The Role of Attention in Predictive Visual Motion Processing

Submitted in total fulfillment for the jointly awarded degree of Doctor of Philosophy

Faculty of Medicine, Dentistry and Health Sciences
Melbourne School of Psychological Sciences
The University of Melbourne

&

Faculty of Human Sciences
The University of Cologne

by

Jeong In (Jane) Yook
ORCID: 0000-0002-0492-6455

February 2024

This dissertation was accepted by the Faculty of Human Sciences of the University of Cologne in July 2024.

The Role of Attention in Predictive Visual Motion Processing

Inauguraldissertation zur Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät
der Universität zu Köln
nach der Promotionsordnung vom 18.12.2018

&

Melbourne School of Psychological Sciences
The University of Melbourne

vorgelegt von

Jeong In (Jane) Yook
aus Südkorea

Februari 2024

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im Juli 2024 angenommen.

Abstract

Our visual world presents us with an abundance of sensory information, yet only a subset can be processed due to the inherent capacity limitations on attention and perception. Prediction and attention are two cognitive processes that assist in making sense of dynamic information: prediction utilizes prior information to guide the processing of new sensory information, while attention selectively enhances the processing of information relevant to behavior. This thesis aimed to explore the relationship between prediction and attention in shaping our perception of moving objects, using a combination of psychophysics and electrophysiological approaches coupled with attentional manipulations.

In **Chapter 2**, we explored how predictions manifest in perception through the flash-lag effect, where a static stimulus appears to lag a behind a continuously moving object despite being physically aligned. We show that this perceptual discrepancy arises from the brain's extrapolation of motion to compensate for neural processing delays. Additionally, we found that the magnitude of the illusion systematically varies with the perceived speed of motion, underscoring the explicit representation of velocity in predictively encoding the position of moving objects.

Following this, in **Chapter 3**, we investigated how attention influences predictions in dynamic environments involving multiple moving objects. We revealed that divided attention among multiple stimuli leads to a stronger flash-lag effect compared to when attention is focused on a single stimulus. This suggests that when the task demands monitoring of multiple moving objects simultaneously, attention serves to guide extrapolation to efficiently compensate for accumulated delays in processing these objects with limited attentional resources.

Finally, in **Chapter 4**, we employed multivariate pattern analyses on EEG data to examine the neural evidence of motion extrapolation within competing stimulus streams. We discovered that when a stream is attended, sensory-like representations of a stimulus' future position can be driven solely by prediction. Conversely, when the stream is ignored, the representation of its future position is consistent only with stimulus presentation. This implies that attention allocates all processing resources towards anticipating the future positions of task-relevant objects, thereby enhancing perception and behavioral performance while suppressing distractors.

Taken together, this research illuminates how attention may be strategically allocated to guide extrapolation to achieve an efficient monitoring of multiple moving stimuli or accurate encoding of task-relevant stimuli. This serves to facilitate successful navigation within dynamic visual environments. I discuss the relationship between prediction and attention within a comprehensive framework consistent with predictive coding theories.

Deutsche Zusammenfassung

Unsere visuelle Welt beinhaltet eine Fülle von sensorischen Informationen. Aufgrund von Kapazitätsbeschränkungen von Aufmerksamkeit und Wahrnehmung kann jedoch nur eine Teilmenge verarbeitet werden. Vorhersagen und Aufmerksamkeit sind zwei kognitive Prozesse, die dabei helfen, dynamische Informationen effizient zu verarbeiten: Vorhersagen steuern die Verarbeitung neuer sensorischer Informationen aus Basis vorheriger Informationen, während Aufmerksamkeit selektiv die Verarbeitung verhaltensrelevanter Informationen fördert. Ziel dieser Dissertation war es, die Beziehung zwischen Vorhersage- und Aufmerksamkeitsprozessen bei der Wahrnehmung bewegenden Objekten zu erforschen, wobei eine Kombination aus psychophysikalischen und elektrophysiologischen Ansätzen in Verbindung mit Aufmerksamkeitsmanipulationen verwendet wurde.

In **Kapitel 2** haben wir untersucht, wie sich Vorhersagen in der Wahrnehmung im sogenannten „Flash-Lag-Effekt“ manifestieren, bei dem ein statischer Stimulus einem sich kontinuierlich bewegenden Objekt hinterherzuhinken scheint, obwohl er physikalisch an einer äquivalenten Position präsentieret wird. Wir zeigen, dass diese Wahrnehmungsdiskrepanz durch Extrapolation von Bewegung durch das Gehirn Kompensation von Verzögerungen in der neuronalen Verarbeitung erklärbar ist. Darüber hinaus haben wir festgestellt, dass das Ausmaß der Illusion systematisch mit der wahrgenommenen Bewegungsgeschwindigkeit variiert, was für eine explizite Repräsentation der Geschwindigkeit bei der prädiktiven Kodierung der Position von sich bewegenden Objekten spricht.

Anschließend haben wir in **Kapitel 3** untersucht, wie Aufmerksamkeit die Vorhersagen in dynamischen Umgebungen mit mehreren sich bewegenden Objekten beeinflusst. Hier zeigte sich, dass die Verteilung der Aufmerksamkeit auf mehrere Stimuli im Vergleich zu einer fokussierten Aufmerksamkeit auf einen Reiz zu einem stärkeren Flash-Lag-Effekt führt. Dies deutet darauf hin, dass bei gleichzeitiger Beobachtung mehrerer sich bewegender Objekte Aufmerksamkeit dazu dient, die Extrapolation zu steuern, um die akkumulierten Verzögerungen bei der Verarbeitung dieser Objekte effizient zu kompensieren.

In **Kapitel 4** schließlich untersuchten wir mit Hilfe multivariater Musteranalysen von EEG-Daten die neuronale Evidenz der Bewegungsextrapolation innerhalb konkurrierender Reizübertragungen. Wir entdeckten, dass die sensorische Repräsentation der zukünftigen Position eines Stimulus allein durch Vorhersage gesteuert werden kann, wenn dieser Reizübertragung beachtet wird. Wird der Reizübertragung hingegen ignoriert, ist die Repräsentation seiner zukünftigen Position nur mit der Stimuluspräsentation möglich. Dies bedeutet, dass Aufmerksamkeit alle Verarbeitungsressourcen auf die Vorhersage der zukünftigen Position aufgabenrelevanter Objekte richtet und dadurch die Wahrnehmung und die Verhaltensleistung verbessert, während Ablenkungen unterdrückt werden.

Zusammengenommen beleuchten dies Forschungsergebnisse, wie Aufmerksamkeit strategisch ausgerichtet wird, um Extrapolation zu steuern und eine effiziente Beobachtung mehrerer sich bewegender Reize oder eine genaue Kodierung aufgaben relevanten Reize zu erreichen. Dies erleichtert die erfolgreiche Navigation in dynamischen visuellen Umgebungen. Ich diskutiere die Beziehung zwischen Vorhersage- und Aufmerksamkeitsprozessen in Rahmen von „predictive coding“ Theorien.

Declaration of Authorship

I hereby declare that,

a) this thesis comprises only my original work towards the degree of Doctor of Philosophy, except where indicated in the preface;
b) due acknowledgment has been made within the text for all other materials used; and
c) this thesis is fewer than 100,000 words in length, excluding figures, tables, footnotes, references, and appendices.

23.02.2024, Jeong In (Jane) Yook

Preface

This thesis includes two manuscripts in preparation for publication and one published article. All data presented in this thesis were collected by the PhD candidate, with assistance from members of the Time in Brain and Behavior Laboratory at the University of Melbourne. The first drafts of all chapters were written by the PhD candidate and subsequently revised and edited in collaboration with the named co-authors. All co-authors have agreed to use the manuscripts in this thesis and have provided signed copies of the co-author authorization form. The work in this thesis was supported by two Australian Research Council grants awarded to Hinze Hogendoorn (DP180102268 and FT200100246) as well as support from Forschungszentrum Jülich. Additionally, the PhD candidate was a recipient of the Melbourne Research Scholarship.

In the course of completing this thesis, it would be remiss not to acknowledge the profound impact of the global COVID-19 pandemic. Specifically, the pandemic significantly disrupted various aspects of research, including research operations, access to resources, and collaborative work environments, leading to long-term delays throughout the majority of the PhD duration. These repercussions further necessitated a reassessment of the project's scope and feasibility. Under the Jülich-University of Melbourne Postgraduate Academy (JUMPA) program, the PhD candidate was required to spend at least 12 months at both the home institution, the University of Melbourne, and the host institution, Forschungszentrum Jülich. Regrettably, due to these constraints, a decision was made to discontinue one major study within the project. The discontinued study aimed to investigate the neural networks underlying the predictive and attentional mechanisms as an extension of **Chapter 3**, using an fMRI approach called dynamic causal modeling. This adjustment was necessary to accommodate both potential and realized challenges encountered during the completion of this thesis and to ensure timely progress despite the imposed constraints.

Acknowledgements

This thesis would not have been possible without the guidance and help of several individuals, and it is a pleasure to thank those who made it happen. Firstly, I would like to thank my supervisors at the University of Melbourne, Hinze Hogendoorn, and at Forschungszentrum Jülich, Simone Vossel and Ralph Weidner. Your supervision, support, and expertise have been instrumental in shaping my PhD over the last four years. It was a very challenging, and rewarding, time working on problems aligned with your joint interests.

I extend my appreciation to my committee chairs, Patrick Goodbourn and Amy Jordan, for their kindness, reliability, and guidance throughout the challenges during my candidature. Special thanks to the JUMPA coordinators, Jessica Piesse, Petra Van Nieuwenhoven, and Maria Han, for being a crucial link to Melbourne.

I am grateful to my collaborators, Shiva Viswanathan, Philippa Johnson, and Will Turner for sharing their knowledge of all things electroencephalographic. My heartfelt thanks to Shiva and Philippa, who always found time to provide advice. Shiva, your rigorous scientific approach and attention to detail have enabled me to become a much-improved scientist and communicator. Philippa, your encouragement has inspired me to get to the end. I also could not have completed my experiments without the contributions of Lysha Lee, Mackenzie Murphy, and Vinay Mepani. Thank you for your invaluable assistance!

I am thankful to a few special people who have been a pillar of strength throughout this PhD. Jennifer, Misha, Chareesa, Kelly, and Eun Ah, thank you for your understanding and for keeping me sane throughout many mental breakdowns. Warm thanks to Helen, Yiu Hong, Sebastian, Yu Qian, Annika, and Lea for the memories beyond the INM-3 walls. To Hanbit friends, DocTeam, all my office mates, my yoga girlies, your support is truly appreciated. Last but not least, I would like to express my deepest gratitude to my family for their unwavering love and support in everything I do.

List of Publications During Candidature

Peer-reviewed Publications

Yook, J., Hogendoorn, H., Fink, G.R., Vossel, S., & Weidner, R. (in press). When visual attention is divided in the flash-lag effect. *Journal of Vision.*

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2022). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. *Vision Research.*

Feuerriegel, D., **Yook, J**., Quek, G. L., Hogendoorn, H., & Bode, S. (2020). Visual mismatch responses index surprise signalling but not expectation suppression. *Cortex*.

Submitted or In Preparation

Yook, J., Viswanathan S., Turner, W., Johnson, P., Weidner, R., Vossel, S., & Hogendoorn, H. (in prep). Meaningful objects: Selective visual attention guides predictive motion extrapolation. *Unpublished material not yet submitted for publication.*

Feuerriegel, D., Lilburn, S., **Yook, J**., Garrett, P., Smith, P., & Bode, S. (in prep). Neural correlates of evidence accumulation during continuous report decisions. *Unpublished material not yet submitted for publication.*

Conference Abstracts

Yook, J., Turner, W., Johnson, P., Weidner, R., Vossel, S., & Hogendoorn, H. (2023). Prediction in the visual perception of moving objects: the role of attention. Invited symposium talk given at the 65th *Tagung experimentell arbeitender Psycholog:innen; Conference of Experimental Psychologists (TeaP)*, Trier, Germany.

Yook, J., Turner, W., Johnson, P., Weidner, R., Vossel, S., & Hogendoorn, H. (2022). Predictive activation of neural position representations for moving objects with and without visual attention. Poster presented at *European Conference of Visual Perception (ECVP),* Nijmegen, The Netherlands.

