brain regions associated with experimental manipulations. The current study uses task PLS to identify brain regions and the functional relations among them that reliably demarcate the limits of mental attention function. We found a distributed pattern of brain activity functionally associated with task-demand exceeding mental attention capacity, suggesting the existence of a load- dependant mechanism within broader mental attention. These findings suggest that the visual attention system differentially orchestrates resources to manage the dynamic representation and manipulation of visual arrays exceeding capacity. These results not only delineate our mechanistic understanding of mental attention but provide targets for future analyses to further develop our models and understanding of visual perception/attention disorders.

EFFECT OF BINOCULAR DISPARITY ON FLOW PARSING [3.16]

Hongyi Guo & Robert S. Allison, York University

During locomotion, optic flow provides important information for detection, estimation and navigation. On the other hand, binocular disparity, which carries compelling depth information, can potentially aid optic flow parsing. We explored the effect of binocular disparity on observers' ability to detect and discriminate object motion during simulated locomotion with two experiments. In each experiment, twelve participants were recruited and tested on our wide-field stereoscopic environment (WISE). In the first experiment, the stimulus consisted of four spherical targets hovering in a pillar hallway presented under stereoscopic, synoptic (binocular but without disparity), and monocular viewing conditions. In each trial, one of the four targets moved either in depth (approaching or receding) or a frontal parallel direction (contracting or expanding). Participants detected the moving target during a simulated forward walking locomotion in a 4-alternative forced choice task. Adaptive staircases were adopted to obtain the thresholds of the target motion speed in each viewing condition. The results showed that when stationary, binocular summation contributed more to motion detection, but when moving, binocular disparity contributed more. In the second experiment, a target was presented each time in the same scene. Participants discriminated between forward and backward target motions in a 2-alternative forced choice task. There were three main findings. Firstly, binocular disparity had a significant effect on the discrimination bias, but binocular summation does not. Secondly, there was an enormous difference on discrimination bias between stationary and moving trials, but there was no significant difference between the absolute values of bias for forward and backward motion. This showed that the presence of locomotion affects flow parsing much more than reversing the locomotion direction. Thirdly, the mean bias resulted in the direction of the target staying approximately constant with respect to the pillars, potentially revealing the strategy being used.

A PILOT STUDY OF THE IMPLICATIONS OF LONG-TERM DANCE CLASSES IN PEOPLE WITH PARKINSON'S DISEASE [2.29]

Sarah Houshangi-Tabrizi, Katy Ghanai, Rebecca Barnstaple, Hedieh Tehrani, Rachel Bar, Karolina Bearss, & Joseph F. X. DeSouza, York University

Parkinson's Disease (PD) is a neurodegenerative disorder that leads to motor and nonmotor dysfunctions. This is due to the depletion of dopaminergic neurons in the Basal Ganglia, specifically substantia nigra pars compacta region which is associated with movement control. Based on the Movement Disorder Society Unified Parkinson Disease Rating Scale (MDS-UPDRS), after five years of being diagnosed with PD, an annual motor decline in the range of 5.2-8.9 is observed. Previous studies have shown the implications of dance as an improvement on the motor and non-motor PD symptoms to slow the progression of this disorder. This study aims to examine the long-term functional and neuroplasticity changes of the brain in people with Parkinson's Disease (PwPD) through a dance-visualization using music they were about to learn in the dance studio, by comparing the combined participants scans of the first week with the last two scanning sessions. We hypothesized that in a group of PD, dance visualization of the learnt dance choreography through listening to music will show activation in the auditory and insular cortices, motor areas, and the visual cortex. We also hypothesized that as PwPD continue learning dance choreography, the activation will increase. To conduct this study, PwPD were recruited (n=10) between the ages of 52-83 over 34 weeks of dance class. Each participant was scanned under the fMRI at four different time points during the 34-week period and while in the scanner, they were assigned to a dance- visualization task cued by music, followed by no music. For the first week, all ten participants were scanned. For the last sessions, two participants were scanned on the third timepoint and four participants were scanned for the last timepoint for a total of six participant for the last two scanning sessions. The functional data of the fMRI scans were analyzed using General Linear Model: Multi-subject in BrainVoyager QX to investigate dance-visualization compared to no music. The fixation (no music) was used to change the blocks of imagery five times for a total of 7.5 minutes. The results of the study demonstrated activation in the network of brain regions (motor areas, auditory and insular cortices) during the first week, and an increase activation of the same areas during the last weeks. Interestingly, there was no activation observed in the visual cortex region in both the first week scans and the last week scans. Both hypotheses were supported apart from no activation in the visual cortex. This study would provide the foundation to learn what is involved in real world movement-visualization in the body of music. Moreover, it provides an insight to the benefits associated with dance classes as a therapeutic measure and how it helps to improve the motor symptoms and other entities associated with PD. Future research can investigate the activation area of different brain region while physically dancing to see all the area associated with physically dancing.

DOGS RELY ON VISUAL CUES RATHER THAN ON EFFECTOR-SPECIFIC MOVEMENT REPRESENTATIONS TO ANTICIPATE HUMAN ACTION TARGETS [1.19]

Ludwig Huber, Lucrezia Lonardo, Christoph J. Völter, & Claus Lamm, Messerli Research Institute, University of Veterinary Medicine Vienna

The ability to predict others' action targets is one of the pillars of social cognition. We investigated the processes underlying this ability by pitting motor representations of the observed movements against visual familiarity. In a pre- registered experiment, we tested an unusual model species: dogs (Canis familiaris). Indeed, dogs cannot perform and hence likely lack representations of certain human movements but at the same time they gain extensive visual experience with such human movements in their everyday life. Using an eye-tracker, we measured the gazing of 16 dogs who observed a human

or a conspecific execute the same goal-directed actions, such as approaching a ball and pushing it away. Considering the first trial, when the human agent performed movements outside dogs' specific movement repertoire (such as bipedal walking and kicking the ball with her foot), dogs' gaze arrived at the target object predictively. When the agent was a conspecific, dogs' gaze shifts to the object were reactive. When the human agent behaved more closely to the dogs' motor possibilities (crawling instead of walking and pushing the ball away using her extended arm), the dogs' gaze arrival times were intermediate between the other two conditions. On average, dogs looked sooner at the target object of the human compared to the dog agent. In a second version of the experiment, showing the same actors approaching and lifting a toy elephant, dogs' gaze arrival times to the object did not significantly differ across the three conditions. When we explored the data by considering also looks at the object that were not preceded by looks at the agents, on the first trial, dogs gazed predictively at the object irrespective of condition, hence irrespective of the species and type of movements performed by the agents. On average, dogs looked predictively and sooner at the target object when the agent was human than when the agent was a conspecific. Further controls are necessary to interpret the influence of the agent, the action and the object on the dogs' looking behavior.

PHOTORECEPTOR BIOPHYSICS ENABLES DEEP LEARNING MODELS TO GENERALIZE ACROSS LIGHT LEVELS [1.26]

Saad Idrees, Greg D. Field, Fred Rieke, & Joel Zylberberg, York University

Under controlled experimental conditions, state-of-the-art deep learning models can predict retinal ganglion cell (RGC) responses to visual scenes well. However, challenges exist for these models in highly dynamic conditions of natural vision. For example, walking into a shaded area on a sunny day can instantly change the intensity of light reaching the retina by several orders of magnitude. These changes in light level dramatically alter RGC responses. Because current deep learning models of retina have no built-in notion of light level, they are unable to accurately predict RGC responses to light levels that they were not trained on. Our goal is to build retina models that (1) perform better than existing state-of-the-art retina predictors at a single light level, and (2) predict RGC responses to visual scenes at different light levels without training the model at each light level. To achieve this, we developed a deep learning model of retina that combines a biophysical front end capturing the phototransduction cascade with convolutional neural networks (CNNs) capturing the inner retina processing. In addition, we developed an adaptive-convolution layer, which dynamically adjusts the gain based on recent input history. The front end photoreceptor-layer with fast and slow adaptational mechanisms provided models with an in-built notion of light intensities by converting stimulus light intensities (rod isomerizations R*rod-1sec-1) into photocurrents. Subsequent CNNs converted the photocurrents into RGC spiking. The adaptive-conv layer that we developed as part of this project further equipped the model to adapt to input dynamics. Here, the layer input was convolved with two temporal kernels which were combined divisively to form the output. The resulting model with a photoreceptor-layer, adaptive-conv layer, conventional convolution layers and a dense output layer, when trained to predict rat RGC responses to stimuli at bright light levels (10,000 R*rod-1sec-1), explained 91% variance in RGC responses. This is a significant improvement over previous state-of-the-art retina predictors based on McIntosh et al., 2016, that could only explain 79% of the variance in RGC responses to the same stimuli. To evaluate if models with photoreceptor-layer trained at bright light level could also predict retinal activity at a darker light level (1 R*rod-1sec-1), we replaced the photoreceptor-layer parameters representing cone dynamics with values representing rod dynamics. The CNN parameters were left unchanged from training at bright light level. This model could explain 54% of the variance in RGC responses. This generalization is exceptional given that this model was never trained on dark light level data. In contrast, retina models based solely on CNNs trained at bright light level failed to predict

responses to stimuli at dark light levels. Teaching artificial neural networks to adapt to dynamics of natural world is a first step in enabling models to generalize across conditions they were not trained on. This is significant for several applications, including visual prosthetics. With a trainable photoreceptor layer and the adaptive-conv layer, our deep learning model is already a better predictor of retinal activity than previous state-of-the-art deep learning models of the retina.

MULTI CAMERA MULTI PERSON TRACKING AND RE IDENTIFICATION IN MOBILE ROBOTS [3.23]

Nizwa Javed & James H. Elder, York University

Intelligent robots can be used to support people in daily tasks like guiding, guarding and elderly care. For a robot to perform effectively, the ability to understand and perceive people in an environment is a key element. To follow a person, robots have to be able to robustly detect and track the position of the target person. While following a person, it often happens that the robot cannot keep track of the person if he/she moves out of the robot's field of view, or is occluded by other people. In such cases, to resume the tracking and following, the robot needs to re-identify the person using a target person model before losing the track This work addresses the problem of detection, tracking and re-identification of people from an ego-centric perspective of a mobile robot equipped with multiple RGBD cameras. In our proposed framework, we first detect people using state of the art object detector YOLO in RGB images from four intel realsense RGBD cameras mounted on top of a mobile robot. Then, the bounding box detections are projected into a robot-centric frame for tracking. DeepSORT algorithm is used for tracking and reidentification. If the robot loses track of the target person, it re-identifies the target person among surrounding people with the help of a CNN based Re-ID model. The entire system is designed in Python and Robot Operating Software (ROS).

IS THE SELF-MOTION-INDUCED BIAS IN TIME-TO-CONTACT ESTIMATION ATTENUATED WHEN OBJECT MOTION IS CONSISTENT WITH GRAVITY? [2.5]

Björn Jörges & Laurence R. Harris, York University

We recently showed that visually-simulated self-motion can bias the perceived speed of an object in the observer's environment [1]. Visual self-motion involves a multisensory conflict between visual and the other vestibular, efferent copy and proprioceptive cues that would normally accompany self-motion. Humans can use an internal representation of Earth gravity to time interceptive actions more accurately [2]. If time-to-contact is estimated partially based on a model of gravity, then biases in perceived speed induced by self-motion should affect the estimated time-to-contact less when the motion is consistent with Earth gravity than when it is not. We tested this hypothesis by immersing participants in a virtual 3D environment and showing them parabolic trajectories in the fronto-parallel plane that was consistent with Earth gravity or with a pseudo-gravitational pull upwards. The target disappeared after reaching the peak of its parabolic path and participants indicated when and where it would have hit a target plank. While they were doing this, participants experienced visually simulated self-motion in one of the four cardinal directions or remained static. Evidence both for the use of an internal representation of Earth gravity and for an effect of visual self-motion on motion extrapolation was inconsistent across the different experimental conditions. Differences between this experiment and the published literature, which may explain these discrepancies, are discussed.

[1] Jörges, B., & Harris, L. R. (2021). Object speed perception during lateral visual self-motion. Attention, Perception, & Psychophysics.
[2] Zago, M., McIntyre, J., Senot, P., & Lacquaniti, F. (2008). Internal models and prediction of visual gravitational motion. Vision Research, 48(14), 1532-1538.

MULTISCALE VIDEO TRANSFORMER [3.25]

Rezaul Karim, Mennatullah Siam, & Richard P. Wildes, York University

Multiscale transformers have been explored in a wide variety of vision tasks. To date, however, the multiscale processing always has been confined to the encoder or decoder alone. We present a unified multiscale encoder-decoder transformer that is focused on dense prediction tasks in videos. 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. We showcase our Multiscale Encoder- Decoder Video Transformer (MED-VT) on Automatic Video Object Segmentation (AVOS), where we outperform state-of-the-art approaches on multiple benchmarks using only raw images and without using explicit optical flow.

ANALYZING AND IMPROVING SHAPE SENSITIVITY OF DEEP CONVOLUTIONAL NEURAL NETWORKS [3.30]

Shaiyan Keshvari, Xingye Fan, & James H. Elder, York University

Recent research suggests that Deep Convolutional Neural Networks (DCNNs) rely more on local textures rather than global object shape when doing object recognition. Human visual perception, on the other hand, relies heavily on shape information. Therefore, a pertinent question is whether we can train DCNNs to better capture shape information. First, we test for existing shape utilization in ImageNet pre-trained DCNNs. Removing global shape information by block-scrambling greatly reduces accuracy, but only if the blocks are small enough to disrupt local texture, suggesting a minimal role for long range shape information. To increase the use of shape information, we propose two fine-tuning methods. One is to fine-tune with hybrid images such that global shape disambiguates local texture information. Hybrid images combine two natural images from the ImageNet dataset: one is coherent, remaining intact while the other is an incoherent image distorted by block scrambling. The two images are combined in alternating blocks into a "checkerboard". This checkerboard technique makes knowing global shape critical to utilize local texture. The second method modifies the fine-tuning loss function to penalize correlation between network outputs to a natural image and its block-scrambled counterpart. We find improvement in increased shape bias, despite mild losses in overall accuracy. The results show that fine-tuning enables the model to differentiate between intact and scrambled images, and that shape utilization is improved. We discuss implications for robustness to occlusion and adversarial attack.