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2021). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Poster presented at the *INM-3 Retreat,* Jülich, Germany.

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2020). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Talk given at *Neuromatch Conference 3.0* (online).

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2020). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Talk given at the *MDHS Graduate Research Conference,* Melbourne, Australia (online).

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2020). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Poster presented at the *UNSW workshop on Expectation, Perception, and Cognition*, Sydney, Australia (online).

Table of Contents

List of Tables 1

List of Figures 2

List of Abbreviations 4

Chapter 1 **General Introduction** 7

Visual processing 8

Compensating for neural delays 11

Attention and perception 16

Attention and prediction 20

Overview of methodology 21

Overview of empirical studies 24

Chapter 2 **Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed** 29

Chapter 3 **When visual attention is divided in the flash-lag effect** 59

Chapter 4 **Meaningful objects: Selective visual attention guides predictive motion extrapolation** 87

Chapter 5 **General Discussion** 135

Summary of empirical findings 136

Theoretical implications 137

Reconciling effects of attention 141

Altered attention states 142

Limitations 143

Conclusion 144

References 146

Author Contributions 167

Curriculum Vitae 169

Lebenslauf 173

List of Tables

Chapter 3

Table 1 Paired-sample *t*-tests (two-tailed) with correction for multiple comparisons for mean FLE magnitude 70
Table S1.2 One-sample t-test (one-tailed) results against zero 78
Table S2 One-sample t-test (one-tailed) results against zero 79

Chapter 4

Table 1 Results of a three-way mixed ANOVA for diagonal-averaged bias values 113
Table S2 Results of one-sample *t*-tests (one-tailed) for the time-agnostic peak CV accuracy values 120
Table S4 Results of one-sample *t*-tests (one-tailed) for the cross-condition classification accuracy values 122

List of Figures

Chapter 2

Figure 1	Schematic illustration of the stimulus configuration and flash-lag sequence in Experiment 1A	36
Figure 2	Schematic illustration of Experiment 1B	38
Figure 3	Results of Experiment 1	39
Figure 4	Change in FLE magnitude plotted as a function of change in perceived speed for (A) filtered noise and (B) pixelated noise textures	40
Figure 5	Results of Experiment 2A	44
Figure 6	Results of Experiment 2B	45

Chapter 3

Figure 1	Experimental paradigm	65
Figure 2	Individual proportions of fixation time across various phases of the experiment (cue, orient, and target), collapsed across attention conditions	70
Figure 3	FLE results	71
Figure S1.1	Individual distributions of FLE magnitudes across attention conditions	77
Figure S3	Hemifield results	80

Chapter 4

Figure 1	Stimuli and experimental paradigm	92
Figure 2	First-level analysis pipeline	99
Figure 3	Position-specific neural responses to single flashes	105
Figure 4	Position classification results of single flashes	106
Figure 5	Cross-condition classification results of apparent motion	108
Figure 6	Main results	110
Figure 7	Temporal pairing results	112
Figure S3	Spatial generalization results	121
Figure S5	Interaction effects	123

Figure S6.1 Replicated main results 125

Figure S6.2 Replicated temporal pairing results 126

List of Abbreviations

ANOVA	Analysis of variance
CV	Coefficient of variation (Chapter 3)
CV	Cross-validation (Chapter 4)
dva	Degrees of visual angle
EEG	Electroencephalography
FLE	Flash-lag effect
fMRI	Functional magnetic resonance imaging
FWE	Familywise error rate
hMT+	Human medial temporal area complex (or visual area V5)
Hz	Hertz
ISI	Inter-stimulus interval
ITI	Inter-trial interval
LGN	Lateral geniculate nucleus
M	Mean
max	Maximum
min	Minimum
mins	Minutes
MOT	Multiple object tracking
ms	Millisecond
MST	Medial superior temporal area
MT	Macaque medial temporal area
MVPA	Multivariate pattern analysis
p	p-value
PSE	Point of subjective equality
RM	Representational momentum
ROI	Region of interest
s	Seconds

SD	Standard deviation
SEM	Standard error of the mean
SOA	Stimulus onset asynchrony
TMS	Transcranial magnetic stimulation

CHAPTER 1

General Introduction

Visual perception is much more complex than what is often imaged on the eye. We are constantly exposed to streams of information, such as the bustling activity of a crowded street, passing cars, and changing lights, which shape how we perceive our visual world. The visual system is adept at processing these changes that unfold before us, allowing us to navigate through complex stimuli. At the heart of this experience is the brain's ability to make sense of dynamic information. Selecting the appropriate subset of information to process is crucial, especially in rapidly changing, dynamic environments. For example, selective attention allows us to focus on features that are relevant to behavior. Our ability to monitor the color of a traffic light while approaching an intersection, while ignoring the shape or size of the signal, exemplifies that attention serves as a filter for specific features of the environment while ignoring others. Another critical mechanism involves prediction, anticipating future events based on prior knowledge of the world. In dynamic environments, probabilistic regularities enable an anticipation of when or where task-relevant information might occur, for example, seeing the traffic lights change from green to yellow is a predictor of an impending red color. These regularities can be exploited to allocate attentional resources in anticipation of relevant information. At the same time, probabilistic regularities could involve seemingly task-irrelevant features – suppose even if a traffic light's color is the key feature of relevance, at a busy intersection the spatial location of the traffic light indicates whether the displayed color is relevant to the observer or not (to drivers in other lanes). In this thesis, I will explore how prediction and attention influence perception, with a focus on the role of attention in predictive motion processing in the visual system.

This introductory chapter is divided into two parts. The first part provides an overview of visual processing and the challenges posed by visual processing delays. It also explores how the brain overcomes the problem of neural delays to perceive dynamic stimuli, such as moving objects, in real time. The second part provides a concise literature review of attention, a crucial top-down mechanism in visual perception.

Visual processing

The anatomical organization of the visual system follows a hierarchical structure (Hubel & Wiesel, 1968). Visual information is initially captured by photoreceptor cells in the retina. Through the optic nerve, visual information from both eyes is integrated before reaching the lateral geniculate nucleus (LGN) of the thalamus. From the LGN, this information is propagated posteriorly, where the processing hierarchy begins within different areas of the occipital lobe. The projection from the eye to the brain is organized into several parallel pathways, each serving distinct submodalities of visual perception (Wässle, 2004). Notably, the magnocellular (M) pathway is primarily involved in motion detection, while the parvocellular (P) pathway is responsible for high-acuity vision, color perception, and form recognition (Masri et al., 2020). Throughout the visual pathway, a consistent retinotopic organization is observed, with cells of the optic nerve projecting to their respective M and P layers of the LGN before continuing up to the primary visual cortex (V1) (see also Ungerleider & Haxby, 1994).

This arrangement is maintained in V1, where the receptive field of each neuron encodes visual information. Receptive fields denote specific regions of the visual field to which neurons respond. These fields vary in size and exhibit spatial selectivity to local motion signals, meaning neurons are more responsive to movement within their immediate vicinity. This spatial selectivity allows neurons to encode motion patterns in V1. Additionally, neurons with nearby receptive fields are represented closer together, creating a spatial mapping of the visual space (Braddick, 2001).

Moving beyond V1, the visual pathway divides into two streams: the dorsal stream, which carries information to the parietal lobe, and the ventral stream, which extends to the temporal lobe. The dorsal stream follows the M pathway, responsible for processing spatial and motion information (“where” pathway), and the ventral stream follows the P pathway, involved in object recognition (“what” pathway) (Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982).

As visual information travels along the ventral stream, it undergoes increasingly complex processing as it propagates to higher-level cortical areas including the inferior temporal (IT) cortex. While V1 encodes basic features such as color, orientation, and retinotopic position (Hubel & Wiesel, 1968), neurons in these areas have larger receptive fields and contribute to the integration of more semantic features (Castelo-Branco et al., 2002), as well as the formation of abstract representations of objects that are robust to changes in rotation, contrast, and position (Mishkin & Ungerleider, 1982; Tanaka, 1996).

Simultaneously within the dorsal stream, the medial temporal lobe specializes in motion perception (Grill-Spector & Malach, 2004; Zeki, 1974; Zeki et al., 1991). For primates, this region is known as area MT. In humans, the hMT+ complex, or V5, is a homolog of the macaque area MT and the adjacent medial superior temporal (MST) area, located on the lateral surface at the occipital-temporal boundary (Zeki et al., 1991). hMT+/V5 receives input from earlier visual areas V1, V2, and V3, as well as subcortical structures like the superior colliculus and the pulvinar. Additionally, it has connections with various parts in the superior temporal lobe, parietal lobe, and frontal lobe, which indicates its involvement in higher-level visual processing such as the analysis of optic flow and motor planning (Culham et al., 2001). Motion-sensitive neurons in V1 tend to have small receptive fields, limiting their detection of motion. MT neurons have significantly larger receptive fields (~10 times), allowing them to integrate information from multiple local motion signals to perceive global motion patterns (Simoncelli & Heeger, 1998). This difference in receptive field size enables approximately 90% of MT neurons to be tuned to the direction and speed of moving objects, compared to the 25% of direction-selective neurons found in V1 (Albright, 1984; Hubel & Wiesel, 1968; Maunsell & Van Essen, 1983).

Therefore, hMT+ is particularly responsive to dynamic stimuli, including flickering checkerboards, moving gratings, moving objects, and moving dot patterns (Culham et al., 1998). It is also activated during visual motion memory and imagery tasks (Cohen et al., 1996; Zaksas & Pasternak, 2006). hMT+ neurons show a preference for coherent over incoherent motion (Rees et al., 2000), although some studies observe the contrary (L. M. Harrison et al., 2007; McKeefry et al., 1997). Conversely, systematic modulation of MT activity can alter motion perception. For instance, studies employing single-pulse transcranial magnetic stimulation (TMS) to hMT+/V5 have demonstrated impaired judgments of motion direction and perception of moving stimuli (Laycock et al., 2007; Maus, Ward, et al., 2013).

Visual processing has traditionally been considered a feedforward process, driven by the hierarchical architecture of the visual pathways outlined earlier. This structure allows incoming sensory input to travel up the processing hierarchy, where visual information is processed and interpreted for conscious perception in a bottom-up manner. However, due to the noisy and dynamic nature of visual scenes, relying solely on this feedforward sweep of information is insufficient to keep up with the constantly changing environment. To address this, top-down processes are required to help make sense of ambiguous sensory information (Friston, 2005).

In addition to feedforward processing, visual processing also involves bidirectional interactions (Lamme et al., 1998). Horizontal connections convey information between neurons within the same processing level, while recurrent feedback connections extend from hierarchically later visual areas to earlier visual areas (Lamme & Roelfsema, 2000). These connections play an important role in ensuring a coherent representation of visual stimuli throughout the processing hierarchy (DeYoe & Van Essen, 1988; Wyatte et al., 2014). For instance, a substantial amount of information from the LGN bypasses V1 and directly reaches hMT+/V5 (Laycock et al., 2007), allowing representations to become available earlier in the processing stream and influence subsequent processing through feedback connections to V1. In line with this, there is growing evidence suggesting that horizontal and feedback signals can activate neural responses (Benvenuti et al., 2020; Kok et al., 2016). Transient stimulations of hMT+/V5 followed by V1 have been shown to induce perceptions of motion initiation, indicating that feedback connections between hMT+/V5 and V1 are involved in shaping our perception of motion (Silvanto et al., 2005).

Neural processing delays

How long does perceptual processing take? The duration of visual processing can vary depending on the complexity of the stimulus and the pathways involved. Typically, feedforward processing requires approximately 30–60 ms from the moment visual information enters the visual system through the eyes (Lamme & Roelfsema, 2000; Maunsell & Gibson, 1992; Schmolesky et al., 1998). In macaques, responses to visual stimuli have been recorded approximately after 60–100 ms (Schmolesky et al., 1998), with a delay of around 10 ms between V1 and MT (Nowak & Bullier, 1997). Median latency in V1 is 45 ms for flash stimuli (Maunsell & Gibson, 1992) and 85 ms for moving stimuli (Raiguel et al., 1999), with a mean latency of 72 ms in V1 and 76 ms in MT (Lamme & Roelfsema, 2000). Processing latencies are typically longer in humans, and the earliest emergence of a visual percept is approximately 150 ms (Thorpe et al., 1996). This implies that neural transmission alone introduces a delay of at least ~100 ms in the initial feedforward sweep.