DETECTABILITY OF IMAGE TRANSFORMATIONS DURING EYE AND HEAD MOVEMENTS [3.20]

Maryam Keyvanara & Robert S. Allison, York University

Users of a VR make frequent gaze shifts and head movements to explore their surrounding environment. In addition to self generated head movements, the virtual camera in a VR may be rotated for purposes such as redirection or repositioning of the users. In a VR study, we compared different extents of saccadic suppression of image transformations, during the (a) saccades that users made during active head rotation, (b) saccades made during a simulated rotation of the virtual camera, and (c) saccades were made with a static head. Our subjects wore a head-mounted display, and reported the directions of a sudden additional scene rotation that occurred either during their saccade as a target object in the 3D scene was displaced or just before the target displacement while the target was fixated. Consistent with our previous studies, we found a stronger suppression during saccades that were made during head rotations (whether real or simulated), compared to when the head was static. In addition, our results showed a stronger suppression of visual changes during saccades, whether accompanied by head rotations or not, than during fixations of the eyes. The results of this study add valuable information for the design of more efficient VR environments, and can be used in scenarios where designers want to introduce a change of viewpoint without the user noticing it.

COGNITIVE FACTORS MAY AFFECT THE UPDATING OF AN OBJECT'S POSITION DURING LINEAR LATERAL TRANSLATION [1.14]

John Jong-Jin Kim & Laurence R. Harris, York University

Updating the egocentric positions of surrounding objects during self-motion is fundamental to the daily activities of people interacting and/or avoiding them as they navigate in the world. Despite the importance of this ability, past studies have shown that people make systematic errors when updating during linear self-motions. To update the positions of surrounding objects, a person needs to know their own movement through space, which requires integrating information from multiple senses including visual, vestibular, somatosensory and motor systems. In Exp1, we measured people's updating errors after passive lateral translation (1m) with visual motion cues only (using an Oculus VR HMD while sitting on a stationary chair in an office-like room). Remembered target positions were shifted in the direction of movement. In Exp2, we translated people again (.46m) this time adding physical motion with and without visual cues (using a 6-DOF moving platform in a test room identical to the VR environment) and also visual cues only. However, the updating errors were the same as when they were stationary, when the person moved (either visually, unlike the finding from Exp1, or physically). Were the differences between Exp1 and 2 due to the difference in the travel distance, or the effect of exposure to a moving platform? Therefore, in the present study (Exp3), we looked for updating errors associated with the two visually simulated travel distances (.46m and 1m). With N=9, a main effect of translation was found where their remembered positions shifted in the direction of the movement as found in Exp1 but not in Exp2. This suggests that cognitive factors (sitting on a platform that is known to be able to move) may impact people's updating, making them more accurate at updating the positions of targets that were initially seen before they moved.

VISUAL HALLUCINATIONS CAUSED BY CHARLES BONNET SYNDROME ARE NOT IMPACTED BY THE PSYCHOSOCIAL FACTORS RESULTING FROM COVID-19 RESTRICTIONS [2.4]

Aysha Kinakool, Stefania S. Moro, & Jennifer K. E. Steeves, York University

Charles Bonnet Syndrome (CBS) is a poorly understood, debilitating phenomenon following vision loss from common diseases such as age-related macular degeneration and currently has no treatment. Individuals with CBS often experience visual hallucinations in the form of images such as people, landscapes, or patterns that do not exist. This experience is the result of an imbalance between the loss of vision and the active visual brain but is often misdiagnosed as dementia or a mental health condition. Psychosocial factors have been suggested to exacerbate CBS visual hallucinations. The ongoing COVID-19 pandemic has resulted in the implementation of public health safety measures including extended periods of total or partial lockdowns to reduce virus transmission. To date, Canadians have undergone several waves of restrictions that may increase the risk of pervasive psychosocial issues from social isolation. The goal of this research is to assess how the social constraints of the current COVID-19 pandemic may affect CBS visual hallucinations. We surveyed 39 individuals (M: 69 years, SD: 16 years) with CBS from the CNIB Foundation. Visual hallucinations, anxiety, social isolation, loneliness, and QoL were assessed using: The Specific Psychotic Experiences Questionnaire and modified versions of the Generalized Anxiety Disorder 7 scale, Steptoe Social Isolation Index, DeJong Gierveld Loneliness scale, and the World Health Organization QoL scale, respectively. 90.6% of patients reported no change (increase: 6.3%; decrease: 3.1%) in average duration and 62.2% reported no change (increase: 21.6%; decrease: 16.2%) in frequency of CBS hallucinations. No significant differences were observed in anxiety ($p= 0.514$); QoL ($p= 0.155$); social isolation ($p= 0.835$); and loneliness ($p= 0.296$) between participants who reported experiencing a change compared no change in hallucinations. The social constraints of the current COVID-19 pandemic measured through anxiety, social isolation, loneliness, and QoL do not affect visual hallucinations caused by CBS. This is consistent

with the notion that CBS hallucinations are associated with an active visual cortex following vision loss.

SIMULATED MOTION IN VIRTUAL ENVIRONMENTS AFFECTS COGNITIVE TASK PERFORMANCE [3.24]

Onoise Gerald Kio & Robert S. Allison, York University

Compelling dynamic visual content in virtual environments can induce the sensation of self-motion (also known as vection) in stationary observers. It has been shown that vection significantly reduces in magnitude when people perform attentionally-demanding visual tasks while simultaneously observing fast-moving virtual reality (VR) content. Is it also possible that fast-moving VR content also affects people's performance of attentionally demanding tasks? In this study we investigated the effects of simulated motion on the performance of attentionally-demanding tasks in a virtual environment. We compared the intensity of vection that stationary observers experienced while performing visual and aural discriminatory tasks with the vection they experienced while they had no attentionally-demanding task. We found that vection was significantly reduced only when they performed the visual tasks. We also found that while the speed of simulated motion did not significantly affect accuracies of both tasks, it significantly affected their response times to the aural task but not the visual task. These suggest that simulated motion in virtual environments causes attentional loading that significantly affects cognitive task performance.

A SCOPING REVIEW OF THE FEAR OF FALLING RESEARCH: CONCEPTUAL FRAMEWORK AND EMPIRICAL GAPS [1.15]

Kamila Kolpashnikova & Shital Desai, York University

Fear of falling (FoF) can profoundly affect the quality of life of older adults. Despite extensive evidence of medical, public health, and sociological literature on individual-level factors and behavioural interventions in assessing and mitigating FoF, very little attention has been devoted to environmental and socio-economic factors to study FoF and its effects on everyday lives of seniors. This scoping review aims to map the conceptual framework and identify empirical gaps in the literature to guide future work on FoF. Using a scoping study framework and PRISMA-ScR guidelines, seven databases (Medline-Ovid, Embase-Ovid, PsycInfo-Ovid, EBM Reviews - Cochrane Database of Systematic Reviews-Ovid, CINAHL, Web of Science, and Scopus) were searched with a relevant keyword in the title, "FoF," and subject headings and keywords relating to the older population. After de-duplication and exclusion criteria checks, 523 articles were analyzed. Using topic modelling with the help of the Non-negative Matrix Factorization (an NLP machine learning technique that creates a summary 'topics' for a large number of text documents) on articles' abstracts, we identify the range of the topics covered in the research of FoF, as well as the areas where more empirical effort would be needed. The results suggest that the topic areas with the largest research focus in literature are exercise and training interventions, balance and FoF, and gait and FoF. The results of the scoping review have implications for future research that focusses on user friendly environments for older adults.

LONGER-TERM HOCKEY PLAYER TRACKING WITH SPLIT-AND-RECONNECT [3.19]

Maria Koshkina & James H. Elder, York University

Computer vision is becoming increasingly important for automatic sports understanding. Long-term player tracking is a key sub- problem, critical for many downstream analysis tasks. However, research on multi-object tracking (MOT) has to date primarily focused on pedestrian and vehicle tracking in urban scenes due in part to the existence of large public datasets such as the MOT Benchmark. Player tracking is a special case of MOT and presents its own unique challenges such as fast motion, frequent

occlusions, and similarity of appearance for players on the same team. Here we introduce a novel hockey player tracking dataset, use it to evaluate state-of-the- art MOT trackers, and propose a post-processing approach that enables longer-term tracking of players. Our evaluation of current state-of-the-art MOT trackers on our hockey player dataset shows that they are able to perform short-term tracking fairly well but fail to maintain player identities through occlusions. To address this limitation, we propose a split-and-reconnect post-processing method, where the initial output of the tracker is analyzed and broken down into smaller, more reliable tracklets. Tracklets are then re-connected into longer tracks using appearance and motion models. We show that this approach improves association accuracy in player tracking. An analysis of the potential upper performance bound of this approach indicates that further improvement to our splitting and reconnection algorithm could lead to further performance gains.

A DEEPER DIVE INTO WHAT DEEP SPATIOTEMPORAL NETWORKS ENCODE: QUANTIFYING STATIC VS. DYNAMIC INFORMATION [3.21]

Matthew Kowal, Mennatullah Siam, Amirul Islam, Neil D. B. Bruce, Richard P. Wildes, & Konstantinos G. Derpanis, York University

Few-shot video understanding aims at conducting spatiotemporal localisation through segmenting actions/actors within a video using few labelled trimmed videos. Towards this goal, we study two main questions: (i) how to enforce temporal consistency of the predictions within a video? and (ii) how to quantify the models ability to capture dynamic information, which is a crucial component when looking into generalization to novel unseen actors/actions. In the first question, we focus on the simpler task of few-shot video object segmentation (FS-VOS) which segments video frames using a few labelled examples of classes not seen during the initial training. We present a simple but effective temporal transductive inference (TTI) approach that leverages temporal consistency in the unlabelled video frames during few-shot inference. Key to our approach is the use of both global and local temporal constraints. The objective of the global constraint is to learn consistent linear classifiers for novel classes across the image sequence, whereas the local constraint enforces the proportion of foreground/background regions in each frame to be coherent across a local temporal window. These constraints act as spatiotemporal regularizers during the transductive inference to increase temporal coherence and reduce overfitting on the few-shot support set. Empirically, our model outperforms state-of-the-art meta-learning approaches in terms of mean intersection over union on YouTube-VIS by 2.8%. In addition, we introduce improved benchmarks that are exhaustively labelled (i.e. all object occurrences are labelled, unlike the currently available), and present a more realistic evaluation paradigm that targets data distribution shift between training and testing sets. In order to investigate the second question, we propose a quantifiable interpretability technique that studies different spatiotemporal models. Currently, there is a limited understanding of what information is captured by these models in their intermediate representations. It has been observed that certain spatiotemporal models are heavily influenced by visual appearance in single static frames. However, there is no quantitative methodology for evaluating such static bias in the latent representation compared to bias toward dynamic information (e.g. motion). We tackle this challenge by proposing a novel approach for quantifying the static and dynamic biases of any spatiotemporal model. To show the efficacy of our approach, we analyse two widely studied tasks, action recognition and video object segmentation. Our study demonstrates the effects of datasets and architectural decisions on these biases. Our study serves as an analysis tool to help us gain insights into how to improve the generalization capability of spatiotemporal models beyond the training classes.

MINIMIZING PEDESTRIAN FALLS ON OUTDOOR WALKWAYS: MINIMUM FOOT CLEARANCE ESTIMATION (MFCE) SYSTEM [1.4]

Anchana Kuganesan, Hamed Ghomashchi, & Tilak Dutta, KITE Research Institute

Falls are the most common cause of injury in seniors with one in three Canadian adults aged 65 years and older experiencing a fall at least once annually. Many outdoor fall-related injuries on sidewalks are caused by tripping that occurs when an individual fails to adjust their gait while negotiating obstacles and raised surfaces. An important indicator of the risk of tripping is an individual's minimum foot clearance (MFC). MFC refers to the instant in the swing phase of the gait cycle when the foot is closest to ground. As an individual gets older, their mean MFC value also decreases putting them at increased risk for falls. Thirteen percent of adults in the risk population have MFC values below 6 mm yet many jurisdictions allow level changes up to 13 mm. The existing literature on MFC values are based on measurements done in the laboratory environment, which is known to result in overestimates of 6.5-16.2% compared to real-world settings. In order to address the need for real-world MFC measurements, new and more efficient approaches for measuring MFC are needed. Our team has recently developed the Minimum Foot Clearance Estimation (MFCE) system. The data collection module is designed to be positioned at ground level next to a walkway to efficiently collect sagittal plane videos of many pedestrians' feet and lower legs. It consists of a pair of synchronized, calibrated video cameras (Z CAM E2, Shenzhen, China) that can be used to estimate the 3D coordinates of objects within their view. The videos recorded by the cameras are used to track the three-dimensional foot trajectory of the pedestrian and extract the MFC value. The video data collected is then processed offline using a computer vision algorithm for automatically estimating MFC values. The objective of the current project is to use the MFCE system to estimate the distribution of MFC values for the population by estimating, collecting, and analyzing foot clearance estimates from pedestrians on outdoor walkways. Data collection locations will include walkways that are level, sloped, and locations with existing level changes of differing sizes. The resulting estimated MFC values from this analysis will be plotted in a set of histograms defining the MFC distribution for pedestrians that will be used to determine the extent to which individuals adjust their foot clearances in each of the conditions evaluated. This foot clearance distribution data will contribute to developing evidence-based guidelines for outdoor walkway design and maintenance and inform other strategies for minimizing the risk of trip-related falls on outdoor walkways.

ROLE OF THE VENTROMEDIAL PREFRONTAL CORTEX IN MNEMONIC DISCRIMINATION [2.7]

Claire Lauzon, Daniel Chiasso, Stevenson Baker, Elisa Ciaramelli, & R. Shayna Rosenbaum, York University

The ventromedial prefrontal cortex (vmPFC) is involved in strategic aspects of memory encoding and retrieval, which may be essential for accurate discrimination of similar memory representations. The vmPFC is also thought to mediate a "feeling of rightness" within familiarity, which influences the ability to judge the appropriateness or veracity of visual stimuli. Individuals with lesions to the vmPFC have previously demonstrated inflated confidence in erroneous memories, possibly due to faulty meta-mnemonic monitoring after injury to this region. While mnemonic discrimination tasks have been traditionally conceptualized as behavioural estimations of hippocampal pattern separation, the known role of the vmPFC in episodic memory might also influence mnemonic discrimination. We examined if mnemonic discrimination critically depends on the vmPFC by testing individuals with selective vmPFC lesions. Participants were administered the Mnemonic Similarity Task (MST; Stark et al., 2015), which assesses the ability to distinguish previously learned images of everyday objects from novel, highly similar images (lures) and dissimilar images (foils). In a modification of the original MST, we also examined response confidence to better understand whether the role of the vmPFC in mnemonic

discrimination is related to disproportionately high confidence. Relative to controls, individuals with vmPFC lesions were selectively impaired in discriminating studied items from similar lures but not foils, resembling findings in individuals with hippocampal compromise. Analysis of confidence ratings additionally suggests that vmPFC patients are overly confident relative to their accuracy in response to similar lures, but not in response to targets or foils. Thus, informative confidence signals and "feeling of rightness" as subserved by the vmPFC may be uniquely necessary for mnemonic discrimination of similar, overlapping information, but not for recognition of discrete stimuli. These findings suggest that mnemonic discrimination may rely on proper functioning of the vmPFC, perhaps specifically because of the region's contributions to metamnemonic monitoring.

MODIFIED OPTIC FLOW GAIN DURING QUIET STANCE [1.18]

Lisa Lavalle & Taylor Cleworth, York University

In healthy individuals, balance is maintained through sensorimotor integration combining visual, vestibular, and proprioceptive feedback to maintain upright stance. When visual feedback is reduced, such as by presenting participants with 4 Hz stroboscopic illumination, postural instability increase. In contrast, previous studies that have magnified visual biofeedback have observed a tighter regulation of center of pressure (COP) within the base of support. However, there is currently limited work examining how postural control is affected by manipulating optic flow directly. Therefore, by using virtual reality (VR) to apply gain factors to amplify and reduce optic flow, this study aimed to better understand how the gain of optic flow influences balance control during quiet, upright stance among healthy adults. 30healthy adults (mean age: 22.9 years, 21 female) stood quietly either directly on a force plate or on foam. Each surface condition included a set of four experimental and two baseline trials, each 60-seconds in duration. During each experimental trial, participants wore a VR head mounted display (Oculus Rift), exposing them to a full field of view (approx. 110°) virtual environment that replicated a real-world setting. Vizard python programming (WorldViz) was used to apply gain factors (0.25, 2, 4, and 16) in a counterbalanced order to manipulate optic flow relative to head motion. Head kinematics (HeadPos) and COP amplitude, velocity, and frequency were used to quantify balance, in addition to mean power across four frequency bands (LOW: 0-0.1 Hz, MED: 0.1-0.5 Hz, MED-HIGH: 0.5-1.0 Hz, and HIGH: 1.0-5.0 Hz). While standing on foam, COP and HeadPos decreased in amplitude as optic flow gain increased. COP mVel increased with optic flow gain, while HeadPos velocity decreased with gain on foam. Frequency analysis showed an increase in COP frequency with an increase in optic flow gain while standing on foam. Additionally, across both surface conditions, COP and HeadPos mean power decreased with optic flow gain in the MED frequency band. In both the MED and MED-HIGH frequency bands, while on foam COP and HeadPos mean power were greatest at a gain of 16 compared to all other gain conditions. When there was an increased reliance placed on the visual system, such as by standing on foam, postural stability appeared to be significantly influenced by optic flow gain. Specifically, an increase in optic flow gain contributed to a tighter control of upright stance. There was also evidence to suggest that visual contributions to balance control may extend to frequencies above 0.1 Hz. Overall, by understanding how participants utilize visual feedback for postural control, we can eventually apply this visuomotor mechanism to study how optic flow gain plays a role in balance deficits and fall risk.

THE INFLUENCE OF NAVIGATIONAL EXPERIENCE ON BOUNDARY EFFECTS IN SPATIAL MEMORY [3.11]

Xuehui Lei & R. Shayna Rosenbaum, York University

Previous studies have shown that spatial memories for complex spaces can be fragmented and distorted by physical boundaries. The current study examined if navigational experience would

modulate the effect of physical boundaries on spatial memories in a complex building. One floor of a building at York University was simulated in a desktop virtual environment. Two groups of participants were tested. One group of inexperienced participants had no prior experience with the building and the other group of experienced participants had prior real-world experience and were familiar with the building. Participants learned the locations of objects in the building, with some objects inside rooms and other objects in open areas. After learning, the objects were removed. Participants completed an onsite pointing task in which they appeared at an object's location and were asked to point to the direction of another target object. The pointing could be across boundaries (i.e., across rooms, between a room and an open area) or within boundaries (i.e., within a room, within open areas). Participants then completed a sketch map task outside the virtual environment, where they created a map of the objects with the exterior walls of the building and the objects provided. The preliminary results showed that both inexperienced and experienced groups exhibited larger absolute pointing errors when pointing across boundaries than within boundaries, although, overall, the experienced group was more accurate than the inexperienced group. The interaction between the groups and the boundary conditions was not significant. Sketch mapping of the objects from the experienced group was closer to the correct object configuration than that from the inexperienced group. These results demonstrated that prior real-world navigational experience facilitated spatial learning of new objects in a complex indoor space, but the effect of physical boundaries on fragmenting spatial memories still existed regardless of familiarity with the space.

CONTRIBUTION OF INDIVIDUAL DIFFERENCES IN IMAGERY TO EPISODIC MEMORY AND SPATIAL NAVIGATION [1.1]

Adrienne Li, Maria Arrieta, & R. Shayna Rosenbaum, York University

Individual differences in mental imagery, specifically aphantasia (APH) - the absence of voluntary visual imagery, have attracted growing interest, but research into its associations with other cognitive domains, such as long-term memory and spatial navigation, remains preliminary. We present data comparing APH individuals to controls (HC) on self-rated measures of spatial navigation, strategy and mood; additionally, we investigate the contribution of imagery ability to performance on qualitative episodic spatial memory tasks. APH participants (n=113) and matched HC (n=110) completed questionnaires on mental imagery, spatial memory, anxiety and navigational strategy. A sub-group of APH (n=65) and HC (n=72) completed the Route Description task (Herdman et al., 2015) assessing episodic and gist-like spatial memory for familiar and unfamiliar routes. We predict that APH will differ significantly from controls on imagery (self-rated and vividness questionnaire) but not spatial navigation, strategy, anxiety, nor on qualitative measures of spatial navigation (i.e., route description). We predict that overall number of details will not differ between APH and HC, but APH will report fewer visual details and lower vividness for both highly familiar (frequently travelled) routes and less familiar (infrequently travelled) routes. Consistent with previous findings, APH do not differ from HC on self-report measures of spatial navigation, but differed on their self-rated ability to sketch a map of routes. APH had lower vividness ratings than HC of familiar and unfamiliar routes and provided fewer sensory details for unfamiliar but not familiar routes on the RD task. Unfamiliar routes are likely to be more "episodic" in nature (experienced only once or rarely). Similar levels of sensory details for familiar routes in APH and HC may point to alternative strategies used by APH to compensate for poor imagery and episodic memory that develop with extensive exposure (e.g., semanticization). Overall, the results in APH suggest that although some aspects of spatial memory and navigation are intact, imagery of sensory details that enables re-experiencing of an environment is compromised. Findings provide further evidence that there are dissociable aspects of spatial navigation supported by regions outside of the hippocampus.

INFLUENCE OF AN ALLOCENTRIC CUE SHIFT ON REACH ACCURACY IN MONKEYS [1.13]

Jennifer Lin, Hongying Wang, Saihong Sun, Xiaogang Yan, & J. Douglas Crawford, York University

Various sources of information, such as object locations, are taken in and further processed within a specific frame of reference, ultimately determining a mental spatial representation of the environment. It has been shown that humans optimally weigh egocentric and allocentric (landmark) cues when pointing (Byrne & 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, behavioural data is collected from one female Macaca mulatta monkey (ML) who is currently being trained to complete a memory guided reaching task. This task requires the monkey to touch 1 of 3 varying locations of a waist level LED bar. The animal reaches and touches a target appearing at 1 of 15 locations on the touch screen after a delay. There are 3 conditions: no landmark, stable landmark and landmark shift. In the stable landmark condition, we found that correlations with target variance for target relative to landmark cue were initially low but increased dramatically to about 0.82 for catch trials and this shows that the animal is reaching towards the target and not the landmark. In the shift condition, the allocentric weight is about 0.2295. The allocentric weights were higher for targets located closer to the upper left corner. This suggests that the monkey is influenced by visual landmarks when reaching to remembered target in a similar way as humans.

NOISE2NOISEFLOW: REALISTIC CAMERA NOISE MODELING WITHOUT CLEAN IMAGES [3.14]

Ali Maleky, Shayan Kousha, Michael S. Brown, & Marcus A. Brubaker, York University

Image noise modeling is a long-standing problem with many applications in computer vision. Early attempts that propose simple models, such as signal-independent additive white Gaussian noise or the heteroscedastic Gaussian noise model (a.k.a., camera noise level function) are not sufficient to learn the complex behavior of the camera sensor noise. Recently, more complex learning-based models have been proposed that yield better results in noise synthesis and downstream tasks, such as denoising. However, their dependence on supervised data (i.e., paired clean images) is a limiting factor given the challenges in producing ground-truth images. This paper proposes a framework for training a noise model and a denoiser simultaneously while relying only on pairs of noisy images rather than noisy/clean paired image data. We apply this framework to the training of the Noise Flow architecture. The noise synthesis and density estimation results show that our framework outperforms previous signal-processing-based noise models and is on par with its supervised counterpart. The trained denoiser is also shown to significantly improve upon both supervised and weakly supervised baseline denoising approaches. The results indicate that the joint training of a denoiser and a noise model yields significant improvements in the denoiser.

INVESTIGATING FACIAL EMOTION RECOGNITION IN PARKINSON'S DISEASE USING COVID-19 FACE MASKS [2.8]

Geneva Mariotti, Judith Bek, & Joseph F. X. DeSouza

A deficit in facial emotion recognition (FER) accuracy has been reported in people with Parkinson's disease (PwPD). This observed deficit has been theorized to be either due to differences in face-related visual scanning strategies in PwPD or to an impairment in facial emotion mimicry in PwPD. Further, preliminary investigations during the COVID-19 pandemic have revealed impairments in FER accuracy within the general population when observed faces are wearing COVID-19 style masks which obscure part of the face. The pandemic presents an ecological opportunity to investigate FER impairment in PwPD with the use of COVID-19 face masks to limit the facial information presented. This study aimed to advance our understanding of the FER deficit in PwPD, while accounting for alternative influential variables reported in previous studies such as depression, disease progression, and PD medications. We

hypothesized that both PwPD and healthy controls (HC) would have reduced FER accuracy when faces were wearing COVID-19 masks compared to when the faces are unaltered, and that this impairment would be larger for PwPD than for HC. To investigate this, individuals aged 65+ (PwPD = 8, HC = 14) were actively recruited from 12 major organizations associated with PwPD and ageing, to complete a 25-minute online-quasi-experimental paradigm. Participants were presented with photos from the FACES database, either with a COVID-19 mask (obscured) or without (unobscured). Participants were asked to identify the gender of each face (control task) or the emotion expressed (experimental task) through a multiple choice format. Participants also completed the Geriatric Depression Scale - Short (GDS-SF) and PwPD completed the Parkinson's Disease Questionnaire-8 (PDQ-8) and reported PD medications. Results revealed no group differences between PwPD and HC's FER accuracy for the identification of emotion in masked and unmasked faces nor in individual emotions. A 2 (Group) X 2 (Mask) mixed methods ANOVA revealed a significant main effect of masks ($F(1, 20) = 34.37$, $p < .001$) suggesting that the use of face masks made FER significantly less accurate ($M = 55$, $SD = 9.51$) compared to unobscured faces ($M = 79.3$, $SD = 9.79$). Further, a 6 (Emotion) X 2 (Group) mixed methods ANOVA revealed a significant main effect of emotion ($F(5, 100) = 21.11$, $p < .001$) in FER accuracy, with some negative emotions (anger, disgust, fear, sadness) being more difficult to identify than some positive emotions. While part of our hypothesis surrounding the mask's role in impairing FER accuracy was supported, no group differences were found between PwPD and HC as hypothesized. A possible explanation for our findings is the small sample, which may be insufficient to detect group differences as PwPD scored low in disease-impairment (M PDQ-8 = 12.13, SD = 7.14) and depression (M GDS-SF = 4.62, SD = 3.37) for a clinical PD population. This population represents the healthiest of PwPD, possibly explaining why expected FER deficits were not observed. Future investigations without online components may yield a more representative sample.