Despite the rapid processing of the visual system, the problem of neural delays becomes apparent when perceiving moving objects. The time it takes for object motion to be processed means that by the time we perceive its position, it has likely shifted to a new location. This creates a discrepancy between events in the world and the representations of these events in our internal world, leading to inaccuracies in perception because it is out of date.

The impact of neural delays can be illustrated by the challenge faced by athletes in intercepting a moving ball, as commonly observed in sports like tennis and cricket (Nijhawan, 1994, 2008). Imagine a scenario where a cricket player bowls a cricket ball with an average velocity of 90 mph. It takes about 100 ms for the batter to perceive the ball's position. According to Nijhawan (2008), this results in a lag of 13.2 ft (4 m) between the perceived and actual ball position at any given moment. Such a discrepancy means that the perceived ball position lags significantly behind its veridical location, requiring the batter to intercept it before seeing the bat make contact with the ball.

Can the batter effectively play a shot? Relying on outdated information from 100 ms ago would make it impractical for the batter to execute a hit accurately. Importantly, despite this, we can localize and interact with moving objects remarkably well, which suggests that the brain has developed mechanisms to compensate for its neural delays.

Compensating for neural delays

In the following section, I will provide a concise overview of the theoretical framework involving predictive coding and motion extrapolation to address neural delays during motion perception. Additionally, I will outline some of the behavioral and neural evidence demonstrating compensation for neural delays within the visual system, exploring potential mechanisms for implementation.

Predictive coding

An influential theory of perception posits that perception is a result of the interaction between sensory input and our internal model of the world (Friston, 2003; Rao & Ballard, 1999; Walsh et al., 2020). Predictive coding models, which are rooted in Bayesian inference, propose that before receiving any sensory input the brain constantly generates beliefs about what it might perceive based on its prior knowledge of the world. These beliefs are represented by a prior probability distribution, and the model generates predictions about what it expects to perceive, guiding the interpretation of forthcoming sensory data (Friston, 2005; Knill & Pouget, 2004). When processing sensory information, the brain combines its predictions with actual sensory data to form an updated model of the world, represented as a posterior probability distribution. If predictions match the sensory input, existing predictions are reinforced. However, if there is a mismatch, the difference between predicted and actual input is represented as a prediction error, which is used to update the internal generative model and refine new predictions (Friston, 2003).

According to predictive coding theories, perception is the best guess of this inference process, where bottom-up sensory signals and top-down prediction signals converge through recurrent feedforward/feedback connections across the visual cortex (Clark, 2013; Rahnev, 2019). This mechanism ensures that neural resources are not wasted on encoding redundant or predictable information when expectations are fulfilled (Rao & Ballard, 1999). Consequently, the iterative process enables us to efficiently maintain up-to-date representations of the world across various levels of visual processing hierarchy and reduce prediction errors over time (Clark, 2013; Friston, 2010; Huang & Rao, 2011; Knill & Pouget, 2004).

In essence, predictive coding and Bayesian inference complement each other (Aitchison & Lengyel, 2017). While predictive coding focuses on the neural representation of prediction errors, it does not specify how predictions are generated and how prediction errors are used. Bayesian inference provides a mathematical framework for computing these signals without specifying the underlying representations across the processing hierarchy (Walsh et al., 2020).

One caveat of classical predictive coding models is their inability to account for temporal predictions (Hogendoorn & Burkitt, 2019). In the case of a time-varying stimulus, sensory information continues to evolve during the inference process, while predictions become progressively outdated at each processing stage. This ongoing mismatch between predictions and incoming sensory input may result in a continuous cycle of prediction errors, preventing the convergence of prediction and prediction error signals (Hogendoorn & Burkitt, 2019). To resolve this issue, the model must be predictive *about the future,* as it is predictive about hierarchical activations. Several dynamic predictive coding models integrate time, each differing in their neural implementations (Friston, 2008; Hogendoorn & Burkitt, 2019; Jiang & Rao, 2023; Kutschireiter et al., 2017; Millidge et al., 2023; Ororbia et al., 2020; Rao & Ballard, 1997). I will not delve into the details of these models within this thesis; however, the underlying idea is that specific temporal dynamics are integrated based on statistical regularities of the world. For example, neural representations of an object's position and its temporal derivatives, such as velocity, acceleration, and jerk can be used to predict its most probable state at the future time-step based on the current time-step.

Motion extrapolation

A specific application within this framework that addresses how neural delays are compensated during motion perception is motion extrapolation. Motion extrapolation refers to a mechanism that allows the brain to predict the future position of a predictably moving object, based on its past and current motion signals (Nijhawan, 1994). Essentially, it extends the current velocity of a moving object into the future to anticipate where it will be located in the future. Predictive motion extrapolation facilitates the visual system in maintaining a continuous and coherent neural representation of moving objects despite neural delays (Hogendoorn, 2020; Nijhawan, 2008).

Recent studies have integrated motion extrapolation into a Bayesian framework (e.g., Burkitt & Hogendoorn, 2021; Jiang & Rao, 2023; Khoei et al., 2017). Within this framework, a probabilistic representation of velocity is combined with a fixed delay (i.e., uncertainty associated with the sensory information), enabling the computation of the likely position of a moving object. This allows all levels of the processing hierarchy to represent the same position aligned in real-time, facilitating compensation of known neural delays at any given moment (Hogendoorn & Burkitt, 2019). These models have been simulated to explain several aspects of the flash-lag effect (see below) but have not been empirically tested yet. It remains an open question whether different levels of the processing hierarchy are actually aligned.

Besides addressing neural delays, motion extrapolation helps the brain maintain the perception of continuous motion even when sensory information is temporarily unavailable or ambiguous due to factors like occlusion or eye blinks (Teichmann et al., 2021). This aligns with extrapolation as a mechanism for filling in perceptual gaps during motion perception.

Evidence of predictive motion processing

Motion extrapolation has been argued to underlie various visual illusions that involve the displacement of objects due to their motion or the motion of nearby objects (Changizi et al., 2008). These illusions highlight the predictive encoding of the position of a moving object, leading to shifts in:

1. Representational momentum: This phenomenon describes the tendency to misperceive a moving object as stopping slightly ahead of its actual position (Freyd & Finke, 1984; Hubbard, 2005),

2. Fröhlich effect: This illusion occurs when the initial position of a moving object is misperceived ahead of its actual position (Fröhlich, 1923),

3. High-Phi illusion: In this illusion, a rotating texture is replaced by a new texture, leading to illusory forward jumps in its perceived position (Wexler et al., 2013),

4. Flash-lag effect: A static object flashed next to a moving object is perceived as lagging behind the moving object, despite being spatially aligned at the time of the flash (Mackay, 1958).

Among these illusions, the flash-lag effect (FLE; Mackay, 1958; Nijhawan, 1994) is particularly noteworthy. The dominant explanation of the FLE is that the visual system extrapolates the position of the moving object along its expected trajectory to compensate, at least partially, for neural delays. This results in the perception of an extrapolated position that is closer to its veridical position (Nijhawan, 2008; Nijhawan & Wu, 2009). In contrast, the flash cannot be extrapolated because there is no need for such compensation for time-invariant stimuli (Khurana & Nijhawan, 1995).

In animals, a neural correlate to the psychophysical effects of motion extrapolation has been observed in a study by Jancke et al. (2004). They recorded neurons in cat V1 (area 17, or human V1) while the cats viewed either briefly presented or moving squares of light in different directions and with different speeds. Using an optimal linear estimator, they reconstructed the position of each stimulus based on the pooled activity of neuronal populations. Notably, they found that the peak population activity of receptive fields shifted earlier in response to the moving stimulus compared to the flash, indicating that the processing latency was 16 ms faster for motion. In addition, the representation of the moving stimulus at slower speeds (4.5°/s) was closely aligned with its veridical location, suggesting full compensation. However, at higher speeds (> 8.8°/s), they noted a spatial discrepancy in the representation and the stimulus position, which increased with increasing stimulus speed.

From V1, functional magnetic resonance imaging (fMRI) studies have revealed similar patterns of neural activity across various areas involved in motion processing including V1-3 and hMT+/V5 in humans (Apthorp et al., 2013; Harvey & Dumoulin, 2016; Maus, Fischer, et al., 2013; M. Schneider et al., 2019).

More recently, Hogendoorn and Burkitt (2018) conducted an electroencephalographic (EEG) study where wedges were presented sequentially, either in a circular path to induce the percept of a single stimulus moving in apparent motion or flashing in a scrambled order. They evaluated latency differences between stimuli across two intervals: [80–90 ms] after stimulus onset, likely reflecting an early position representation, and [140–150 ms] after stimulus onset, indicative of a position representation originating in higher-level visual areas. During the latter interval, Hogendoorn and Burkitt observed that the representation of the moving wedge was available 15 ms earlier compared to the same wedge when it was flashed, a finding consistent with previously reported results in cat V1 (Jancke et al., 2004). Building upon these findings, Blom et al. (2021) demonstrated using the same stimuli that a latency advantage for the moving stimulus emerged from the second presentation within the motion sequence, or when the direction of motion was predictable (i.e., the first presentation has a 50:50 chance moving in either direction).

Neural mechanisms of predictive motion processing

The earliest stage at which compensation has been found is in the retina (B. Liu et al., 2021). Berry et al. (1991) conducted experiments where they recorded salamander and rabbit retinal ganglion cells while presenting a bar that either flashed or moved in a horizontal direction. They observed that when a stationary bar was presented, these cells responded with a 50 ms latency. By contrast, when the bar moved across the visual field, the peak in the firing rate response of retinal ganglion cells occurred before the peak response to the flash. This shift in firing pattern was mediated by a contrast-gain control mechanism, where a cell responds strongly when the stimulus enters its receptive field, followed by desensitization and reduces firing as it leaves the receptive field. This implies that the neural image transmitted from the retina to the brain is already extrapolated along its motion trajectory so the position of the stimulus is represented closer to its veridical position (Berry et al., 1999). This mechanism has also been shown in the thalamus and V1 (Johnston & Lagnado, 2015), which suggests that neural delays could be compensated throughout the visual system via a feedforward gain control mechanism.

Moving on to horizontal connections, Jancke et al. (2004) demonstrated a similar mechanism in cat V1. They recorded neurons responding to flashed or moving squares of light and observed shorter response latencies when the neurons were presented with the moving light compared to the flash. Specifically, the activation of neurons by the moving light also triggered a sub-threshold wave of activity that propagated horizontally to neighboring neurons across the visual cortex, known as activation spreading. While this activity did not reach firing thresholds to modulate neuronal activity, it served a critical function in priming neurons further along the motion trajectory to coordinate responses when the light arrived at their receptive fields (Jancke & Erlhagen, 2010; Müsseler et al., 2002). Similar facilitation has been demonstrated in macaque V1 (Benvenuti et al., 2020) as well as the retina (Souihel & Cessac, 2021), suggesting that neural delays could be compensated by anticipation of moving stimuli. However, it is important to note that faster processing of moving stimuli aligns more closely with a theory of differential processing latencies between moving and static objects (see Flash-lag effect), rather than predictive motion extrapolation.

Finally, feedback connections could facilitate a predictive mechanism to compensate for neural delays. Silvanto et al. (2005) conducted a series of TMS experiments where sequential stimulation over hMT+/V5 followed by V1 induced perception of moving phosphenes. Conversely, reversing the stimulation order disrupted the induced perception, indicating that feedback connections from hMT+/V5 significantly contribute to motion perception. To assess whether these disruptions in these feedback connections impair the ability to infer the position of moving stimuli through motion-based prediction, Maus, Ward, et al. (2013) administrated single-pulse TMS over hMT+/V5, V1, and V2 at different times of a flash-lag task. They found that TMS to hMT+/V5 diminished the FLE, while stimulations of V1 or V2 did not affect the perceived position of the moving stimulus relative to flashes. Furthermore, Wang et al. (2022) applied anodal and cathodal transcranial direct current stimulations to hMT+/V5 before observers performed a flash-lag task. These stimulations, which modulate cortical excitability in hMT+/V5, led to systematic enhancements and reductions in the FLE magnitude. These results suggest that compensatory mechanisms could be facilitated through predictive processes.

Attention and perception

Another process that shapes our perception is visual attention, as our capacity to process sensory information is limited (Norman & Bobrow, 1975). Selective attention serves an important function in enabling a preferential processing of sensory information that is most relevant to behavior, such that it is striking how little we perceive when we do not pay attention. Therefore, it is important to understand how attention prioritizes task-relevant information from streams of incoming visual information and various task demands.

The control of attention is commonly understood as bottom-up and top-down mechanisms. Bottom-up processes are driven by sensory input, leading to automatic and involuntary shifts in attention. They are triggered by salient or unexpected stimuli in the visual environment, regardless of their relevance to the current task (Theeuwes, 2010). In contrast, top-down processes are guided by the observers' goals, expectations, and prior knowledge, resulting in voluntary and selective allocation of attention. Top-down attention is directed towards specific stimuli or features that are task-relevant, which is the focus of this thesis.

In the following sections, I will provide an overview of top-down attention, including popular theoretical frameworks and the underlying mechanisms of selective attention in motion perception.

Spatial attention

Attention can be directed to specific locations in the visual field, enhancing the processing of all features at the attended location. A classic experimental paradigm used to study selective spatial attention is the Posner cueing paradigm (Posner, 1980). In this paradigm, probabilistic cues are used to guide attention to one of two likely locations where a target is expected to appear. Observers are instructed to detect the target as quickly as possible. Responses to targets presented at the cued location are reported to be faster and more accurate compared to targets presented at uncued locations (Posner, 1980). Several studies have shown that this behavioral facilitation is accompanied by strong neuronal responses when attention is directed to locations corresponding to their retinotopic position and reduced responses to unattended locations (Luck et al., 1997; Treue & Maunsell, 1996). Spatial attention has been found to enhance perceptual sensitivity to attended locations, thereby reducing interference from distractors (Luck et al., 1997), and improving spatial acuity or resolution (Carrasco, 2018) and trial-to-trial variability (Arazi et al., 2019).

Non-spatial attention

There is compelling evidence indicating that attention can be directed to specific features, even when they overlap in space. In the classical Stroop task, observers are presented with words printed in colored ink and asked to report either the color of the ink or the semantic meaning of the word (Jensen & Rohwer, 1966; Stroop, 1935). Importantly, the words often conflict with the ink color (e.g., the word “green” printed in red ink), and observers must selectively attend to the ink color while ignoring the word’s meaning, requiring them to suppress the automatic process of reading the word. Feature-based attention can also be directed to a non-spatial feature, such as stimulus color, shape, or motion, which enhances responses towards the attended feature across the entire visual field (Maunsell & Treue, 2006). This has also been shown to reduce neuronal response latencies and behavioral response times to task-relevant changes in the speed of moving stimuli (Galashan et al., 2013).

Finally, attention can also facilitate features that cannot be separated by location through object-based attention. When attention is directed to a specific feature of an object, it enhances all features constituting the object, even if they are task-irrelevant (O’Craven et al., 1999). In line with this, Ekman et al. (2020) demonstrated that when spatial attention is directed to a part of an object, the attentional modulation spreads to other locations that are part of the same object. These findings suggest that the visual system prioritizes processing visual stimuli as coherent entities rather than individual features or spatial locations (Cavanagh et al., 2023). Thus, the visual system must possess the ability to bind these features together. Pre-attentive object representations, however, fail to achieve this integration and remain as a collection of disjointed features.

Theories of selective attention

Theories of selective attention are useful for understanding why and how we prioritize only certain aspects of the visual environment over others. A notable concept in this regard is the attentional spotlight (James, 1950), which is analogous to the “Cocktail Party” effect observed in auditory perception (Cherry, 1953). Just as we can selectively focus on a single conversation amidst a noisy party, our attentional spotlight in vision enables us to focus on specific visual stimuli while filtering out irrelevant sensory information. However, the spotlight sometimes leads to overlooking unexpected stimuli or events when our attention is engaged elsewhere, as demonstrated in studies of change blindness (Simons & Chabris, 1999).

Hence, Broadbent's filter model represents one of the earliest bottleneck theories of attention. In this theory, attention acts as a selective filter at an early stage of processing, where it prioritizes incoming sensory information based on basic physical features such as location, direction, and color (Broadbent, 1958). Task-relevant information is allowed to proceed for higher-level processing involving memory and decision-making, while task-irrelevant information is completely filtered out. By focusing on the most relevant stimuli through the processing hierarchy, this theory effectively manages our limited attentional capacity. Treisman's attenuation model revises Broadbent's filter model by proposing that unattended information is not completely excluded from further processing but rather attenuated (Treisman, 1964). According to this theory, all incoming sensory information is processed to some extent, with the degree of processing determined by its task relevance. While unattended information may be weakly represented, it remains accessible, making it difficult but not impossible to extract information from task-irrelevant stimuli. Biased competition theory offers a different perspective on selective attention, where stimuli in the visual field compete for cortical representation and processing resources in the brain (Desimone & Duncan, 1995). The outcome of this competition is influenced by top-down factors such as goals and expectations. Task-relevant stimuli are prioritized for processing while irrelevant or distracting stimuli are suppressed or inhibited.

In summary, it is possible attention may operate from an early stage, so observers do not perceive unattended stimuli, or low-level processing occurs for all stimuli but higher-level integration is reserved for attended stimuli, making it difficult to perceive anything more than the physical features of unattended or ignored stimuli, or attended stimuli are processed at the expense of unattended stimuli.

Neural mechanisms of selective attention

Historically, the prevailing belief was that attention primarily influenced later processing stages, with significant attentional modulations observed in areas like MT (Treue & Maunsell, 1996, 1999), with areas such as V1 were thought to be minimally affected (Luck et al., 1997). However, subsequent studies challenged this belief by revealing differential activations of attended and unattended moving stimuli even in earlier visual areas V1 and V2 and areas closely connected to hMT+/V5 like the thalamus and superior colliculus (Büchel, 1997; Büchel et al., 1998; Culham et al., 1998). These findings suggest that attentional effects are not confined to later processing stages but manifest early and increasingly throughout the processing hierarchy, extending to various regions within the visual system (Cheong et al., 2012; Seidemann & Newsome, 1999).

Attentional modulation has been demonstrated to exert profound effects on the activity of individual neurons, influencing their firing rates and response properties throughout the visual cortex (Reynolds & Chelazzi, 2004; Treue, 2001). McAdams and Maunsell (1999) recorded single neurons in macaque V4 during an orientation-matching task involving Gabor and colored Gaussian stimuli, with one appearing on each side of the display. When attention was directed to one of the two locations, the overall responsiveness of orientation-selective neurons within the attended receptive field increased. This enhancement can be attributed to a multiplicative increase in neural responses along the tuning curve, strengthening the representation of the task-relevant information without sharpening or shifting their tuning curves towards the attended location (McAdams & Maunsell, 1999; Treue & Maunsell, 1996, 1999). Treue and Martínez Trujillo (1999) extended these findings to direction-selective neurons in MT.

Furthermore, several studies have demonstrated that spatial attention impacts the spatial properties of receptive fields such as size, shape, or location, independently of their response properties. For instance, when attention is allocated to one of multiple locations inside a neuron's receptive field, the receptive field shifts its center towards the attended location. Additionally, it is observed to shrink around the attended stimulus, effectively excluding unattended stimuli from within the receptive field (Moran & Desimone, 1985; Womelsdorf et al., 2006, 2008). Computational studies have hypothesized that these changes are due to the multiplicative gain of neural responses in early visual areas, such as V1 and V2 (Compte & Wang, 2006; Fox et al., 2023; Klein et al., 2014), suggesting that a response gain mechanism may underlie the neural correlate of the attentional spotlight.

In the case of feature-based attention, a non-multiplicative modulation of tuning curves has been shown to operate based on a feature similarity gain mechanism (Martinez-Trujillo & Treue, 2004). In experiments conducted by Maunsell and Treue (1996, 1999), two oppositely moving dots were presented within the receptive fields of direction-selective neurons. They consistently observed stronger neuronal responses for the preferred direction compared to the opposite direction, even when the attended stimulus was located outside of the receptive field. This suggests that attention enhances the responsiveness of neurons tuned to the attended feature (i.e., direction), regardless of whether the attended stimuli are within their receptive fields. Treue and Martínez Trujillo (1999) presented two random dot patterns within the same receptive fields, with one pattern moving in the anti-preferred direction and the other moving randomly in one of 12 possible directions. They found a response enhancement for the preferred direction in MT neurons accompanied by a response suppression for the anti-preferred direction, indicating that attention could modulate the processing of task-relevant information while suppressing irrelevant information.

Finally, object-based attention shares common neural mechanisms with feature-based and spatial attention but focuses on enhancing the processing of features within their spatial context (see also Cavanagh et al., 2023). In the context of motion, similar to spatial attention, object-based attention may also involve shifts in receptive field centers or changes in receptive field sizes to boost the neural representation or responses to the grouped features of the attended object including its speed, direction, and location. This would allow the receptive fields to align more closely with the motion trajectory of the attended object, to capture its motion-related features better. By adjusting receptive field properties, object-based attention may enhance the perception and tracking of moving objects and their changing positions over time.

So far, I have presented evidence of selective attention in conditions involving more than two stimuli, where visual attention is allocated to one of the stimulus locations or features. Divided attention is a related concept where attention is allocated to multiple stimuli simultaneously. However, fundamentally, divided attention shares common mechanisms with processes of selective attention (Hahn et al., 2008).

Attention and prediction

Many studies of attention reviewed here have reported heightened neural responses to stimuli at expected (i.e., attended) locations (Reynolds & Heeger, 2009), which seems contradictory to predictive coding theories, which propose that responses to expected stimuli should be suppressed (Summerfield & Egner, 2009). To reconcile this tension, recent predictive coding theories have proposed integrating attention and prediction within a common framework (Feldman & Friston, 2010).

This integration is premised on the understanding that sensory signals are inherently noisy and unreliable (see **Predictive coding**). To minimize the impact of unreliable sensory information (i.e., uncertainty) on our perception, the Bayesian inference process must factor in the reliability of sensory signals (Friston, 2009; Rao & Ballard, 1997). In predictive coding, sensory signals and prediction errors are functionally equivalent, as prediction errors represent unexplained sensory information. This means prediction errors may be more reliable in indicating whether our current model of the world needs to be updated or if the incoming sensory input is simply too noisy. Therefore, prediction error signals should be weighted by their precision (or inverse of signal variance), with less weight given to less reliable information. Attention is proposed as a mechanism to optimize these precision estimates (Clark, 2013; Feldman & Friston, 2010; Hohwy, 2012), assigning more weight to reliable prediction errors and facilitating the brain in accurately modeling attended stimuli or events. Consequently, attention modulations in the encoding of prediction errors have been demonstrated through enhanced processing of surprising sensory inputs compared to expected ones (Kok et al., 2012; Marzecová et al., 2017; Smout et al., 2019).

Furthermore, according to the prediction error minimization approach, the precision weights must sum effectively to zero (Hohwy, 2012). This suggests that prediction errors associated with unattended sensory input may be weaker when attention is focused on task-relevant stimuli. Even if prediction errors are generated by surprising unattended input, they would be too weak to bias the model towards a more precise, up-to-date representation of task-irrelevant stimuli. This could potentially impact the awareness or perception of the task-irrelevant stimulus, similar to the theory of biased competition of attention (Spratling, 2008).

Traditionally, attention has been understood to select relevant information for further processing. An outstanding question in this context is whether this further processing includes motion extrapolation. The overarching aim of this thesis is to investigate whether attention can enhance predictive processing, fine-tune motion extrapolation, or modulate the allocation of cognitive resources during complex tasks, as outlined previously.