CHARTQA: A BENCHMARK FOR QUESTION ANSWERING ABOUT CHARTS WITH VISUAL AND LOGICAL REASONING [3.12]

Ahmed Masry, Do Xuan Long, Jia Qing Tan, Shafiq Joty, & Enamul Hoque, York University

Data visualizations such as bar charts and line charts have become popular in analyzing data and making informed decisions. To analyze data, often people ask complex reasoning questions about charts involving arithmetic and logical operations. Answering such questions requires a significant amount of perceptual and cognitive efforts as people need to combine multiple operations such as retrieving values, comparing values, finding maximum, calculating sums and differences of values (e.g., Which year has the most divergent opinions about Brazil's economy?). The goal of a Chart Question Answering (ChartQA) system is to help users by taking a chart and a natural language question as input and predicting the answer. This task differs from other QA tasks such as QA on texts and tables because the input for ChartQA is a visual representation of data that can draw a reader's attention to various prominent features such as trends and outliers. Also, people tend to ask questions by referring to visual attributes of marks (e.g., What is the peak value of the orange line?) While the task of ChartQA has received growing attention in recent years, existing datasets have several major limitations: (i) the questions are generated automatically using pre-defined templates which lack naturalness, (ii) the charts are created automatically using a programming tool like Matplotlib which do not reflect the diverse styles of many real-world charts, and finally, (iii) in most datasets, the answer comes from a small fixed-sized vocabulary (e.g., chart axis labels, 'yes', 'no'). Consequently, existing models usually treat the task as a classification problem and rely on dynamic encoding techniques with the questions and answers encoded in terms of spatial positions of chart elements (e.g., x-axis-label-1). Such approaches do not work when the OCR model generates errors or when the question refers to chart elements using synonyms (e.g., US vs. United States). PlotQA attempts to support open vocabulary

questions by applying a TableQA model but it does not consider any visual features of a chart which are critical for answering visual reasoning questions. To address these limitations, we present a large-scale benchmark covering 9,608 human-written questions focusing on logical and visual reasoning questions. Since human annotations are costly, we also generated another 23,111 questions automatically from human-written chart summaries, which have rich variations in language, using a T5 model and manually validated a subset of it for quality assurance. Our benchmark consists of 20,882 charts which are curated from four different online sources to ensure variety in visual styles and topics. To address the challenges introduced in our benchmark involving visual and logical reasoning over charts, we present two transformer-based models that combine visual features and the extracted underlying data table of the chart (by adapting the ChartOCR model) in a unified way to answer questions. Our models achieve the state-of-the-art results or stand on par with the previous models on the previous datasets as well as on our newly created benchmark.

WHEN TRAVELING BETWEEN REAL AND VIRTUAL WORLDS THE CURRENT ENVIRONMENT DIRECTS POINTING BEHAVIOUR [2.19]

Meaghan McManus, Franziska Seifert, Immo Schütz, & Katja Fiehler, Justus Liebig University

Sometimes when removing a virtual reality (VR) headset and coming back to the real world (RW) we are surprised to find that we are facing a different direction than expected. When in VR it is easy to lose track of our remembered location in the RW. In the current study we were interested in how well humans can build and maintain spatial representations of virtual and real environments when switching between both. In particular, we examined the participants' ability to point to the remembered locations of targets that were seen in a VR or RW environment while they were currently in the other environment. Additionally, we tested how pointing endpoints would be affected if the VR environment was a virtual representation of the same RW environment (lab, matched) or another environment (kitchen, not-matched). Participants were asked to memorize the positions of 4 target objects in either the RW or VR environment. They then entered the other environment by either putting on or taking off the VR headset, respectively. The targets were not visible in the second environment and participants were asked to point to the remembered location of the targets in the first environment. Importantly, the VR environment could either be aligned to the RW (0°) or unnoticeably misaligned (2.6° pitched up or down). This resulted in four conditions, VR to RW and RW to VR where the VR environment could be matched or not-matched to the RW environment. We hypothesized that in the unmatched conditions participants would be more likely to ignore the second environment and therefore more accurately point to the remembered positions of the targets from the first environment. In the matched conditions participants would be more likely to integrate the two environments resulting in a greater pointing error. A LMM was run to compare the effect of pitch (0°, 2.6°, -2.6°), order (VR-RW and RW-VR), and VR environment (matched vs not-matched) on the average pointing error. There was a significant interaction between order and pitch. When participants switched from VR to RW there was no effect of the previously seen pitch, but when they switched from RW to VR the pointing error was significantly affected by the current pitch of the environment. We then subtracted each participant's pointing error at 0° from their pointing error during pitch and calculated the ratio between pointing error and pitch (average pointing error/pitch). When the switch occurred from RW to the VR the pointing error was 87% of the pitch, and from VR to RW it was 4.7%. Our results show that the remembered locations of the targets are based primarily on the environment where participants perform the pointing task and not the environment where they saw the targets. Participants are likely relying mainly on the allocentric relationship between the targets and the environment to guide their pointing. This would suggest that when moving between different environments participants may be only maintaining information about the allocentric relationships between the objects in the previously seen environment.

ADAPTATION TO PONG BOUNCE PERTURBATIONS IS QUICK AND INDEPENDENT FROM WALL TILT [1.17]

Laura Mikula, Bernard Marius 't Hart, & Denise Y. P. Henriques, York University

The human motor system can adapt to unexpected perturbations during ongoing movements. Except for target jump studies, most of the research focused on adaptation to perturbations applied to the hand, such as force field or visuomotor rotation. But less is known about how we adapt to perturbations affecting objects that we interact with. For instance, do participants take into consideration visual cues in the environment to reduce their errors and correct subsequent motor commands? To investigate this question, we used an online version of the pong game in which participants intercepted a ball using a paddle controlled by their cursor. We manipulated the post-bounce ball trajectory to be congruent or not with the orientation of the bouncing wall. The "trained tilt" group (n = 34) adapted to the congruent condition while the "trained horizontal" group (n = 36) adapted to the incongruent condition. If visual cues are effectively used by participants, the "trained tilt" group should exhibit faster and/or more complete adaptation than the "trained horizontal" group. Our results showed that the perturbation significantly decreased the interception success rate. Both groups showed typical markers of motor adaptation: large initial errors upon perturbation introduction, followed by rapid error reduction and aftereffects (errors in the opposite direction) when the perturbation was removed. However, we did not observe significant differences in interception success rates or errors between the "trained tilt" and "trained horizontal" groups. Our findings suggest that participants quickly adapted to the dynamics of the pong ball although the tilt of the bouncing surface did not contribute to their performance. Furthermore, this experiment provides evidence that adaptation to external perturbations applied to a moving object is possible in online settings. These results encourage further research on motor adaptation using more naturalistic stimuli and gamified tasks, in real-world or virtual reality environments.

THE DENTATE GYRUS SUBREGION OF THE HIPPOCAMPS IS INVOLVED IN PERCEPTUAL DISCRIMINATION: INSIGHTS FROM DENTATE GYRUS- AND CA1-LESION CASE STUDIES [3.18]

Krista Mitchnick & R. Shayna Rosenbaum, York University

Visual perception is typically understood to be subserved by the ventral visual stream (VVS), whereas the hippocampus (HPC) plays an integral role in episodic memory. Nevertheless, some evidence exists to suggest that the specific subregions within the HPC differentially contribute to underlying processes. Specifically, the dentate gyrus (DG) subregion is involved in pattern separation, orthogonally representing similar information to facilitate the formation of separate, fine-resolution memories. Given that the 'pattern separation' function of the DG is believed to relate to sparse encoding facilitated by DG neurogenesis, the DG might also be involved in non-mnemonic, or perceptual, discrimination, as well as general information encoding. On a series of neuropsychological measures requiring fine-grained perceptual discrimination of abstract designs, line lengths, space, orientation, or size, and on measures requiring categorical perception of morphed common objects or emotional expressions, a unique brain-damaged individual, BL, with bilateral selective DG lesions, performed below expectations on all but one measure. Conversely, BR, an individual with a bilateral lesion specific to the CA1 area of the HPC, exhibited no difficulty on any of these tasks. On measures of verbal memory (e.g., stories, list learning), BL's encoding/immediate recall memory was impaired, but his retention of what he did encode was generally intact. By contrast, BR exhibited average encoding but poor retention, presenting more classically amnestic. The findings are unlikely to be explained by intelligence, attention, working memory, language abilities, or visuospatial abilities, as performance in these domains was intact in the two cases. Collectively, these results strongly support an involvement of the DG, but not CA1, in non-mnemonic discrimination and encoding, and CA1, but not DG, in retention and/or retrieval. Our data

additionally indicate deficits in complex visual perception outside the VVS.

EFFECTS OF VISUAL CUES ON ADAPTATION TO INTERNAL AND EXTERNAL ERRORS [1.29]

Shanaathanan Modchalingam, Bernard Marius 't Hart, & Denise Y. P. Henriques, York University

When performing motor tasks, we improve performance by modifying future movements to correct for observed errors. The assigned source of the errors can affect many aspects of adaptation including its generalizability and the updating of internal models. Adaptation to errors assigned to internal sources (e.g., our arm) is often specific to the arm and is poorly generalized when acting with a different effector. Adaptation to errors assigned to external sources (e.g., the environment) on the other hand, is agnostic to the effector being used but may be specific to the environment. Since the cause of an error is often ambiguous, sensory cues can be used to estimate the likely source of the error. We developed a task in which motor errors could be assigned to internal or external sources. Participants made arm movements to roll a ball toward targets in a head-mounted virtual reality environment. We induced errors by either modifying the mapping between the arm movement and the initial movement of the ball, or by applying a change in the ball path only after the release of the ball. Additionally, we used informative visual cues to signal changes in the environment to increase external error attribution. Pilot data show visual cues can facilitate the assignment of errors to external sources. The visual cues can be dynamics of the ball movement or global environmental changes. External error attribution can in turn allow for rapid switching between motor memories.

OBJECTIVE AND SUBJECTIVE IMPACT OF CHROMATIC ABERRATION COMPENSATION ON COMPRESSION ARTIFACTS [3.29]

Sanjida Sharmin Mohona, Domenic Au, Laurie M. Wilcox, & Robert S. Allison, York University

In virtual and augmented reality displays, lenses focus the near-eye display at a far optical distance and produce a large field of view. These lenses exhibit considerable distortion and cause chromatic aberration. As these are typically corrected by pre-processing the image with the opposite distortion before sending it to the display, they are not apparent to the user. To correct the pin-cushion distortion the pre-processing involves pre-warping source images with an inverse pin-cushion (barrel) transform for each colour channel. Most image compression algorithms use a colour space conversion before compression as it normally improves compression performance by reducing the degree of correlation between components. However, as lens pre-distortion processing is colour specific the spatial correlation between colour channels is disrupted. We performed objective analyses which suggest that the colour space conversion may not be beneficial under these conditions. A set of 10 computer-generated stereoscopic high dynamic range images were tested. Images spanned a wide range of content and were designed to challenge the codecs. The pre-processing workflow involved pre-distorting the images with color shifting (to simulate color separation) and /or barrel distortion, compressing with codec, and finally uncompressing and undistorting the image. The main image manipulations were the codec used and whether the colour transform was bypassed (bypass-on) or not (bypass-off). Images were compressed at the codec's respective nominal production level. We found that with normal color processing, codec performance degrades with any color shift tested. With color transform bypass on, performance appears to be relatively independent of color shift. Normal color processing outperforms color transform bypass with no shift but is worse with any shift. It was found that degradation due to color shifting is generally larger than for the more realistic Brown distortion model as corners of a radial distorted image have chromatic aberration, but the center does not. Overall, it is advisable not to perform perceptual color transformation before compressing RGB images with precorrection for chromatic aberration.

DIFFERENTIAL PROCESSING OF REFLECTION AND ROTATION SYMMETRIES IN VISUAL TEXTURES [2.23]

Rachel Moreau, Peter Kohler, Erez Freud, Alasdair Clarke, & Nihan Alp, York University

Symmetries of various types are prevalent in the natural world. Psychophysical studies show that reflection symmetry (found in faces, bodies) can be detected preattentively, requiring less cognitive resources than other symmetry types such as rotation (found in flower petals, snowflakes). The distinction between symmetry types is important to our understanding of how symmetries contribute to perception of scenes and objects. Visual search has previously been used to probe the distinction between serial and parallel processing of cues to object shape (Enns and Rensink, 1991). Here we employ a visual search paradigm with symmetries that are embedded within regular textures. Our goal is to enhance our understanding of the mechanisms responsible for perceiving symmetries in textures, and differentiate between reflection and rotation symmetry. Based on previous findings, we hypothesize that reflection will elicit more parallel processing than rotation. We conducted four visual search experiments in which participants were presented with arrays of tiles containing symmetries. Participants were asked to report the presence of a target tile that had no symmetry and thus disrupted the regularity. We used four different array sizes (total # tiles: 9, 16, 25, 36), and trials with each array size were presented in an interleaved fashion. In Experiment 1, the non-target tiles contained reflection symmetry (N=133), while in Experiment 2 they contained rotation symmetry (N=148). In Experiments 3 and 4 we introduce a "jitter" component to our arrays which disrupts the textural regularity of these displays. This allows us to manipulate if the effects of symmetry type change when embedded in textural regularity or when they are isolated. In all four experiments we found that accuracy was reduced and RT was increased as array sizes got larger, consistent with serial processing. Importantly, this array size effect was reduced for reflection symmetry relative to rotation symmetry. Additionally, in our jitter experiments, where textural regularity was disrupted, we found a consistent pattern but higher RT and reduced accuracy overall. These results suggest that reflection symmetry elicits more parallel processing than rotation and when embedded in regular texture, processing of symmetries is more efficient than without.