Overview of methodology

How important is attention for motion-based prediction? While attention is widely acknowledged as indispensable to our visual perception, it remains unclear its specific contribution to the compensatory mechanisms outlined in the first part of this chapter. To address this gap, we have employed a combination of behavioral and EEG approaches to investigate how attentional resources influence motion extrapolation and shape our perception of moving objects in dynamic visual environments involving the flash-lag effect (FLE) and apparent motion.

Flash-lag effect

A motion extrapolation account of the FLE posits that the illusion occurs due to the visual system compensating for the distance a moving object travels during the time it takes for its motion to be processed, accounting for neural delays. In contrast, the flash, having no motion trajectory, does not require such compensation through extrapolation. In addition to motion extrapolation, several hypotheses have been proposed over the past 20+ years to explain the FLE (Hogendoorn, 2020). These alternative explanations primarily focus on the relationship between the flash and moving object, rather than directly addressing neural delays.

The differential latency hypothesis suggests that the perceived displacement in the FLE is not due to actual differences in spatial positions, but rather differences in the processing times of moving and static objects (Purushothaman et al., 1998; Whitney et al., 2000; Whitney & Murakami, 1998).

When a moving object and a flash are presented simultaneously, the perceived position of the moving object appears ahead because motion signals are processed faster than position signals. The differential latency hypothesis is consistent with a larger FLE when a flash is presented in low contrast (Kanai et al., 2004; Purushothaman et al., 1998), influencing the processing speed of the flash (Berry et al., 1999; Schultz & Eriksen, 1978; Töllner et al., 2011).

The attentional shifting hypothesis implies that attention shifts from the moving object to the flash when it appears (Baldo & Klein, 1995, 2010). The moving object continues to move while the flash reaches awareness, resulting in the perceived flash-lag. Similar to differential latency, the attentional shifting mechanism predicts a reduction in the FLE if the flash can be anticipated (Baldo et al., 2002; Chappell et al., 2006; Namba & Baldo, 2004; Shioiri et al., 2010). Cues that allocate attention to the flash before it appears can minimize some processing latency, leading to a smaller FLE. The attentional shifting hypothesis is supported by findings showing a larger FLE with increasing eccentricities between stimuli and a smaller FLE when attention is pre-cued to the location or timing of the flash (Baldo et al., 2002; Chappell et al., 2006; Namba & Baldo, 2004). However, the attentional shifting hypothesis appears to be more relevant to object-based prediction, rather than motion-based prediction (Duncan, 1984; Khurana & Nijhawan, 1995).

The temporal averaging hypothesis suggests that the perceived position of the moving object is an average of its varying positions sampled over a temporal window (Krekelberg & Lappe, 1999, 2000). This window has been estimated to vary from a 60-ms sample after the flash (Brenner & Smeets, 2000), to a ~50-ms window (Whitney et al., 2000) and up to 500 ms (Krekelberg & Lappe, 1999). In contrast, the flash's perceived position is consistent with its veridical position, suggesting that the flash is not subject to the same temporal averaging as the moving object.

The postdiction hypothesis expands upon the theory of temporal averaging, proposing that the position information is similarly integrated over time but reset at the onset of a salient event such as a flash (Eagleman & Sejnowski, 2000). The perceived position of the moving object is estimated from a window of ~80 ms after the flash, and then retrospectively attributed to the moment of the flash. What is intriguing about postdiction is its similarity to extrapolation (prediction) mechanisms, as both involve integrating motion information over time. However, they differ in the temporal window before or after the flash event (Van Heusden et al., 2019). The postdiction hypothesis potentially explains why the FLE does not occur when the moving object reverses direction at the time of the flash (Eagleman & Sejnowski, 2007; but see Nijhawan, 2008).

Finally, the discrete perception hypothesis suggests that the FLE occurs because perception samples information discretely over time in brief windows of [~100–150 ms] rather than continuously (K. A. Schneider, 2018). Each discrete sample represents a momentary snapshot of the visual scene, allowing stimuli to be perceived at their final position within that moment. These discrete samples are integrated over time to form a cohesive perception of motion, with the positions of the moving object interpolated based on the sampled information, similar to postdiction. As the flash is also sampled in the same way, if the flash coincides precisely with the *last* instant of the sampling window and aligns with the final position of the moving object within that moment, there would be no perceived lag between the two stimuli. However, if the flash occurs at a different moment within the sampling window, a lag may be perceived.

In **Chapter 2**, I will revisit these hypotheses to empirically determine which one provides the most parsimonious explanation for the observed illusion when perceived speed is manipulated. In **Chapter 3**, I will further explore how divided attention influences the FLE in a dynamic visual environment, extending our understanding of the mechanisms elucidated in **Chapter 2**.

Apparent motion

Apparent motion is a perceptual phenomenon in which static stimuli presented in rapid successive at different locations create the illusion of continuous motion along a trajectory (Wertheimer, 1912). Typically, this involves a series of discrete frames of stimulus positions, with each frame briefly followed by a blank screen. With appropriate spacing and timing between frames, the brain infers the presence of a single moving object (Ramachandran & Anstis, 1986). This suggests that the brain integrates discrete inputs over time into a cohesive perception of motion between spatially separate stimuli, similar to the aperture problem. The ability to maintain a neural representation of the object’s identity or characteristics is crucial for our perception of apparent motion (Chong et al., 2016; Muckli et al., 2005; Yantis & Nakama, 1998).

The perception of apparent motion closely resembles the perception of smooth motion, despite their technical disparities (Yantis & Nakama, 1998). Engaging similar cortical areas such as MT (Goebel et al., 1998; T. Liu et al., 2004; Muckli et al., 2002), apparent motion also activates V1 as akin to real motion (Muckli et al., 2005). This illusory motion is likely caused by feedback connections from motion processing areas to V1 (Muckli et al., 2002; Sterzer et al., 2006). Notably, disruptions caused by double-pulse TMS over hMT+/V5 diminish apparent motion perception, suggesting the involvement of motion-based prediction (Vetter et al., 2015). Consequently, larger temporal or spatial gaps between frames can distort apparent motion (Burr, 1979; Gepshtein & Kubovy, 2007).

Recent studies have explored compensation for neural delays during apparent motion, with multivariate pattern analyses (MVPA) methods emerging in EEG research. Originally developed for fMRI data, MVPA involves analyzing patterns of neural activity to decode the content and representation of information in the brain. In the context of EEG, MVPA takes advantage of the high-temporal-resolution data provided by EEG recordings to track the rapid changes in brain activity that occur during tasks (Carlson et al., 2020). By leveraging this temporal resolution, MVPA methods can characterize the evolving patterns of neural activity and uncover the underlying representations across time. Specifically, this entails training machine learning classifiers to distinguish between EEG responses to different stimuli or conditions. These trained classifiers are tested on new data, elucidating the information contained in the neural responses, rather than their magnitude. This process can be repeated across time (and conditions) to reveal the time course of information processing in the brain (Grootswagers et al., 2017).

Relevant to apparent motion, MVPA has been used to investigate the neural representations of changing positions of an apparent motion stimulus (Hogendoorn & Burkitt, 2018). Hogendoorn and Burkitt observed predictive information corresponding to the expected position of the stimulus within a motion sequence. In another study, observers tracked or imagined where an apparent motion stimulus would be as it moved behind an occluder (Robinson et al., 2020). Robinson et al. found that neural activity patterns during mental tracking of the occluded stimulus closely mirrored the patterns of activity when the stimulus was visible, reflecting a predictive representation of its anticipated position in the absence of sensory input. These findings resonate with motion extrapolation, where neural populations generate predictions about a stimulus' future position. At the same time, the observed n qeural patterns also align with prediction error signals, arising from the discrepancy between predicted and observed sensory inputs, or their absence. In this context, the visual system's expectation of continuous motion clashes with the unexpected occlusion. As the visual system would not have predicted an occlusion based on the apparent motion of the stimulus, the deviation leads to a prediction error. Consequently, this discrepancy highlights the brain's capability to predict forthcoming events and reconcile unexpected events, thereby ensuring a seamless perception of continuous (apparent) motion even during interruptions (Agostino et al., 2023; Teichmann et al., 2022) or mental imagery (Robinson et al., 2020).

In **Chapter 4**, I will extend these findings and investigate whether these neural representations persist when apparent motion is selectively attended or ignored in a dynamic visual environment.

Overview of empirical studies

The body of this thesis composes three original research articles and manuscripts that empirically test the following questions.

How are moving objects processed compared to static objects?

In **Chapter 2**, I describe a study that used the FLE to understand how moving objects are encoded in the visual system. The FLE is a promising paradigm to investigate how the visual system compensates for its delays through motion extrapolation.

Besides motion extrapolation, alternative explanations of the FLE also offer compelling explanations under specific conditions and are relevant to broader aspects of visual perception. For instance, there is a reported latency advantage of approximately 20 ms for moving objects compared to static objects (Arnold et al., 2009; Hogendoorn & Burkitt, 2018; Jancke et al., 2004). Despite this, the observed difference is insufficient to fully account for the FLE (Arnold et al., 2003). Although the processing speed of the moving object may contribute to the illusion, other factors likely play a role as well. It is worth noting that these explanations may not be as mutually exclusive as originally posed and could even all contribute to the FLE (Holcombe & Corbett, 2023).

However, a critical distinction between motion extrapolation and the other explanations lies in the nature of the perceptual error they propose. Motion extrapolation suggests a spatiotemporal discrepancy, where the perceived position of the moving object is a prediction based on its prior motion. In contrast, other explanations imply a temporal discrepancy, where the perceived difference between the moving object and the flash arises from misalignments in time rather than actual differences in spatial positions. Therefore, motion extrapolation uniquely involves the computation of velocity and position information, allowing the visual system to compute the future position of a moving object along its expected trajectory (Maus, Ward, et al., 2013). Thus, in **Chapter 2**, we verified whether changes in the perceived speed of the moving object modulate the magnitude of the FLE as predicted by motion extrapolation.

Is motion extrapolation modulated by visual attention?

In **Chapter 3**, I describe a study that explored how attention modulates motion extrapolation in the FLE. The previous chapter aimed to investigate the evidence supporting the existence of motion extrapolation within the visual system, where the positions of moving objects are predictively encoded. Moving forward, we shifted our focus to dynamic conditions where multiple objects move concurrently. Given that real-world visual environments often involve multiple independently moving objects, this prompted us to investigate whether and how motion extrapolation persists when attention is distributed among these objects.

To understand the potential role of attention, we can infer from studies of multiple object tracking (MOT; Pylyshyn & Storm, 1988). In MOT tasks, observers are presented with a display containing several identical stimuli, such as dots, which move simultaneously. A subset of these stimuli is highlighted as targets beforehand, and the observers' task is to (covertly) attentionally track these targets among distractors. The starting positions and motion direction of stimuli are random and independent of each other. Typically, the dots move in random motion, but, for context, I will only review studies that used predictable motion (e.g., linear path) here (Adamian & Andersen, 2022; Howe & Holcombe, 2012; Luu & Howe, 2015). In the subsequent display, observers identify the targets based on where they would be, not where they were last seen to be. That is, the predictability of target motion allows observers to predict its likely position with deterministic certainty. There is evidence that observers employ extrapolation to anticipate the future position of targets in MOT, and this ability declines with an increasing number of targets to be tracked, suggesting limitations in attentional capacity. Despite these insights, it was unclear how limited attention precisely leads to variations in the FLE.

In the FLE, only a handful of studies have manipulated attention in paradigms that divert attention from the moving object or allocate attention selectively to various aspects of the display. For example, in dual-task paradigms, observers perform a second task concurrently with a flash-lag task, drawing attention from the moving object. They could be asked to detect brief stimuli presented at fixation while judging the position of the moving object (Sarich et al., 2007). In attention cueing paradigms, spatial cues are used to direct attention to specific locations, such as the probable location of the flash. Attention could be selectively directed towards or away from the moving object (Baldo & Klein, 2010; Namba & Baldo, 2004; Vreven & Verghese, 2005; but see Khurana et al., 2000). In divided attention paradigms, the number or characteristics of moving objects are manipulated. Increasing perceptual load could potentially reduce attentional resources available for processing each object (Shioiri et al., 2010). Across these manipulations, it has been consistently observed that reduced attention results in a larger FLE. This has been explained as attentional enhancements, but the exact mechanisms by which attention influences processing remain unclear. In **Chapter 3**, we investigated how divided attention impacts the magnitude and variability of the FLE.

Does motion extrapolation require visual attention?

In **Chapter 4**, I describe a study that used EEG to investigate how positions of moving objects are represented in the brain. Specifically, during visual tasks, we are often faced with distracting information that interferes with perception and behavioral performance. In the scenario presented at the beginning of this chapter, the presence of a seemingly task-irrelevant feature poses a tension between attending to regularities that might provide valuable information and the consequence of increasing the subset of information to be attended. We were interested in how the brain manages these regularities. In a task where the color of a moving object is relevant, we empirically tested whether position information is predictively encoded even when it is unattended or ignored.

Predictive mechanisms have been observed to manifest as predictive information in the neural representations of stimuli across various modalities and paradigms (e.g., Aitken et al., 2020; Albers et al., 2013; Demarchi et al., 2019; Ekman et al., 2017; S. A. Harrison & Tong, 2009; Kok et al., 2017; Simanova et al., 2016). We know from these studies that prediction shapes perception. In the context of visual motion, these representations refer to an expected position or a future position along the motion trajectory of a moving object.

Recently, Blom et al. (2020) attempted to decode the position of an apparent motion stimulus following the end of a motion sequence. Using a combination of EEG and MVPA methods, Blom et al. trained classifiers to distinguish neural activity patterns of stimulus presentations corresponding to two positions within the motion sequence: the previous position (P_{-1}) and the next position (P_{+1}) relative to the final position (P) of a sequence. Importantly, P_{+1} was predicted from the apparent motion but was never actually presented. The trained classifiers were tested on position P, which allowed them to discern whether representations of previously presented stimuli or predictions of the forthcoming stimulus at were present in the ongoing neural activity. Consistent with the studies mentioned above, the findings revealed evidence of neural representation corresponding to P_{+1} within the neural response to P, even in the absence of sensory input pertaining to P_{+1}. This suggests that these representations were solely driven by the prediction of the stimulus' future position, or there was predictive information about P_{+1}. Additionally, the study observed that the anticipatory representation of P_{+1} was activated before the onset of the stimulus at P. This forward shift in time suggests that predictive mechanisms were at play, compensating for delays in perceptual processing, consistent with motion extrapolation.

This neural evidence for motion extrapolation is has only been investigated in the context of single, attended stimuli (Battistoni et al., 2017). In **Chapter 4**, we explored whether predictive mechanisms drive sensory-like representations when apparent motion stimuli are selectively attended or ignored.

CHAPTER 2

Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed

This chapter is based on the work published in *Vision Research* (Yook et al., 2022). Formatting has been adapted to suit the thesis.

Abstract

In the flash-lag effect (FLE), a flash in spatiotemporal alignment with a moving object is misperceived as lagging behind the moving object. One proposed explanation for this illusion is based on predictive motion extrapolation of trajectories. In this interpretation, the diverging effects of velocity on the perceived position of the moving object suggest that FLE might be based on the neural representation of perceived, rather than physical, velocity. By contrast, alternative explanations based on differential latency or temporal averaging would predict that the FLE does not rely on such a representation of perceived velocity. Here we examined whether the FLE is sensitive to illusory changes in perceived speed that result in changes to perceived velocity, while physical speed is constant. The perceived speed of the moving object was manipulated using revolving wedge stimuli with variable pattern textures (Experiment 1) and luminance contrast (Experiment 2). The motion extrapolation interpretation would predict that the changes in FLE magnitude should correspond to the changes in the perceived speed of the moving object. In the current study, two experiments demonstrated that perceived speed and FLE magnitude increased in the dynamic pattern relative to the static pattern conditions, and that the same effect was found in the low contrast compared to the high contrast conditions. These results showed that manipulations of texture and contrast that are known to alter judgments of perceived speed also modulate perceived position. We interpret this as a consequence of motion extrapolation mechanisms and discuss possible explanations for why we observed no cross-effect correlation.

1. Introduction

In the process of visual perception, delays are incurred as information encoded by the retina is transmitted to the visual cortex (Schmolesky et al., 1998). Neural processing subsequently takes time, so for a time-varying stimulus such as a moving object, its position information would be outdated when it is available in visual cortical areas (De Valois & De Valois, 1991). Because the object would continue moving in the physical world, a lag between its perceived and veridical position would be inevitable at any given time (Maunsell & Gibson, 1992). Intuitively, the impact of such delays should be significant when localizing moving objects, but despite this, we seem to be able to accurately pursue and interact with even fast-moving objects (Smeets & Brenner, 1995b). One proposed explanation is that neural delays are partly compensated perceptually by a process of motion extrapolation (Nijhawan, 1994, 2002; motor compensation is also observed in Kerzel & Gegenfurtner, 2003 and outlined in Nijhawan & Wu, 2009). The motion extrapolation hypothesis proposes that the brain continuously extrapolates the trajectory of a moving object to predict where it is now, as closely as possible to its veridical position (Cavanagh, 1997; Nijhawan & Wu, 2009).

A strong case for visual motion extrapolation has been made on the basis of motion-induced position shifts, many of which indicate a strong coupling between motion and position signals. For example, the Fröhlich effect describes an illusory forward shift in the starting point of the motion trajectory (Fröhlich, 1924), and the representational momentum phenomenon involves a forward displacement in the stopping point of a moving object (Freyd & Finke, 1984; Hubbard, 1995, 2005, 2018). Similarly, when an object moves into a blindspot, the final position of the occluded object is extrapolated in the direction of motion beyond its vanishing point (Maus & Nijhawan, 2008); or if a moving object contains a moving texture, the perceived position of the object appears shifted toward the direction of motion of the texture, as opposed to motion of the object itself (Arnold et al., 2007; De Valois & De Valois, 1991; Roach et al., 2011). Of these illusions, this paper is concerned with the related flash-lag effect (FLE). The FLE has received extensive attention since Nijhawan (1994) reported that a flashed object is misperceived as lagging behind a physically aligned, continuously moving object. Explanations based on motion extrapolation mechanisms propose that the perceived position of the moving object is extrapolated, evoking an apparent spatial offset even though the two objects are always physically aligned (Nijhawan, 1994, 2002). Although different possible causes have been proposed for the FLE (see Hubbard, 2014; Maus et al., 2010; Nijhawan, 2002 for reviews), convergent evidence from neural, behavioral, and computational studies demonstrate how extrapolation could be implemented in the visual system – and that this could compensate for neural delays and potentially explain the above illusions (Hogendoorn, 2020). Were these mechanisms to underlie the FLE, then such a shift should be expected, indeed, for the moving object but not the flash (Maus & Nijhawan, 2006; Nijhawan, 2002).

According to the motion extrapolation interpretation, the predictive mechanism that results in the FLE and related illusions derives from a process of constantly estimating and updating motion and/or position signals (Kwon et al., 2015). Arguably, early extrapolation signals have been shown to carry an estimate of an object's velocity, in order to gauge the distance for the object to (have) travel(ed) (Crick & Koch, 1995; Nijhawan, 2008; Nijhawan & Wu, 2009; Pollen, 2003; Rosenbaum, 1975). Importantly, in both representational momentum (Freyd & Finke, 1985; Hubbard & Bharucha, 1988) and the FLE (Brenner & Smeets, 2000; Krekelberg & Lappe, 1999, 2000; López-Moliner & Linares, 2006; Murakami, 2001; Nijhawan, 1994; Nijhawan et al., 2004), the perceived offset increases linearly with the velocity of the moving object. However, not all FLE studies are consistent. Kanai and colleages (2004) observed that the FLE was unaffected by velocity when FLE magnitudes was measured in spatial units and, in fact, decreased with increased velocity in units of time. Cantor and Schor (2007) showed that the FLE decreases for velocities greater than one degree of visual angle (hereafter denoted as dva throughout the manuscript) per second, while Wojtach et al. (2008) tested a larger range of 3 – 50 dva/s and found that the FLE varied as a logarithmic function of velocity, so the linear effects consistent with previous observations were only shown at the lower velocity ranges. Such inconsistent effects led us to question whether these FLE findings might be due to changes in perceived, rather than physical velocity (Finke et al., 1986; Makin et al., 2009). Although physical velocity and perceived velocity are naturally closely correlated, the distinction is that (physical) velocity could be perceived differently depending on properties such as object contrast (Stone & Thompson, 1992), transiency/duration (Treue et al., 1993), as well as spatial (Smith & Edgar, 1990) and temporal (Shen et al., 2003) frequency. For this reason, we investigate the nature of the velocity representation that contributes to the neural computation that generates the perceived offset in the FLE.

In the present study, we test the hypothesis that the magnitude of the FLE depends on perceived velocity, resulting from changes in perceived speed, even when the physical velocity of the moving object is unchanged. Unique among explanations of the FLE, under the motion extrapolation model, a neural representation of velocity is necessary for an explicit computation that allows us to correctly perceive objects in their current and future positions despite neural delays. Other models of the FLE which attribute the FLE to temporal processes (based on differential processing latency and attentional shifts; Baldo & Klein, 1995; Öğmen et al., 2004; Purushothaman et al., 1998; Whitney et al., 2000; Whitney & Murakami, 1998; or based on temporal averaging of positions, sampling, and postdiction; Brenner & Smeets, 2000; Eagleman & Sejnowski, 2000, 2007; Krekelberg & Lappe, 1999, 2000) would also expect the FLE to increase with faster physical speed – but only as a temporal offset evident as a spatial error (Kanai et al., 2004; Krekelberg & Lappe, 1999, 2000). These models therefore do not require an explicit representation of velocity, and as such, would not be expected to depend on perceived speed when physical speed remains constant.

2. Experiment 1

One method of manipulating perceived speed is by dynamically modulating the texture of a moving object (Treue et al., 1993). When Carlson and colleages (2006) presented a smaller square with a static noise texture superimposed on top of a larger square with dynamic noise and the two objects moved together, the area of static noise was perceived to consistently lag behind the area of dynamic noise – causing an illusory percept that they called the Floating Square illusion. Despite having the same physical velocities, observers perceived the dynamic noise object to move faster and separate from the static noise object. The difference in perceived speed of the two patterns was sufficient to generate large perceptual errors when their relative positions were in fact identical. The results from this experiment are consistent with the idea that motion extrapolation is influenced by perceived speed. We apply this manipulation to an FLE paradigm, where our motion stimulus contains either a static or a dynamic pattern.

Participants carried out a two-alternative forced-choice (2AFC) spatial localization task where they were asked whether a revolving stimulus or a flashed target appeared to lead in the direction of motion, at the moment of the flash. Hereafter we refer to this as the flash-lag task (Experiment 1A). In a control speed-discrimination task (Experiment 1B), participants were asked whether the patterns differed in perceived speed. We hypothesized that if the FLE were sensitive to a change in perceived speed, then both perceived speed and perceived offset would be greater for the dynamic stimulus relative to the static stimulus. Conversely, if the FLE were not dependent on perceived speed, then we would expect no difference in FLE magnitudes as an addition of a dynamic pattern.

2.1 Materials and Methods

2.1.1 Participants

A total of 99 observers participated in one or both parts of Experiment 1: 30 observers participated only in Experiment 1A, and 20 observers participated only in Experiment 1B. To investigate possible correlations between the FLE measure obtained in Experiment 1A and the perceived speed measure obtained in Experiment 1B, a further 49 observers participated in both tasks. Two observers were excluded from Experiment 1A because their results significantly violated normality (Shapiro-Wilk: $p < 0.001$; Kolmogorov-Smirnov: $p < 0.001$). One observer was excluded from both tasks for incorrectly performing the tasks. Together, this yielded a total of 76 observers in Experiment 1A (23 male, age [M: 24.2 years, SD: 6 years]) and 68 observers in

Experiment 1B (23 male, age [M: 24.2 years, SD: 5.6 years]), with 48 of these observers having completed both tasks.