AFFECTIVE PRIMING AS A BEHAVIOURAL MEASURE OF IMPLICIT COVID-19 ATTITUDES IN PEOPLE WHO ARE VACCINE HESITANT [3.1]

Stefania S. Moro & Jennifer K. E. Steeves, York University

Public health safety measures including precautionary behaviours such as border closures, total or partial lockdowns, social distancing, and face mask mandates have been employed to reduce the spread of the novel coronavirus (COVID-19). The enforcement of these types of precautions has been shown to contribute towards increasing the risk for pervasive mental health problems and psychological fear-related responses. Conversely, high-risk behaviours such as, ignoring recommendations for social distancing (observed through sustained group gatherings) and continuing to travel despite restrictions have also been observed accelerating the spread of the disease. An adaptive response is generated through human information processing where incoming stimuli are rated as pleasant or unpleasant, prior to cognitive analysis of the stimulus. The affective priming paradigm has been shown to be effective in illustrating this type of implicit evaluative response by investigating whether the assessment of a first stimulus (the prime) affects the processing of a subsequent stimulus (the target). A facilitation effect, observed as a faster response time, emerges when a polarized target word is preceded by a congruently-polarized prime word rather than an incongruently-polarized prime word. Recently, through the observation of prime-target word pairs that belonged to either pleasant or unpleasant affective categories we measured affective priming as an indirect behavioural measure aimed at evaluating implicit COVID-19 attitudes (Moro & Steeves, Scientific Reports, 2021). Additionally,

we directly measured COVID-19 attitudes through the COVID-19 Pandemic Mental Health Questionnaire (CoPaQ). Explicitly, participants rated their overall risk perception associated with contracting COVID-19 significantly lower compared to their perception of necessary precautions and overall adherence to public health measures. During baseline trials, participants explicitly rated COVID-19 affiliated words as unpleasant, like traditional unpleasant word stimuli. Despite rating the COVID-19 affiliated words as unpleasant, affective priming was not observed for congruent prime-target COVID-19 affiliated word pairs when compared to congruent prime-target pleasant and unpleasant words. Overall, these results provide quantitative evidence that COVID-19 affiliated words do not invoke the same implicit attitude response as traditional pleasant and unpleasant word stimuli, despite conscious explicit rating of the COVID-19 words as unpleasant. In this study, we utilized the same experimental procedure to evaluate COVID-19 attitudes in participants who are vaccine hesitant compared to those who are pro-vaccine. Both groups explicitly rated their overall risk perception associated with contracting COVID-19 significantly lower compared to their perception of necessary precautions and overall adherence to public health measures. Pro-vaccine participants rated their risk perception associated with contracting COVID-19 and their perception of necessary precautions higher compared to vaccine-hesitant participants. During baseline measures, both groups classified COVID-19 affiliated words as unpleasant. Affective priming was observed for congruent prime-target COVID-19 word pairs for pro-vaccine participants but not for vaccine hesitant participants. Results from this study complement the previous findings indicating an attitude-behaviour discrepancy that may contribute to decreased fear-related behaviours and extend them to further examine whether the increase incidence in risky-behaviour is more prevalent in vaccine hesitant demographic of the population.

FUNCTIONAL BRAIN NETWORKS OF EGOCENTRIC AND ALLOCENTRIC MEMORY-GUIDED REACHING [2.17]

Lina Musa, Amirhossein Ghaderi, Ying Chen, & J. Douglas Crawford, York University

The location of a remembered reach target can be encoded in egocentric and/or allocentric reference frames. While the differences in the cortical activation of these two representations have previously been identified (Chen et al., 2014; Neggers et al., 2006), differences in the functional organization of brain networks have not been described. Chen et al. (2014) demonstrated that allocentric and egocentric reach mechanisms use partially overlapping but distinct cortical substrates, that differ in directional selectivity of the target during the memory delay and response. Higher activation in dorsal brain areas (the parietofrontal cortex) was characteristic of egocentric reaching, while allocentric task also involved activity in ventral brain areas (inferior temporal gyrus and inferior occipital gyrus). It is thus expected that the size and connectivity of those functional brain networks will differ, reflecting more widespread connectivity between dorsal and ventral brain areas in the allocentric task. Here, we performed a secondary analysis of an event-related fMRI design to distinguish human brain networks involved in these two forms of representation. 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). The properties of brain networks involved were analyzed using graph theory to derive the efficiency of the network and the clustering between the nodes. Despite higher clustering of brain areas in the egocentric task, we observe similar global efficiency in the allocentric task. The need for economy in brain network organization should lead to more global efficiency in a network with more adjacent nodes (in the egocentric representation), however we find a similar global efficiency in the allocentric task, perhaps due to an adaptive trade-off of task benefit with energy costs.

SPATIAL CODING IN PREFRONTAL VISUAL RESPONSES DURING A REACH TASK [3.7]

Veronica Nacher, Parisa Abedi Khoozani, Vishal Bharmauria, Harbandhan Arora, Xiaogang Yan, Saihong Sun, Hongying Wang, & J. Douglas Crawford, York University

Eye-Centered visual codes have been observed throughout the dorsal visual stream and cortical saccade system, but the nature of visual coding in prefrontal cortex is less clear. We examined this question by recording single neurons from dorsolateral prefrontal cortex (DLPFC) in two monkeys trained to perform a head-unrestrained reaching paradigm. Animals touched one of three central LEDs at waist level while maintaining gaze on a central fixation dot and were rewarded if they touched a target appearing at one of 15 locations in a 40° x 20° (visual angle) array. Preliminary analysis of 509 responsive neurons in both monkeys showed an assortment of target/stimulus, gaze, pre-reach and reach related responses in DLPFC. Most neurons could be described as falling into three main groups: 'Early' (increased firing rate during the target presentation and gaze onset), 'Late' (increased firing rate near the end of the reach), and 'Sustained' responses that spanned both periods. Here, we focused on analysis of the 'visual' response of 92 spatially tuned 'early response' neurons, 80-180ms after visual target onset. We first tested for gaze, head, and hand gain fields during the different neuronal responses and after removing the gain field effects, we fitted the residual data against various spatial models related with target, eye, head, and arm. We found that the visual response best encoded the target relative to the space (Ts) and target relative to the head position (Th) at the population level with the target relative to the eye (Te) being significantly eliminated. At the single unit level, preferred fits were distributed across all three visual models (Te, Th, Ts) and when other motor models were tested, some 'visual' responses actually preferred parameters like future head position. These data suggest that early 'visual' responses in DLPFC show complex levels of spatial processing during reaching, perhaps for action planning.

EXTENDING THE RANGE OF DEPTH CAMERAS USING LINEAR PERSPECTIVE [3.4]

Tasneem Naheyan & James H. Elder, York University

Stereoscopic depth cameras are applicable to a variety of tasks including 3D SLAM and machine perception for autonomous robotics. These cameras provide a cheaper alternative to LiDAR for depth sensing. However, while performing reasonably well in the near field, accuracy of depth estimation is unreliable at greater distances. The human visual system addresses this problem by taking into consideration monocular visual cues in addition to binocular cues. The aim of this research is to extend the range of depth cameras by incorporating the monocular cue of linear perspective. Depth cameras with greater range will improve a robot's 3D SLAM capabilities, enabling smooth and reliable autonomous navigation, and will allow for better scene understanding. We focus on recovering depth in the built environment and exploit the Manhattan constraint that assumes that the scene contains three families of mutually orthogonal planar surfaces. The edges of these surfaces in turn form three families of mutually orthogonal line segments, generating three vanishing points in the image, and detection of these vanishing points allows the 3D orientation of the line segments to be estimated. Regressing onto the stereo depth estimates allows us to localize these line segments in 3D space, providing both interpolation and extrapolation of the stereo depth estimates. We evaluate this approach on the Deep Depth Completion (DDC) dataset.

CHARACTERIZING VISUAL ATTENTION NETWORKS: A NOVEL TECHNIQUE FOR RESTING-STATE FUNCTIONAL CONNECTIVITY ANALYSES [2.16]

Katherine M. Newman, Vicente Alba Suarez, Gary Turner, Mark Halko, & W. Dale Stevens, York University

Changes in visual attention (e.g., alerting, orienting, and sustaining attention) are observed in healthy aging and a variety of neuropsychological conditions. As the average lifespan extends, a corresponding rise in neurological events (e.g., ischemic strokes) that contribute to visual deficits (e.g., spatial neglect) can be expected. A rise in visual attention deficits suggests a need to advance our understanding of the functional mechanisms underlying such processes. Innovative methods in resting-state functional magnetic resonance imaging (rsfMRI) have facilitated the study of intrinsic functional brain networks underlying visual attention. The bilateral frontal eye fields (FEF), intraparietal sulci (IPS), and superior parietal lobules (SPL) are highly interconnected cortical regions of the dorsal attention network (DAN), which facilitates the allocation and control of visual attention. The inferior frontal gyri (IFG) and temporoparietal junction (TPJ) are central nodes of the ventral attention network (VAN), which directs attention to salient but unexpected stimuli in the external environment. The functional specificity of these regions and the ability to modulate their activity with repetitive transcranial magnetic stimulation (rTMS) make them clinically relevant. Standard resting-state functional connectivity (RSFC) analyses of intrinsic functional networks typically average the blood oxygen level-dependent (BOLD) signal time-series across all voxels/vertices within a parcel (i.e., network node) based on group-level or literature defined network topology. Importantly, these techniques ignore individual differences in network- and node-level spatial topology that have been shown to be critical in accurately characterizing and quantifying network properties and targeting nodes for rTMS. Our objective was to develop a new approach to localizing key network nodes at the level of the individual to improve the characterization and quantification of intrinsic network properties using resting-state functional connectivity (RSFC) analyses. We believe our approach will provide insight into individual differences in functional networks related to visual attention while improving rTMS treatment efficacy. To begin, structural MRI data underwent cortical reconstruction and rsfMRI data were processed with the Group Prior Individual Parcellation (GPIP) method to identify subject-specific parcel boundaries. GPIP uses an anatomical atlas and group-based functional parcellation as prior constraints before defining parcel boundaries using individual rsfMRI data. For each parcel, patterns of BOLD activation were correlated for all pairs of vertices to identify the vertex most highly correlated with the rest of the parcel. We hypothesize that this vertex best represents the dynamics of the parcel, making it a strong candidate site to define connectivity between nodes and modulate using rTMS. Here we present a proof-of-concept analysis using data from the Human Connectome Project. The time-series of the most correlated vertex in each DAN and VAN node from GPIP was used to quantify the relationships between nodes. These relationships were also defined using literature-based coordinates and compared to our method. Preliminary analyses suggest that this technique will improve the quantification of visual attention network properties and localization accuracy of critical nodes for rTMS intervention.

CORTICAL INTEGRATION OF MULTIMODAL CUES FOR REACH AND GRASP PLANNING [1.9]

Gaelle Nsamba Luabeya, Ada Le, Erez Freud, Simona Monaco, & J. Douglas Crawford, York University

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 focused our analysis on the second delay period and the execution phase in order to capture the integration of the two cue information. Therefore, we conducted a Representation Similarity Analysis to assess which theoretical models best represent the pattern of activities observed in reach- and grasp-related brain. We expected to see more simple models explain the pattern of activities in the early visual regions, and more complex models would explain the pattern of activities in the higher-level reach and grasp regions. We found that the activities in the Early Visual Cortex were best explained by the location model, and as the model became more complex by addition orientation and order, there was more noise. In addition, the left Superior Parietal Occipital Cortex showed best fits for location & order model, and left dorsal premotor for location, order & orientation model. Nevertheless, as we looked at higher-level visual and action- related regions, the models were less effective at explaining the observed activity. The next steps in our analysis are to examine the execution phase data and test if reach-grasp integration is revealed in functional connectivity of the system.

PERCEPTION OF SELF- AND EXTERNALLY-GENERATED VISUAL STIMULI: EVIDENCE FROM EEG AND BEHAVIOUR [1.8]

Edward Ody, Benjamin Straube, Yifei He, & Tilo Kircher, University of Marburg

Efference copy-based forward model mechanisms may help us to distinguish between self- and externally-generated sensory consequences. Previous studies have shown that neural activity and perception associated with self-initiated sensory consequences are reduced (sensory suppression). For example, event-related potentials (ERPs) elicited by tones that follow a button press are reduced in amplitude relative to ERPs elicited by passively attended tones. However, previous EEG studies investigating visual stimuli in this context are rare, provide inconclusive results, and lack adequate control conditions with passive movements. Furthermore, although self- initiation is known to modulate behavioural responses, it is not known whether differences in the amplitude of ERPs also reflect differences in perception of sensory outcomes. In this study, we presented to participants visual stimuli consisting of grey discs following either active button presses, or passive button presses, in which an electromagnet moved the participant's finger. Two discs presented visually 500-1250ms apart followed each button press, and participants judged which of the two was more intense. Early components of the primary visual response (N1 and P2) over the occipital electrodes were suppressed in the active condition. Interestingly, suppression in the intensity judgement task was only correlated with suppression of the visual P2 component. These data support the notion of efference copy-based forward model predictions in the visual sensory modality, but especially later processes (P2) seem to be perceptually relevant. Taken together, the results challenge the assumption that N1 differences reflect perceptual suppression and emphasise the relevance of the P2 ERP component.

EXPLORING VOLUMETRIC VIDEO FOR INCREASED REALISM IN NURSE DEBRIEFING [3.15]

Colin Orian, Jakob Anderson, Andrew Hogue, Alvaro Quevedo, & Adam Dubrowski, Ontario Tech University

In nursing, debriefing allows trainees to revise, discuss, and reflect how a task was performed. The debriefing process typically involves video recordings and may include feedback from peers and/or instructors. The video recordings allow trainees to observe actions done without having to rely on

memory recollection of the task. However, such an approach is limiting as the camera only captures the actions from a fixed point of view, thus missing critical information if areas of interest are occluded by the trainees or equipment. Therefore, we propose using volumetric video (VV) by recording actions employing multiple cameras with depth information that can create a virtual reality (VR) environment for immersive and interactive debriefing without the limitations of a single camera point of view. We hypothesize that the use of VV will be received well in focus groups and improve the debriefing experience. We currently use RGB-D cameras (Azure Kinect) to create a VV capture, whereby the captured camera data can then be processed to create a temporally updating 3D model that reflects the performance captured. This 3D model can then be imported into game engine tools such as Unity3D to add interactivity and the ability to view the video from any spatial location using VR. The system uses Soar to record the VV and utilizes Soar's Unity package for video playback. Communicating with potential users found that tools to interact with the VV may be beneficial. Potential users also explained that the time between the procedure and the debriefing should be short so that trainees can retain more information from the procedure. We expect that trainees will have an improved debriefing experience through the use of the VR environment, as they are more immersed in the experience and able to observe the procedure in any position. Our proposed methodology is the following: We will first collect demographic info of nurse students participating in the study. The students will be split into two groups. Each group will perform a simulated medical procedure, such as CPR. The control group will debrief using a normal video. The experimental group will debrief using VV in a VR environment and be in the same physical space as peers and instructors. The experimental group will use a head mounted display to be immersed in the VR environment. The VR environment will also have tools to interact with the video. Both debrief groups will have an instructor leading it. Similar to Verkuyl et al., the Debriefing Experience Scale will be used to measure potential improvements in using the VR environment. A multiple choice test will be conducted to measure the knowledge acquired by trainees. The KR-20 test will measure how reliable the knowledge is.

VIDEO-BASED TRAFFIC ANALYTICS AT INTERSECTION [1.5]

Sajjad Pakdamansavoji & James H. Elder, York University

Rising traffic challenges due to lack of proper traffic management need to be addressed with intelligent solutions like Traffic Analytics Software with integrated video-analytics which can effectively aid traffic management. The necessity of using computer vision in such a software, stems from the fact that conventional traffic monitoring schemes mostly gather visual data such as videos and images. Therefore, designing this software based on the current traffic monitoring infrastructure rules out the need for additional hardware installation. Our system leverages existing traffic monitoring cameras and applies computer vision techniques to provide detailed traffic analysis results. It can localize, segment, and track vehicles, pedestrians, traffic signs and other related objects in the intersection. By doing so, it provides informative data regarding traffic volume, movements of interests, rule violations, congestions, and so on. The Traffic Analytics Software follows a pipeline consisting of four main blocks: detection and segmentation, tracking, reprojection, and analysis. The first two blocks are responsible for capturing relevant information regarding the dynamics of the targeted objects. These two blocks commonly work separately, meaning that the output of the first one is fed into the second one on test time while their training is performed independently. Such a decoupled structure is an obstacle preventing the end-to-end fine tuning of the pipeline hence in this study we will attempt to integrate these blocks to one unified differentiable unit performing both tasks jointly; aiming to resolve some of the most fundamental challenges in this area including long-term tracking and id-switch due to partial occlusion. The third block in the pipeline is responsible for reprojecting points from the image coordinates onto the real-world coordinates. To do that we will use calibrated cameras and pair-

matching to estimate the required homography matrix for such a reprojection. Our intuition as well as the qualitative results show that reprojecting points and trajectories onto the ground plane will improve the analysis; hence we will conduct ablation studies to confirm this hypothesis with quantitative results as well. As for the last block, analysis, depending on the downstream task several algorithms are used. For instance, in the case of estimating traffic flow and recognizing movements of interest (MoI) Contour Mapping (CMM) is used to match a trajectory into a commonly traversed path. In such an algorithm automatically choosing a common path is significant as selecting them via human annotation requires both time and money. As such, we will explore unsupervised methods such as clustering as alternatives for the traditional approaches. Beside these main blocks, our software will consist of a digital intersection visualizer. This visualizer can map events in the intersection such as a car crossing or a rule violation into a fabricated model of the intersection. Having this visualizer will facilitate the dissimilation of information to users while providing the developers with a simple approach of testing and improving the software itself.

LIGHTNESS AND BRIGHTNESS CHARACTERIZED VIA DECISION SPACES, IN REAL AND RENDERED SCENES [1.7]

Jaykishan Patel, Khushbu Patel, Emma Wiedenmann, & Richard F. Murray, York University

Lightness and brightness have extensive research literatures, but their relationship is controversial. We used decision spaces to characterize them and test computational models. In Experiment 1, we used a custom-built apparatus where adjustable reflectance patches were visible through two apertures, and illumination at the two apertures could be set independently. On each trial, reflectance and illuminance at the reference aperture were set to one of three settings. Reflectance and illuminance at the test aperture were randomly set to +/- 50% of the values at the reference aperture. In the lightness and brightness conditions, observers judged which aperture had a higher reflectance or luminance, respectively. For each of the three reference stimuli, we plotted the probability that the observer judged the test stimulus as lighter (or brighter), as a function of test reflectance and illuminance. Each such decision space was approximately divided in two by a straight line whose orientation varied across conditions. In the lightness task, the decision spaces were consistent with partial lightness constancy, with Thouless ratios around 0.80. In the brightness condition, Thouless ratios were lower, but decision spaces still indicated judgements closer to reflectance than to luminance judgements. In Experiment 2, we repeated this procedure with a rendering of the same apparatus on a monitor. Decision spaces were similar to those in Experiment 1, but indicated judgements more strongly influenced by luminance. Finally, we simulated computational models of lightness and brightness: ODOG, a high-pass model, a contrast normalization model, and two retinex models. All models' decision spaces were highly inconsistent with those of human observers. We conclude that (a) lightness and brightness judgements are more similar than expected from previous work, (b) brightness is nothing like an estimate of luminance, and (c) current computational models can fail on even simple lightness and brightness judgements.

LIGHTNESS CONSTANCY IN REALITY, IN VIRTUAL REALITY, AND ON FLAT-PANEL DISPLAYS [1.16]

Khushbu Patel, Laurie M. Wilcox, Laurence T. Maloney, Krista A. Ehinger, Jaykishan Y. Patel, Emma Wiedenmann, & Richard F. Murray, York University

Virtual reality (VR) technology is being used in an increasing number of applications. However, research has shown that we often perceive surface properties differently in real and virtual environments. To evaluate how well virtual platforms support realistic lightness perception, we measured lightness constancy in a physical scene, 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 on a custom-built apparatus 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 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 for lightness constancy were significantly higher (indicating greater constancy) in the physical condition (mean and 95% confidence interval 0.87 ± 0.04) than in the flat-panel condition (0.79 ± 0.08). Thouless ratios were not significantly different in the VR condition (0.83 ± 0.08) than in the physical condition or the flat-panel condition. Thus in the simple scenes considered here, lightness constancy is moderately lower in virtual environments than in physical environments. This discrepancy should be considered when developing applications where realistic performance is critical, but our results suggest that VR can be a flexible alternative to flat panel displays and a reasonable proxy for real environments.

AUTOMATED USABILITY ASSESSMENT OF INDIVIDUALS THROUGH OUTDOOR STREET CROSSINGS AND WALKWAYS [1.11]

Jakson Paterson, Hamed Ghomashchi, Zeyad Ghulam, & Tilak Dutta, KITE Research Institute

Seasonal variations can cause a wide range of dangerous and difficult situations for sidewalk and street crossing users in many countries, including Canada. The fluctuation of temperature and precipitation creates hazardous ground surface conditions which discourages older adults and persons with disabilities from outdoor excursions due to the risk of falling. Outdoor fall-related hospitalizations increase in the winter with a wide range of injuries including fractures, joint dislocations, concussions, and other head/spine injuries. In addition, outdoor activity decreases when snow or ice are present leading to individuals becoming isolated and refraining from leaving their homes due to a fear of falling. The isolation experienced in winter has been shown to lead to deconditioning, increased risks of mortality, increased frailty, and can lead to more falls in the future due to the reduced daily physical activity. Despite the magnitude of the problem, there has been limited research focused on understanding how outdoor environment design can facilitate the needs of sidewalk and street crossing users. To address the limitations with current assessment methods, our team is developing an automated detection and tracking system to isolate the individuals most at-risk of falling for manual observation. This system will use pre-recorded bird's-eye-view video of a downtown Toronto intersection and apply computer vision methods to assess the behaviour of individuals in different weather conditions and around environmental barriers. This approach will allow for real-time quantitative research on the perception of environmental influences and focus on the experience of the most at-risk group of individuals. Our lab proposes that current computer vision approaches can be used for the following tasks: i) Develop an automated detection and tracking system for estimating the walking speed and trajectory of sidewalk and street crossing users with birds-eye-view video recordings. ii) Identify differences in individual walking trajectory and behaviour for individuals who fall into the 5th percentile for walking speed during different adverse weather conditions. iii) Individual trajectories and behaviours will also be compared to the top two quartiles of walking speeds to identify opportunities to improve the design, maintenance, and perception of street-crossings. This project will demonstrate the potential of computer vision approaches for automating the real-time

observation of community-dwelling sidewalk and street crossing users. This work will help develop a better understanding of the complex relationship between individuals with mobility limitations and the built environment so that the accessibility of outdoor spaces can be improved.

DOWNWARDS VERSUS UPWARDS MOTION - MORE ARISTOTELIAN THAN WE THINK [2.18]

Mai Huong Phan, Laurence R. Harris, John Jong-Jin Kim, & Frederick Kingdom, York University

Although studies have measured sensitivity to visual acceleration and deceleration as a function of motion direction, an unanswered question is whether our perception of acceleration/deceleration is biased depending on motion direction. We tested the hypothesis that objects moving downwards were perceived to accelerate less and/or decelerate more than objects moving upwards with the same acceleration/deceleration, on the grounds that we have adapted to the more commonly experienced downward acceleration. Observers viewed an image of a ball moving from one side of a room to the other in one of four directions - up, down, left, right - in an image of a square room with a floor, ceiling, wallpapered walls, a door, human figures etc. On each trial participants indicated whether they perceived the ball to accelerate or decelerate. The ball moved at one of two average speeds, taking either 1.0s or 1.5s to traverse the room, and there were several values of acceleration and deceleration. Psychometric functions were fitted to the proportion of perceived accelerations as a function of physical acceleration/deceleration. The point-of-perceived-constant-velocity, or PPCV, was estimated as the physical acceleration/deceleration corresponding to the point at which acceleration is perceived 50% of the time. Our hypothesis predicted that for participants to perceive a downwards moving object as moving at constant speed it needed to accelerate more or decelerate less than that of an upward moving object. Data from 11 participants showed a significant bias in PPCVs in the predicted direction. No significant differences in bias were found between rightward and leftward moving objects, and no significant differences were found for any pair of directions in the slopes of the psychometric functions i.e., in sensitivities. We conclude that downwards acceleration is more likely to be identified as constant velocity than upward acceleration.

A METHOD FOR IMPROVING REGRESSION AND CORRELATION COEFFICIENT ESTIMATES IN THE PRESENCE OF NOISE [1.12]

Jason Pina & Joel Zylberberg, York University

A key challenge in neural and vision sciences is estimating relationships between biological, behavioural, or cognitive variables in the presence of noise. Such noise, or measurement error, arises from uncertainties due to either recording device limitations or intrinsic biological variability, both of which are ever-present in biological and psychological experiments. This noise can greatly reduce estimated linear regression and correlation coefficients, as well as the fraction of explained variance (or R^2 value), compared to their true values. Despite the large impact this *regression dilution* can have on downstream inferences that result from these reported estimates, it is seldom accounted for. However, in many experiments, data that can be leveraged to eliminate this bias is already collected, as the relevant variables are often averages of multiple observations. We present a simple, easy-to-implement method that utilizes these multiple measurements to estimate the noise variance and allow for the regression dilution effect to be removed. Using simulated data, we show that the confidence intervals from our unbiased estimator indeed consistently capture the underlying regression and correlation coefficients, in sharp contrast with those from the uncorrected estimates. This demonstrates that our estimator is directly applicable to any correlation or linear regression analysis for which the data are based off of averages, including neuronal, behavioural, and psychophysical or psychological data that are averaged over subjects, space (pixels/neurons), or time. Some common examples from neuroscience collected from spike counts, fMRI, EEG, or PET scans include noise and signal correlations,

inter-subject correlations (ISCs), functional connectivity analyses, and representational similarity matrices (RSMs). As a specific example of how our estimator can lead to new neuroscientific insights, we applied our method to 2-photon calcium imaging data from VisP from recent experimental work. In the experiment, neuronal responses to two classes of novel, visual stimuli were compared across separate recording sessions in the same neurons to determine whether changes consistent with learning were observable. Detecting these changes required ruling out regression coefficients equal to one (corresponding to no change), which standard methods are unable to do: due to regression dilution, the coefficients estimated from noisy data will be less than one even when the neural responses do not change. In contrast, our approach successfully identified several conditions with low uncorrected regression coefficients as being consistent with a true, underlying regression coefficient of ~1, while ruling out several others as indeed having true coefficients well below 1, findings that better delineate the experimental conditions under which individual neuronal changes occur. Thus, our simple method to account for the effects of noise on estimates of correlation and linear regression coefficients sheds light on learning in the context of novel, visual stimuli, an active area of interest in neuroscience.

BIG DATA MEET DEEP DATA: CHARACTERIZING SPATIAL NAVIGATION IN HIPPOCAMPAL AMNESIA [2.28]

Sara Pishdadian, Antoine Coutrot, Michael Hornberger, Hugo Spiers, & R. Shayna Rosenbaum, York University

Amnesic patients with bilateral hippocampal damage have helped elucidate the role of the hippocampus in learning to navigate new environments, informing major theories such as Cognitive Map Theory (CMT) (O'Keefe & Dostrovsky, 1971). However, this neuropsychological approach is constrained by limited normative data. The mobile video game Sea Hero Quest (SHQ) has been administered to over 4 million people globally, and measures path integration and allocentric map abilities. SHQ performance is sensitive to national-level income and gender inequality (Coutrot et al., 2018) and correlates with real-world navigation (Coutrot et al., 2019). It can also distinguish individuals based on APOE ε4 status who are at higher genetic risk of Alzheimer's Disease, while also demonstrating acceptable test-retest reliability (Coughlan et al., 2020; Spiers et al, 2021). Combining the case study method with crowdsourced, benchmark data has the potential to distinguish subtle individual differences in navigation and strategy use in individuals with hippocampal amnesia. In this study we report the performance of two well-characterized individuals with amnesia due to bilateral hippocampal lesions of varying extents on SHQ. We compare their performance against data from thousands of age-, gender-, and geographic-matched control participants. Results reveal that while both individuals with hippocampal amnesia perform worse than their peers on most levels, they are not universally impaired on the tasks as may be expected. Their performance on trials of difficulty is marked by revisiting past locations, a 'backtracking' pattern not seen in control participants. Individuals with amnesia's performance is relatively better on trials where maps are not visible prior to navigating, demonstrating that amnesic individuals are not encoding the maps, consistent with their unique neuropsychological profiles. Overall, findings from this study speak to the SHQ's sensitivity and specificity to individual-level cognitive abilities and also illustrate the value of big data for understanding cognitive deficits at an individual level.