All observers were naïve to the purpose of the experiment, were right-handed, and reported normal or corrected-to-normal vision. Observers gave informed consent and were reimbursed 10 AUD per hour for their participation in the experiment. The study was conducted in accordance with the Declaration of Helsinki and approved by the University of Melbourne Human Research Ethics Committee of the Melbourne School of Psychological Sciences (Ethics ID 1954146.2).

2.1.2 Stimuli

All stimuli were presented on a 24.5-inch ASUS ROG PG258 monitor (ASUS, Taipei, Taiwan) with a resolution of 1920 × 1080 running at 200 Hz, controlled by an HP EliteDesk 800 G3 TWR running MATLAB R2017b (The MathWorks Inc., 2017) with PsychToolbox 3.04.14 extensions (Kleiner et al., 2007). Participants sat in a light-attenuated room and viewed the stimuli from a headrest at 60 cm from the screen.

The motion stimulus consisted of a wedge segment superimposed on a static annulus, which was centered at a white fixation point and displayed on a uniform 50% gray background (**Figure 1A**). The wedge's inner and outer edges were 4.05 dva and 6.75 dva away from the fixation point and subtended 45 degrees of polar angle (hereafter denoted as ° throughout the manuscript; whereas, dva denotes degrees of visual angle) along the radial axis. The wedge revolved at a fixed speed of 200°/s. The direction of motion varied randomly between clockwise or counter-clockwise across trials but was constant within a given trial.

The wedge and annulus contained one of two textures. Filtered noise was created by applying a low-pass filter on white noise at 0.5 cycles/deg, and pixelated noise was created using an array of 5 x 5 pixel blocks with random luminance. The dynamic pattern was added by replacing the textures with a new random set of pixels, of no correlation with the previous set, at every frame. This configuration resulted in a flickering appearance of the texture. The background of the annulus always had the static pattern, because otherwise the wedge would be indistinguishable from the annulus if both textures were patterned dynamically (second-order motion).

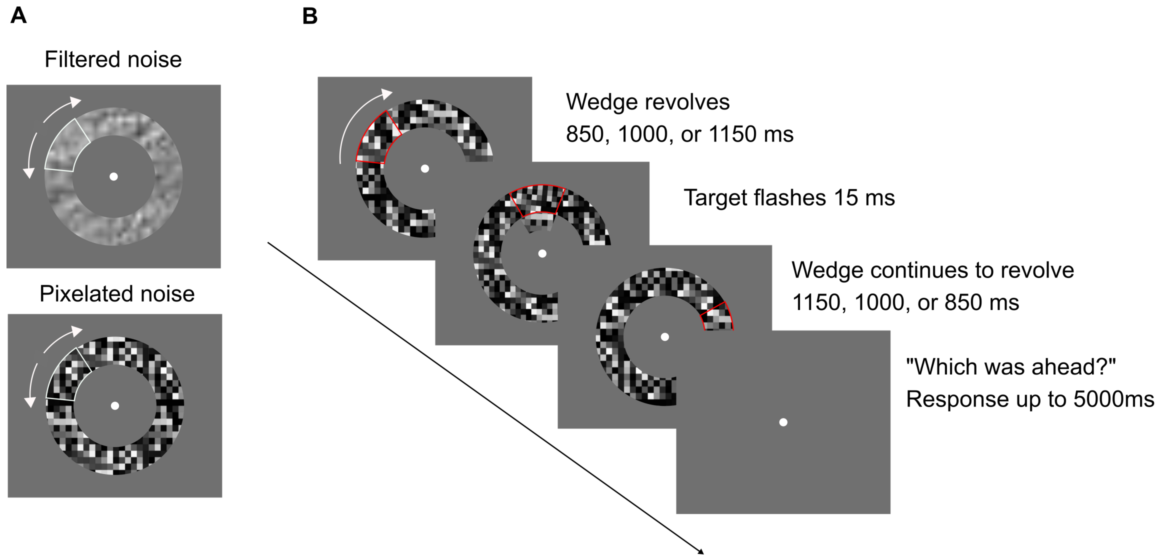

Figure 1. Schematic illustration of the stimulus configuration and flash-lag sequence in Experiment 1A. (**A**) Wedge and annulus containing filtered noise (top) and pixelated noise (bottom) texture (see also videos of the stimuli used in this experiment; speed reduced for illustrative purposes). (**B**) During a given trial, the wedge revolved within the annulus (shown with a pixelated noise texture and outlined in red in the figure). The target flashed after half the time of the overall sequence, and after a delay, observers were asked to respond using a keypress. A complete trial lasted up to 7 s, with the presentation sequence lasting 2 s. In Experiment 2A, the motion stimulus was a wedge segment without the annulus.

In Experiment 1A, observers compared the position of the wedge to a static, target stimulus. The target was a stationary wedge with the same static texture (**Figure 1B**). Its inner and outer edges corresponded to 2.43 dva and 3.78 dva from fixation and subtended 45° along the radial axis.

In Experiment 1B, observers compared the speed of the wedge to a comparison. The comparison was a solid black wedge. The speed of the wedge was constant at 200°/s, but the speed of the comparison varied across trials using a staircase procedure (see **section 2.1.3** below). Both stimuli always revolved in the same direction, which was randomly chosen at the start of each trial.

2.1.3 Procedure

2.1.3.1 Experiment 1A: Flash-lag task

On each trial, the wedge was presented from a random starting position of the annulus and revolved for one of three possible durations (850, 1000, or 1150 ms chosen at random). After half the time of the overall sequence, the target was presented for 15 ms along the inner circumference of the annulus, after which the wedge continued revolving for the remaining 1150, 1000, or 850 ms of its trajectory (cumulative presentation duration of 2 s). Then, observers were instructed to report whether the wedge or target was spatially ahead of the other when the target had appeared (2AFC). Keypress responses were recorded in the next 5 s – 'i' as in inner (target) or 'o' as in outer (wedge) – and prompted the next trial.

The relative spatial location of the wedge and target was driven by a one-up, one-down staircase procedure, and trials were randomly drawn from two interleaved staircases. At the first trial of each condition block, one staircase started at a large difference of + 20° (moving ahead of flash). A –20° starting point was omitted in this procedure as a reversing, flash-lead presentation would be redundant for the purpose of our hypothesized flash-lag illusion. Instead, the second staircase started at a smaller difference of 0° (physically aligned). Depending on the response, the difference was adjusted by ± 2° in the following trial drawn from that staircase. Observers completed 120 trials from each staircase, and the final 60 points of each staircase were averaged to estimate the point of subjective equality (PSE): the displacement at which the wedge and target were perceived as aligned.

Trials were blocked into 4 conditions. Each condition consisted of 120 trials (each lasting up to 7 s), for a total of 480 trials over 4 blocks. The conditions were determined based on a combination of *pattern type* (static or dynamic) and *texture type* (filtered noise or pixelated noise). The order of condition was randomized for observers who completed only Experiment 1A and counterbalanced if they completed both Experiment 1A and 1B. All observers completed 12 trials as practice before each experimental block, followed by a self-paced break. This resulted in a total testing duration of approximately 30 minutes.

2.1.3.2 Experiment 1B: Speed-discrimination task

On each trial, the comparison was presented from a random starting position against a 50% gray background and revolved for a randomly chosen duration of 1500, 1750, or 2000 ms (**Figure 2**). Then, the comparison was removed from display, and the wedge from Experiment 1A was presented from a random starting position of the annulus and revolved for 1500, 1750, or 2000 ms also randomly chosen. Presentation duration and starting position were varied so that observers could not reliably base their responses on the difference between the distance traveled by each wedge. After the sequence ended, observers were instructed to report whether the comparison or the wedge revolved faster (2AFC). Keypress responses were recorded in the next 5 s – '1' for first (comparison) or '2' for second (wedge) – and prompted the next trial.

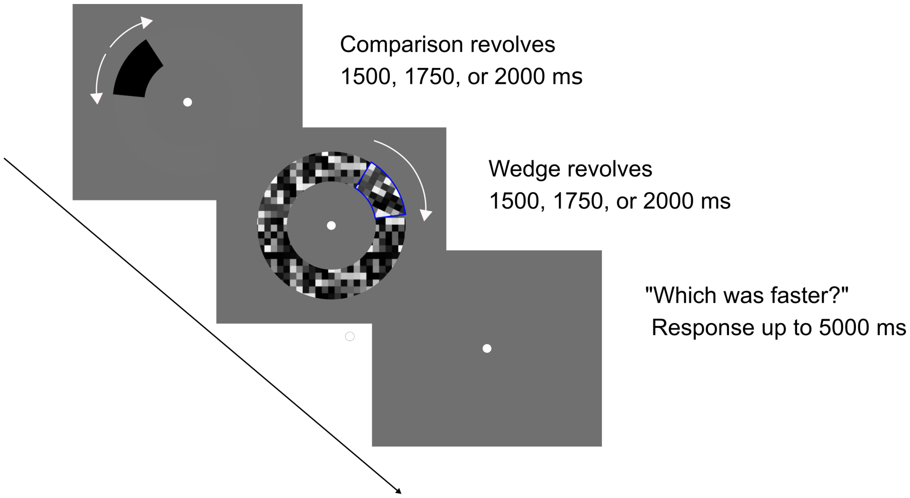

Figure 2. Schematic illustration of Experiment 1B. During a given trial, the comparison and the wedge within the annulus (shown with a pixelated noise texture and outlined in blue in the figure) were presented sequentially. After a delay, observers were asked to respond using a keypress. A complete trial lasted up to 9 s, with the presentation sequence lasting up to 4 s. In Experiment 2B, the motion stimulus was a wedge segment without the annulus.

While the speed of the wedge was fixed, the speed of the comparison was determined using a similar staircase procedure using the ratio of these two stimulus speeds. Trials were drawn from one of two interleaved staircases. The first staircase was initialized at $1.075^{7°}$, and the second started at $1.075^{-7°}$, with a step size set to 1.075° for both. The staircase value of a given trial was proportionate to the relative speed of the comparison and the wedge. Because velocity is represented logarithmically in perception (Nover et al., 2005; Priebe & Lisberger, 2004), the steps were adjusted by means of multiplying and dividing, rather than adding and subtracting. For example, $1.075^{7°}$ (~ 1.66) indicated that the comparison revolved at 166% of the speed of the wedge (166% of 200°/s). Depending on the response, the speed of the comparison was adjusted by 1.075° in the following trial drawn from that staircase. Observers completed 72 trials from each staircase, and the last 36 points of each staircase were averaged to estimate the PSE: the speed at which the comparison and wedge were perceived to match.

Trials were blocked into the same 4 conditions of Experiment 1A: *pattern type* (static or dynamic) and *texture type* (filtered noise or pixelated noise). All observers completed a total of 4 blocks. Observers who completed only Experiment 1B completed 72 trials per block (for a total of 288 trials), and observers who completed both Experiment 1A and 1B completed 144 trials per block (for a total of 576 trials). The total testing duration was 25 and 50 minutes, respectively.

2.2 Results

We tested for differences in the magnitude of the FLE and perceived speed in each condition using a two-way repeated measures ANOVA with factors of *pattern type* (static or dynamic) and *texture type* (filtered noise or pixelated noise). Post hoc paired-samples *t*-tests (two-tailed) with Bonferroni corrections estimated these differences.

2.2.1 Experiment 1A: Effect of noise texture on perceived flash-lag

A repeated measures ANOVA revealed a significant main effect of pattern type ($F_{1,75} = 5.14$, $p = 0.026$), with greater FLE magnitudes observed for dynamic pattern than for static pattern wedges (**Figure 3A**). There was no significant main effect of texture type ($F_{1,75} = 1.61$, $p = 0.209$) and no interaction effect between the two factors ($F_{1,75} = 0.81$, $p = 0.37$). The mean PSE was significantly larger for dynamic pattern compared to static pattern wedges for the filtered noise texture ($t_{75} = 2.49$, $p = 0.015$), but the difference did not reach significance for the pixelated noise texture ($t_{75} = 1.03$, $p = 0.309$).