CONSUMER-LEVEL FACIAL TRACKING PROTOTYPE FOR NON-IMMERSIVE VIRTUAL REALITY REMINISCENCE THERAPY [1.6]

Daniel Presas, Rabia Akhter, Winnie Sun, Alvaro Uribe-Quevedo, Ramiro Liscano, & Sheri Horsburgh, Ontario Tech University

Currently, more than 55 million people worldwide live with dementia, which is the leading cause of

death, disability, and dependency among the elder according to the World Health Organization. RT is often conducted employing pictures, videos, and audio for memory recollection purposes. However, this approach lacks interactivity and immersion found in interactive media including mobile applications, websites, or virtual reality that can further impact triggering memories. This work presents a consumer-level facial tracking prototype employing OpenCV, Unity, a web camera, and a neural network for the recognition of seven facial expressions. The facial expressions allow capturing reactions to the digital content, which can help advance the utilization of reminiscence therapy. A preliminary study with three patients with dementia was conducted at the Ontario Shores Center for Mental Health Sciences using a reminiscence therapy web tool that integrates multimedia from the patient. The process had a therapist navigating the RT content with a patient seated in front of a 50-inch screen and a web camera for facial tracking within 20-30 cm. During the RT session, facial expressions and time stamps were capture to identify responses to content shown. After concluding the session, the therapists were interviewed to gather their perceptions about the use of this technology to support dementia care. The interviews provided the following suggestions towards improving the overall usability and complexity of the facial tracking: i) user-friendly data visualization matching the content, timestamps, and facial expressions since currently, a spreadsheet is used; ii) a virtual notepad to add notes during the therapy to provide context associated to the content and facial expressions; iii) improve the recognition of subtle facial expressions relevant to patients with conditions such as Bell's palsy and other neurological disorders. From a technical standpoint, we also noticed that the system struggled differentiating the patient from the caregiver and detecting subtle facial expressions. Future work will focus on improving the calibration process, flow of actions, reliability, and leverage digital twins for overcoming the limitations of public datasets.

EVIDENCE FOR A TEMPORAL CORTEX ATTENTION NETWORK IN HUMANS AND MONKEYS [1.10]

Hamid Ramezanpour & Mazyar Fallah, York University

Attention is an indispensable component of active vision. Contrary to the widely accepted notion that temporal cortex processing primarily focuses on passive object recognition, a series of very recent studies emphasize the role of temporal cortex structures, specifically the superior temporal sulcus (STS) and inferotemporal (IT) cortex, in guiding attention and implementing cognitive programs relevant for behavioral tasks. The goal of this theoretical paper is to advance the hypothesis that the temporal cortex attention network (TAN) entails necessary components to actively participate in attentional control in a flexible task-dependent manner. First, we will briefly discuss the general architecture of the temporal cortex with a focus on the STS and IT cortex of monkeys and their modulation with attention. Then we will review evidence from behavioral and neurophysiological studies that support their guidance of attention in the presence of cognitive control signals. Next, we propose a mechanistic framework for executive control of attention in the temporal cortex. Finally, we summarize the role of temporal cortex in implementing cognitive programs and discuss how they contribute to the dynamic nature of visual attention to ensure flexible behavior.

DIFFERENCES IN RESTING-STATE FUNCTIONAL CONNECTIVITY UNDERLIE VISUOMOTOR TASK PERFORMANCE DECLINES IN OLDER ADULTS WITH A GENETIC (APOE E4) RISK FOR ALZHEIMER'S DISEASE [1.2]

Alica Rogojin, Diana J. Gorbet, Kara Hawkins, & Lauren E. Sergio, York University

Non-standard visuomotor integration requires the interaction of large networks in the brain. Previous findings have shown that non-standard visuomotor performance is impaired in individuals with specific dementia risk factors (family history of dementia and presence of the APOE e4 allele) in advance of any cognitive impairments. These findings suggest that visuomotor impairments are associated with early

dementia-related brain changes. The aim of this project is to examine the underlying resting state functional connectivity (RSFC) associated with impaired non-standard visuomotor performance, as well as the impacts of dementia family history, sex, and APOE status. Participants included right-handed older adults with a family history of dementia (n=24, 11 female) or no family history of dementia (n=24, 12 female). APOE genotyping was determined from salivary measures. Participants were tested on four visuomotor tasks where reach and gaze were increasingly spatially dissociated using two linked touchscreens. These included a standard condition requiring direct interaction with visual targets, and three non-standard conditions (involving either visual feedback reversal, plane-change, or plane-change feedback reversal). To quantify RSFC within networks of interest, an echo planar imaging (EPI) sequence sensitive to blood oxygenation level dependent (BOLD) contrast was collected. The networks of interest were the default mode network (DMN), somatomotor network (SMN), dorsal attention network (DAN), ventral attention network (VAN), and frontoparietal control network (FPN). Individuals with the e4 allele showed abnormalities in RSFC between posterior DMN nodes that predicted poorer non-standard visuomotor performance. Specifically, multiple linear regression analyses revealed lower RSFC between the precuneus and posterior cingulate cortex, a posterior functional core of the DMN, and the left IPL ($\beta = -16.303$, $p < 0.05$) and left parahippocampal cortex ($\beta = -27.278$, $p < 0.05$). Presence of the APOE e4 allele also modified the relationship between mean DAN RSFC and visuomotor performance for two of the behavioural measures in the plane-change feedback reversal condition. There was a significant interaction effect of mean RSFC in the DAN and APOE status on endpoint error scores ($\beta = -25.543$, $p < 0.01$) and corrective path length ($\beta = -59.268$, $p < 0.05$), both indicative of worse visuomotor performance, only in APOE e4 carriers. There were otherwise no effects of family history, APOE e4 status, or sex on the relationship between RSFC and visuomotor performance for any of the other resting networks. The preliminary findings provide insight into the impact of Alzheimer's disease pathology on neural networks underlying complex visuomotor transformations, and demonstrate that the non-standard visuomotor task paradigm discussed in this study may be used as a non-invasive, easily accessible assessment tool for dementia risk.

THE ROLE OF ATTENTION IN ACQUIRING FRAMES [2.1]

Mohammad Shams-Ahmar, Peter Kohler, & Patrick Cavanagh, York University

The perceived locations of probes flashed within a moving frame are dramatically displaced as if they are seen in frame coordinates rather than world coordinates. Here we demonstrate how a moving frame generates these effects. First, we will demonstrate that the frame's effect on the flashed probes can be engaged exogenously: when several frames move back and forth simultaneously but in different directions, the one that becomes salient determines the relative position of two alternatively flashing probes. Second, we will show that the frame's effect can be initiated endogenously: when a set of frames that can be organized into either horizontal or vertical motion is presented, the relative location of the probes depends on which motion path was attended. Finally, we will show a series of videos that suggest that besides motion, the size and location of the "attention field" is key to perceiving the frame effect at its maximum power.

TOWARDS DEVELOPING MECHANISMS AND ANALYSIS TOOLS FOR FEW-SHOT VIDEO UNDERSTANDING [3.32]

Mennatullah Siam, Matthew Kowal, Amirul Islam, Richard P. Wildes, & Konstantinos Derpanis, York University

Few-shot video understanding aims at conducting spatiotemporal localisation through segmenting actions/actors within a video using few labelled trimmed videos. Towards this goal, we study two main questions: (i) how to enforce temporal consistency of the predictions within a video? and (ii) how to quantify the models ability to capture dynamic information, which is a crucial component when looking

into generalization to novel unseen actors/actions. In the first question, we focus on the simpler task of few-shot video object segmentation (FS-VOS) which segments video frames using a few labelled examples of classes not seen during the initial training. We present a simple but effective temporal transductive inference (TTI) approach that leverages temporal consistency in the unlabelled video frames during few-shot inference. Key to our approach is the use of both global and local temporal constraints. The objective of the global constraint is to learn consistent linear classifiers for novel classes across the image sequence, whereas the local constraint enforces the proportion of foreground/background regions in each frame to be coherent across a local temporal window. These constraints act as spatiotemporal regularizers during the transductive inference to increase temporal coherence and reduce overfitting on the few-shot support set. Empirically, our model outperforms state-of-the-art meta-learning approaches in terms of mean intersection over union on YouTube-VIS by 2.8%. In addition, we introduce improved benchmarks that are exhaustively labelled (i.e. all object occurrences are labelled, unlike the currently available), and present a more realistic evaluation paradigm that targets data distribution shift between training and testing sets. In order to investigate the second question, we propose a quantifiable interpretability technique that studies different spatiotemporal models. Currently, there is a limited understanding of what information is captured by these models in their intermediate representations. It has been observed that certain spatiotemporal models are heavily influenced by visual appearance in single static frames. However, there is no quantitative methodology for evaluating such static bias in the latent representation compared to bias toward dynamic information (e.g. motion). We tackle this challenge by proposing a novel approach for quantifying the static and dynamic biases of any spatiotemporal model. To show the efficacy of our approach, we analyse two widely studied tasks, action recognition and video object segmentation. Our study demonstrates the effects of datasets and architectural decisions on these biases. Our study serves as an analysis tool to help us gain insights into how to improve the generalization capability of spatiotemporal models beyond the training classes.

MODULATION OF CORTICAL ACTIVITY FOLLOWING DANCE TRAINING IN PARKINSON'S: AN FMRI CASE STUDY [2.26]

Royze Simon, Judith Bek, Rebecca Barnstaple, Rachel Bar, & Joseph F. X. DeSouza, York University

Dance is associated with a range of motor and non-motor benefits in people with Parkinson's disease (PD), and a recent 3-year longitudinal study found evidence that regular dance participation may delay progression of motor symptoms in people with mild PD. However, little is known about the neurobiological effects of dance interventions in PD. Neuroimaging in professional dancers has shown altered patterns of activation in supplementary motor area (SMA) and auditory cortex that reflect learning of a new choreography over time. The present study explored potential neuroplastic changes in a 69-year-old male with PD (disease duration 4-years) learning dance over 8 months of weekly classes. Functional MRI was performed at four timepoints, in which the participant performed first-person motor imagery of the choreography practiced during training, while listening to the corresponding music. Using GLM, neural activity was compared between imagery task blocks and fixation blocks. Functional data from the four sessions were used to create an averaged SPM functional map. Regions of interest were defined based on surviving clusters after applying a cluster threshold (k=22) and p-value threshold of P<.001 (Bonferroni corrected). Average BOLD signals from the four scanning sessions were extracted from the ROIs, and linear modelling analysis was performed to determine any modulation. Data analysis suggest signal changes in SMA and right superior temporal gyrus, showing a pattern of increasing followed by decreasing activation across scans. The results could provide new preliminary evidence on the neuroplastic effects of dance as a therapeutic activity for people living with PD.

DYNAMIC EBBINGHAUS VS THE CONTRACTING-EXPANDING SQUARE ILLUSIONS: SO SIMILAR AND YET NOT THE SAME [2.27]

Saki Takao, Katsumi Watanabe, & Patrick Cavanagh, York University

When a test disk is surrounded by smaller disks, it appears larger than when it is surrounded by larger disks (Ebbinghaus Illusion). Mruczek et al. (2015) reported that the strength of this illusion was almost doubled when the stimulus was in motion as the surround disks changed from large to small then reversed direction as the surround disks expanded again. This stimulus has deep similarities to the expanding-contracting square illusion (Anstis & Cavanagh, 2017) where a background texture repeatedly expands and contracts while a test square was flashed when the background was at its smallest size and then again when the background was at its largest size. The apparent size of the test square was twice as large when flashed on the small background compared to when flashed on the large background. We previously showed that this expanding-contracting square illusion is driven principally by the motion after each flash, not before (Takao et al., 2021). Here we examine whether the Dynamic Ebbinghaus version also shows this strong dependence on motion after versus before the test. In the experiment, four surround disks expanded and contracted continuously while the whole configuration moved back and forth diagonally, being smallest in size at one end and largest at the other. Rather than leaving the central test disk on continuously, it flashed briefly at each reversal point. The trajectory was broken into half cycles and the test disk flashed only at one end or the other to evaluate the influence of motion before versus after. Contrary to the results for the expanding-contracting square illusion, no differences were found in the Dynamic Ebbinghaus illusion strength for motion after vs motion before. This study indicates that the two illusions, despite their strong similarities, have very different causes.

BINOCULAR DEPTH AND DISTANCE CUES ENHANCE TOLERANCE TO VIRTUAL MOTION GAIN [3.10]

Xue Teng, Laurie M. Wilcox, & Robert S. Allison, York University

Motion parallax (from visual direction and optic flow) in natural environment supports both depth and distance perception. What happens if we receive conflicting visual/kinesthetic information? We manipulated motion gain using a VR headset and a two-phase task to assess perceived depth and distance. Observers first viewed a "fold" stimulus, a wall-oriented dihedral angle covered in Voronoi texture. The task was to adjust the dihedral angle until it appeared to be 90 degrees (perpendicular). We occluded the top and bottom edges of the fold and varied the width to make the edges of the fold uninformative. On each trial, following the angle adjustment, a second scene appeared which contained a pole that extended from a ground plane. In this phase, the task was to match the position of the pole to the remembered position of the apex of the previously seen fold. We tested observers in two viewing conditions (binocularly and monocularly) and two motion conditions (stationary and moving). When moving, observers swayed laterally through 20 cm in time to a 0.5 Hz metronome; the motion gain varied from 0.5 to 2.0 times the actual self-motion. We found that increased gain caused an increase in the adjusted angle or equivalently a decrease in associated depth of the fold, especially when viewed monocularly. In addition, perceived distance (adjustment of the pole) decreased with increasing gain, irrespective of viewing condition. That is, the fold was perceived as smaller and closer when gain was larger than 1. The effect of the gain manipulation was much weaker under binocular viewing. These data show that perceptual distortions due to differences between actual and virtual head motion are compensated for by binocular viewing, and to a lesser extent monocular, depth and distance cues. These flexible compensatory mechanisms make the human visual system highly tolerant of visual/kinesthetic mismatch.

MODULARITY AND SACCADE INFLUENCES IN THE CORTICAL VISION NETWORK [2.9]

George Tomou, Bianca Baltaretu, Amirhossein Ghaderi, & J. Douglas Crawford, York University

Considerable regional evidence has accumulated for dorsal-ventral modularity in the visual system, but it is not clear how these modules function at the whole brain network level, or how these networks are influenced by naturally occurring saccades. We addressed these questions using graph theory analysis of fMRI data collected during a task where participants had to remember, then discriminate between two different object features. Seventeen participants judged whether a remembered object changed shape or orientation with or without an intervening saccade. BOLD activation from 50 cortical nodes was used to identify local and global network properties, which indicated greater interconnectivity and efficiency of information transfer during saccades. A network modularity analysis revealed three sub-networks during fixation: a bilateral dorsal sub-network linking areas involved in visuospatial processing and two lateralized ventral sub-networks linking areas involved in object feature processing. Importantly, when horizontal saccades across the remembered object required visual comparisons between hemifields, the two lateralized ventral sub-networks became functionally integrated into a single bilateral sub-network. Comparisons of betweenness centrality between conditions identified several significant hub regions in occipital, parietal, and frontal cortex involved in linking distant network nodes during saccades. These results provide support of a ventral and dorsal stream distinction in human perception and show how hemispheric sub-networks are modified to functionally integrate information across saccades.

STEREOSCOPIC DISTORTIONS WHEN VIEWING GEOMETRY DOES NOT MATCH INTER-PUPILLARY DISTANCE [2.13]

Jonathan Tong, Laurie M. Wilcox, & Robert S. Allison

Virtual Reality (VR) headsets have displays and lenses that are either fixed in place, or adjustable to match a limited range or set of user interpupillary distances (IPDs). This means that often there will be mismatches between these settings and the viewer's IPD. Projective geometry predicts distortions of perceived scale when the baseline of the virtual cameras used to render stereoscopic images differs from that of the eyes. However, misalignment between the eyes and the lenses might also affect binocular convergence, which could also distort perceived depth. This possibility has been largely ignored in previous studies. Here, we evaluated this phenomenon in a VR headset in which the inter-lens and inter-axial camera separations are coupled and adjustable. In a baseline condition, both were matched to the observer's IPD. In two other conditions, the inter-lens and inter-axial camera separations were set to the maximum and minimum allowed by the headset. In each condition, observers were instructed to adjust a fold created by two intersecting, textured surfaces until it appeared to have an angle of 90°. The task was performed at three randomly interleaved viewing distances, both monocularly and binocularly. In monocular viewing, observers consistently over-estimated the fold angle (underestimated slant) and there was no effect of viewing distance on their settings. In binocular viewing, when the inter-lens and display separation was less than the observers' IPD they overestimated the fold angle (underestimated slant) relative to baseline; when the inter-lens and display separation was greater than their IPD, they underestimated the fold angle (overestimated slant) relative to baseline. These trends were predicted by a geometric model that incorporates shifts in binocular convergence due to the prismatic effect of decentered lenses, as well as the offset of dual displays relative to the eyes. However, our model tended to overestimate the effect of IPD mismatch on perceived slant at larger viewing distances. We determined that the model was able to better fit the data with an implausibly large virtual image distance (at least double the typical value reported for these devices). We considered an alternative possibility that the contribution of monocular cues might explain the geometric model's overestimation of slant bias. When the model was revised to include

relative weighting of monocular and binocular information, the fit to the data was much improved. Our results demonstrate that improperly matching users' IPD in VR headsets produces distortions in the perception of slant. Furthermore, our geometric model provides researchers and designers of VR-systems a means of predicting depth perception when the optics of head-mounted displays are misaligned with users' eyes.

EFFICIENT SINGLE-VIEW 3D VEHICLE GROUND-TRUTHING FOR VISUAL TRAFFIC ANALYTICS [3.6]

Trong Thao Tran, Yiming Qian, & James H. Elder, York University

Computer vision has become increasingly important for autonomous driving and traffic analytics applications. These applications demand accurate estimates of the 3D locations, poses, and shapes of motor vehicles on the roadway. Deep networks can be trained to estimate these 3D quantities from 2D imagery. However, training requires access to accurate 3D ground truth, typically obtained from LiDAR. Unfortunately, LiDAR systems are expensive, and returns become sparse in the far field. High-resolution RGB cameras are, on the other hand, relatively cheap and provide superior angular resolution. In this project, we explore whether human annotation of monocular data from such a camera can lead to ground-truth estimates of the 3D location, pose and shape of motor vehicles that are superior to estimates from LiDAR in the far field. We make use of the fact that motor vehicles are generally bilaterally symmetric, having a vertical symmetry plane along their long axis. Accordingly, pairs of annotated left-right symmetry points in the image (e.g., headlights) are connected by lines that converge to a common vanishing point on the horizon. These lines are parallel and normal to the symmetry plane. Assuming a fully calibrated camera, this vanishing point can be used to algebraically determine the orientation of both the symmetry plane and the vehicle. To identify the location of the vehicle, we make use of annotated image points projecting from points of contact between the vehicle's tires and the road. Since we know the height of these points, their 3D locations are known immediately. If the left and right tire contact points are visible and annotated, the location of the symmetry plane and hence the vehicle can be immediately determined algebraically. If tire contact points are visible and annotated on only one side of the vehicle, we make use of one or more pairs of extremal symmetry points, which are symmetric pairs of points that lie on the vehicle aligned laterally with the tires. When combined with one or more tire contact points located on one side of the vehicle, these can be used to localize the symmetry plane and the vehicle. Having fully identified the symmetry plane, algebraic estimates of the 3D coordinates of all annotated symmetry points can be determined and then refined iteratively to minimize reprojection error subject to symmetry constraints. We will demonstrate results on a traffic dataset derived from a field site in the Toronto region.

EVALUATING GAZE PERCEPTION WITH SIMULATED MOTION PARALLAX [1.30]

Viswajit Vembukumar & Nikolaus F. Troje, York University

Recently, the frequency of human interaction mediated by screens has exponentially increased. 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. A study conducted in virtual reality has shown that motion parallax had a greater effect in evoking the sense of presence when compared to stereopsis. The presence of motion parallax provides context to directional cues such as hand gestures or eye gaze. This study aims to examine whether the addition of simulated motion parallax increases participants' sensitivity to the gaze direction of faces on a screen. If adding motion

parallax can increase the sensitivity to gaze perception, cheaper and more widely accessible solutions could be developed to integrate the depth cue into standard video communication, thereby enhancing communication efficiency. By using motion capture technology to track a user's head location, we control a virtual camera whose movement in a virtual environment corresponds with the user's own motions. This allows the image on the screen to be dynamically rendered based on their position. The study examines two conditions: one uses head tracking to simulate the view onto a 3D head behind a window framed by the computer screen. In the other condition we present static images of the same head on the screen. In the Window condition the avatar's head and eye gaze are set to varying angles and users are asked to move themselves into the line of sight of the avatar. Once they reach a location at which they perceive eye contact their head location is recorded. The angular difference between the participant's head location and their expected position is then analyzed. In the static condition, the avatar's total gaze (head gaze plus eye gaze) is shown at angles between -11° and 11° deg around the fronto-parallel view, and participants indicated whether they perceived the head 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. From pilot data, in the static condition, participants had an average mean of 0.87° and an average standard deviation of 6.2°. In the motion parallax condition, the average mean was 3.5° and the average standard deviation was 3.1°. The difference between the standard deviations in the two conditions was found to be statistically significant ($p=0.02$).

DOES VISUALIZING LOGICAL RELATIONSHIPS HELP INHIBIT BIASED JUDGMENTS? [2.20]

Michael Truong, Thanujeni Pathman, & W. Dale Stevens, York University

Visualizations are frequently espoused as an effective way to communicate findings and demonstrate logical relationships. For example, early in the COVID-19 pandemic, visualizations of 'flattening the curve' were widely circulated to communicate the importance of slowing the spread of the virus to reduce its impact. Here, we investigate the importance of visualizations in educating people to inhibit their biased judgments by comparing the effectiveness of matched learning materials with and without Venn diagrams. Specifically, we test the effects of our training materials on people's judgments on the Linda problem, a notorious decision-making problem where many people make biased judgments violating the rules of logic through their use of the representativeness heuristic. We also investigate whether training effects are maintained when participants are not cued to apply their learning, and the role of individual differences to learning and decision making. Cues to apply learning were manipulated between-subjects through the presence or absence of a prompt that what they learned is related to the upcoming problem—the Linda problem was embedded within a battery of other non-learning related heuristics & biases questions to conceal the purpose of the learning material. We used two measures of individual differences known to be predictive of performance on decision-making problems – reflectivity and cognitive ability – and one exploratory measure of memory that we hypothesize to predict learning ability. Our findings may inform future educators and policy makers about whether visualizations are effective in reducing biases in people's judgments and whether the effectiveness varies based on individual differences.

UPDATE ON CURRENT RESEARCH: MULTIMODAL INTEGRATIONS OF ARTIFICIAL INTELLIGENCES ACROSS MIXED REALITY EXPERIENCES [1.27]

Sarah Vollmer, York University

Applications of augmenting a variety of lived experiences through artificial intelligence models are explored. Connections are made across several interdisciplinary projects developed over the past year. Common themes include natural language processing, computer vision, virtual and augmented reality, edge computing, and IoT(E) wearables and sensors.

ATTENTION STRATEGIES FOR LEARNING UNDER REDUCIBLE AND IRREDUCIBLE UNCERTAINTY [1.21]

Marcus R. Watson, Mazyar Fallah, & Thilo Womelsdorf, York University

In feature learning, uncertainty about feature values is reduced. Selective attention can help this, implying that agents should focus attention more under greater expected uncertainty about action outcomes. Little work tests this "attention-for-learning" prediction, and in particular it is unknown whether attention-for-learning is sensitive to the degree to which uncertainty can actually be reduced. Here we tested the attention-for-learning hypothesis in a naturalistic learning task that manipulated both reducible and irreducible forms of uncertainty, and quantified the strength of selective attention using attention-augmented reinforcement learning (RL) models. Human participants performed a 2-AFC object selection task in which multidimensional objects with a particular feature were more likely to be rewarded. Reducible uncertainty was manipulated between blocks by having objects vary along either two or five possible feature dimensions (different arms, body shapes, patterns, textures, or colors). Irreducible uncertainty took the form of different reward probabilities, either 0.70 or 0.85. As expected, when either form of uncertainty was higher, response times were longer, learning was slower, and asymptotic performance was lower. On blocks where one form of uncertainty was high and the other was low, these performance measures did not differ. However model results show that this similar performance was the result of different mechanisms. Specifically, when reducible uncertainty was high and irreducible uncertainty was low, participants had narrower attentional focus and greater exploratory biases than in the opposite condition. These results demonstrate that attention flexibly adjusts to the specific type of decision uncertainty. When faced with high levels of reducible uncertainty, attention becomes more focused and exploration increases, but the reverse is true for irreducible uncertainty, even when the resulting behaviour is highly similar. Taken together, these findings provide quantitative evidence for flexible adjustment of attention during learning to specific types of experienced uncertainty.

(UN)CONDITIONAL STEREO VIEW SYNTHESIS [3.17]

Jason Yu, Marcus Brubaker, & Kosta Derpanis, York University

How can general knowledge about scenes be represented? Great advances have been made in representing individual scenes, but methods for representing distributions (priors) over scenes are not as well developed. Generative modelling is a long standing problem, and modern methods are capable of producing high quality images. In our work, we seek to leverage the advances in generative modelling to represent priors over scenes as stereo image pairs. Specifically, we use diffusion models; a powerful class of generative models based on a probabilistic denoising process. This class of generative models not only allows for the generation of new and unseen stereo image pairs, but allows for conditional generation of a second view, given a single real image. We experiment with synthetic datasets generated using computer graphics (CLEVR) to test our proof of concept. We also explore a new method of generating consistent sets of views with more than two views, only using a stereo model. Our current proof of concept is promising, and we plan to extend the model to operate over unverified multi-view image sets.

A NOVEL TECHNIQUE FOR ANALYZING DYNAMIC FUNCTIONAL CONNECTIVITY OF LARGE-SCALE FUNCTIONAL BRAIN NETWORKS [2.21]

Amir Zarie & W. Dale Stevens, York University

Resting-state functional connectivity (RSFC) analysis of functional magnetic resonance imaging (fMRI) data involves correlating spontaneous fluctuations of the blood oxygen level-dependent (BOLD) signal from various brain regions during rest. It has been shown that RSFC can predict individual differences in the ability to process information in the visual domain. However, there are limitations associated with the traditional approach used for RSFC. Notably, the boundaries defining each brain network are assumed to be the same for all participants, incorrectly assuming that the spatial topography of functional networks are the same across individuals. Furthermore, this method of conducting FC assumes that the measures are stationary and do not fluctuate over time. Alternatively, dynamic FC (DFC), quantifies how FC changes over time, thus providing a finer grained picture of FC. Studies have shown that a significantly higher proportion of individual variability in measures of visuo-spatial reasoning is accounted for by DFC compared to standard stationary FC analysis. However, typical DFC techniques still fail to account for individual differences in network topology; for example, the commonly used software "Group ICA of fMRI Toolbox" (GIFT) uses group independent component analysis (ICA) to create a single brain network for all participants. As such, it is insensitive to individual variability. The aim of the current study was to develop and validate a novel method for analyzing DFC that accounts for individual variability in network topology. The technique Group Prior Individualized Parcellation (GPIP) was used to accomplish this. DFC was conducted using both GPIP and GIFT for comparison. There was a notable difference between the two approaches, indicating that taking individual differences in brain networks into account can impact measures of DFC. This study provided an early proof-of-concept for a more flexible and precise way to conduct RSFC analysis. Further research will be done to validate this technique using a larger sample of participants.

NOTES

Contribute to our Kudoboard celebrating former CVR director, Laurence Harris:
https://www.kudoboard.com/boards/G2QLwpBH

Thank you to our sponsoring partners.

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

Meta

Qualcomm

AMD

UBISOFT
TORONTO

CrossWing

CLOUD
CONSTABLE

Thank you also to the Faculty of Health, Faculty of Science, Faculty of Liberal Arts & Professional Studies (LA&PS), School of the Arts, Media, Performance & Design (AMPD), and Lassonde School of Engineering.