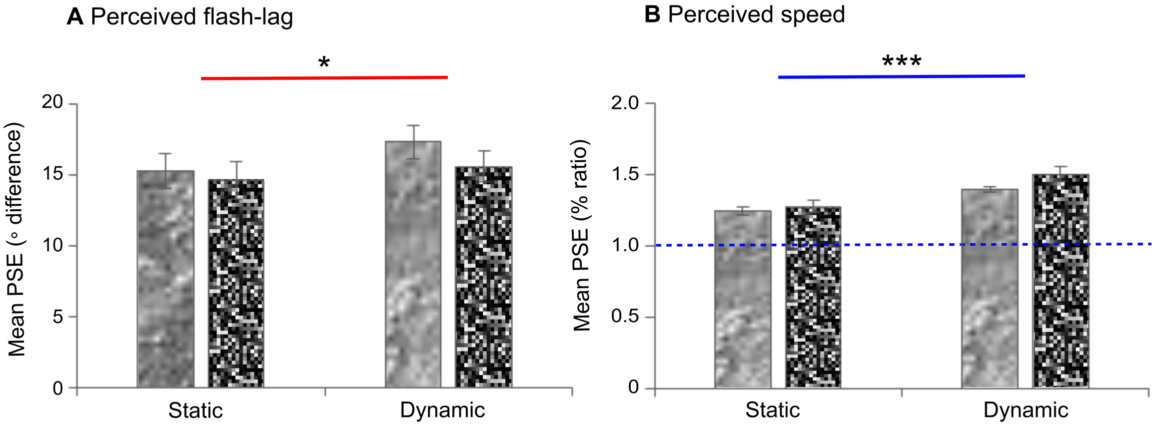

Figure 3. Results of Experiment 1. For illustrative purposes, the bars contain corresponding textures in the figure. (**A**) Mean FLE magnitudes of static pattern and dynamic pattern wedges across the texture conditions in Experiment 1A. (**B**) Mean perceived speeds of static pattern and dynamic pattern wedges across the texture conditions in Experiment 1B. Perceived speed was calculated as a ratio of the comparison speed relative to the wedge speed (taller bars mean that objects in that condition were perceived as faster). Baseline speed is indicated by a horizontal dashed line at y = 1. Mean perceived speeds were significantly different from zero and between pattern conditions. Error bars represent standard errors across observers. Asterisks above the bars indicate a statistically significant main effect of pattern type. * denotes $p < 0.05$. *** denotes $p <$

2.2.2 Experiment 1B: Effect of noise texture on perceived speed

A repeated measures ANOVA showed a significant main effect of pattern type on perceived speed ($F_{1,67} = 35.82$, $p < 0.001$), with higher perceived speeds reported for dynamic pattern than for static pattern wedges (**Figure 3B**). This effect was significant for both filtered noise ($t_{67} = 4.08$, $p < 0.001$) and pixelated noise ($t_{67} = 5.33$, $p < 0.001$) textures. In addition, there was a significant main effect of texture type ($F_{1,67}= 11.5$, $p = 0.001$) but no significant interaction effect ($F_{1,67} = 2.518$, $p = 0.117$).

2.2.3 Cross-effect correlation

In order to further evaluate the possible relationship between the two measures, we investigated whether a dynamic pattern affected perceived speed and perceived position (flash-lag) in the same way. We were interested in the relation between one effect and the other, and if this was consistent between observers in a cross-effect correlation analysis. Among 48 observers who completed both tasks, we hypothesized that observers for whom the dynamic pattern textures greatly increased perceived speed would be expected to report larger changes in the illusion, and conversely, observers with minimal (or negative) changes in perceived speed would be expected to show corresponding small or negative effects in the illusion. However, Pearson's correlation (two-tailed) did not reveal a significant correlation between the two tasks in either filtered noise ($r = 0.15$, $p = 0.3$) or pixelated noise ($r = 0.07$, $p = 0.64$) conditions (**Figure 4**).

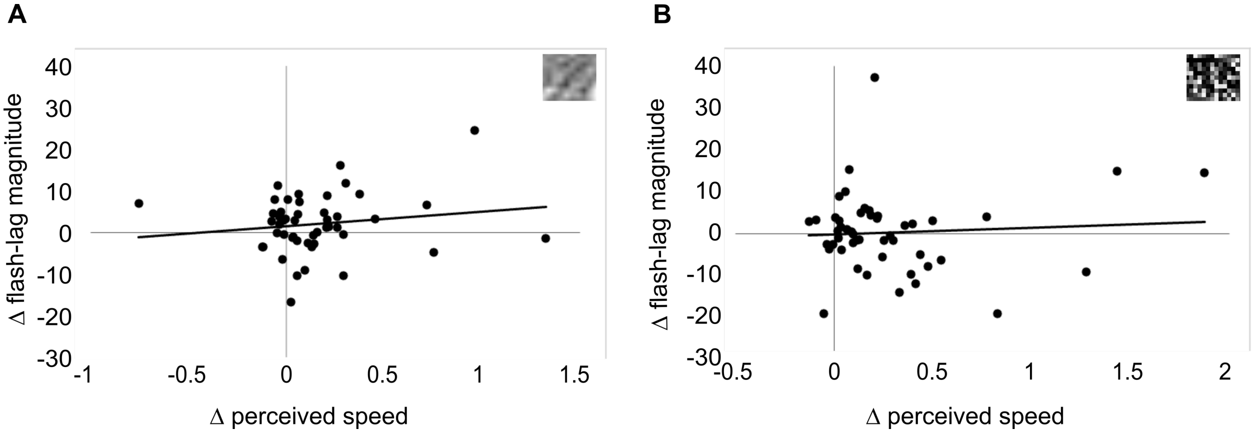

Figure 4. Change in FLE magnitude plotted as a function of change in perceived speed for (**A**) filtered noise and (**B**) pixelated noise textures. Legend for texture type is shown in the top right corner of each panel.

3. Experiment 2

Overall, Experiment 1 demonstrates that the dynamic modulation of a moving object's texture, which is known to influence perceived speed, also influences the perceived position of the moving object – systemically in the same direction in the FLE, as predicted by the motion extrapolation model. To provide confidence that this effect is due to changes in perceived speed rather than unknown low-level factors, we carried out a parallel set of experiments to investigate whether a different, independent manipulation of perceived speed similarly affects the FLE.

The well-known 'Thompson effect' describes how speed is underestimated at low contrasts (Anstis, 2003; Blakemore & Snowden, 1999; Snowden et al., 1998; Thompson, 1982). As a natural example, this has been used to explain why drivers frequently drive too fast under foggy conditions, compensating for their incorrect estimates of how fast (or slow) they and other cars are moving (Anstis, 2003; Snowden et al., 1998; but see Owens et al., 2010; Pretto et al., 2012). Theoretically, this effect is consistent with the findings of Berry et al. (1999), which showed that anticipatory neural responses for the leading edges of a moving object varied based on the contrast of the object. It also resembles the Hess effect (Hess, 1904), a similar illusion to the Floating Squares illusion and the FLE, in which a brighter object is perceived to lead a dimmer object when the two are actually aligned and have the same physical velocities. However, the effects of contrast on perceived speed depend on properties such as object luminance or physical velocity, and a handful of studies previously demonstrated that the effect can reverse under certain circumstances. For example, perceived speed can be overestimated at higher velocities (>8 dva/s; Hawken et al., 1994; Pretto et al., 2012; Thompson et al., 2006) and low luminance levels (Vaziri-Pashkam & Cavanagh, 2008, 2011).

The effects of contrast have also been specifically investigated in the context of motion extrapolation. In the FLE, Kanai et al. (2004) found increases in FLE magnitude with decreases in contrast of moving objects with the background. Kanai et al. suggested that the FLE was modulated by positional uncertainty, as visibility is affected when the contrast of a target and background is reduced. Maus and Nijhawan (2006, 2009) also reported that the forward displacement of a moving object increases if the contrast of the target gradually decreased, while Hubbard and Ruppel (2014) reported that representational momentum decreases with low or decreasing contrast of the target and background. Finally, Vaziri-Pashkam and Cavanagh's (2011) FLE paradigm presented a moving patch of random dots against the background containing static random dots. While they previously found that the perceived speed of moving objects increases at low luminance, caused by motion blurring for faster speeds (Vaziri-Pashkam & Cavanagh, 2008), there were no additional effects on FLE magnitude due to these changes in perceived speed.

We reexamined the effects of contrast in Experiment 2. In line with Experiment 1, the primary (moving) stimulus revolved at a speed of 200°/s, but presented at either high (100%) or low (10%) contrast. Based on the divergent effects of contrast on perceived speed in the previous literature as briefly discussed above, we had no a priori directional hypothesis about the effect of contrast on perceived speed. Instead, we hypothesized that any effect of contrast on perceived speed would be mirrored in an effect on FLE magnitudes in the same direction. In other words, mean FLE magnitudes would increase in high contrast relative to low contrast conditions if perceived speed increased with high contrast, or in the opposite direction, it would increase for low contrast compared to high contrast conditions if perceived speed increased with low contrast. Conversely, if the FLE were not driven by perceived speed, then we would expect an effect of contrast on perceived speed, but no effect on FLE magnitudes.

3.1 Materials and Methods

3.1.1 Participants

Seven observers (3 males, mean age 25.3 years, SD 3.4 years) participated in the experiment. All were experienced observers and completed both Experiment 2A (Flash-lag) and 2B (Speed-discrimination). All inclusion criteria were identical to Experiment 1.

3.1.2 Stimuli

Stimuli were similar to those used in Experiment 1 with the following exceptions (see also videos of the stimuli used in this experiment). In Experiment 2, the motion stimulus consisted of a wedge segment without the annulus, against a uniform gray background. In doing so, we aimed to control for a potential confounding effect, as otherwise observers would be required to detect the contrast of the wedge on top of the background of the annulus against the background, while tracking the wedge's motion (Blakemore & Snowden, 2000).

Observers viewed the stimuli from a distance of 50 cm. The inner and outer edges of the wedge and the comparison were 4.75 dva and 7.92 dva away from the fixation point and subtended 45° along the radial axis. The inner and outer edges of the target were 3.46 dva and 4.44 dva from fixation, subtending 45° along the radial axis. Stimuli textures were filtered noise or pixelated noise and always presented as a static pattern. The main experimental manipulation of Experiment 2 was that on each trial, the wedge was presented at high (100%) or low (10%) contrast relative to the gray background. The target was always presented at 100% contrast.

3.1.3 Procedure

Due to the COVID-19 pandemic, observers participated from their own homes. All observers were provided with identical-speed ASUS ROG PG258Q monitors (1920 x 1080 resolution, 200 Hz), and completed the experiment in a dimly lit room. A chinrest was not used for this experiment, but observers were instructed to sit approximately 50 cm from the screen and make sure the fixation point was at eye level.

The experimental conditions were determined by a combination of *contrast level* (high or low) and *texture type* (filtered noise or pixelated noise). All observers completed a total of 4 blocks: 120 trials per block (for a total of 480 trials) in Experiment 2A and 144 trials per block (for a total of 576 trials) in Experiment 2B. The total testing duration was approximately 80 minutes.

3.2 Results

As in Experiment 1, we probed FLE magnitudes for each condition in Experiment 2A and perceived speed in Experiment 2B. We submitted all data to a two-way repeated measures ANOVA with factors of *contrast level* (high or low) and *texture type* (filtered noise or pixelated noise) and *t*-tests. All relevant assumptions were met.

3.2.1 Experiment 2A: Effect of contrast on perceived flash-lag

A repeated measures ANOVA revealed a significant main effect of contrast level ($F_{1,6} = 7.32$, $p = 0.035$), with greater FLE magnitudes observed for low contrast than for high contrast wedges (**Figure 5A**). There was a main effect of texture type ($F_{1,6} = 28.87$, $p = 0.002$), with greater FLE magnitudes observed for pixelated noise than for filtered noise conditions. There was no significant interaction effect ($F_{1,6} = 0.47$, $p = 0.52$). The mean PSE was significantly larger for low contrast compared to high contrast wedges for filtered noise ($t_6 = 3.47$, $p = 0.013$), but the difference did not reach significance for pixelated noise ($t_6 = 2.12$, $p = 0.078$). For individual data, the FLE consistently increased in low contrast compared to high contrast conditions for all but two observers (P6 for filtered noise and P3 for pixelated noise; **Figure 5B**).

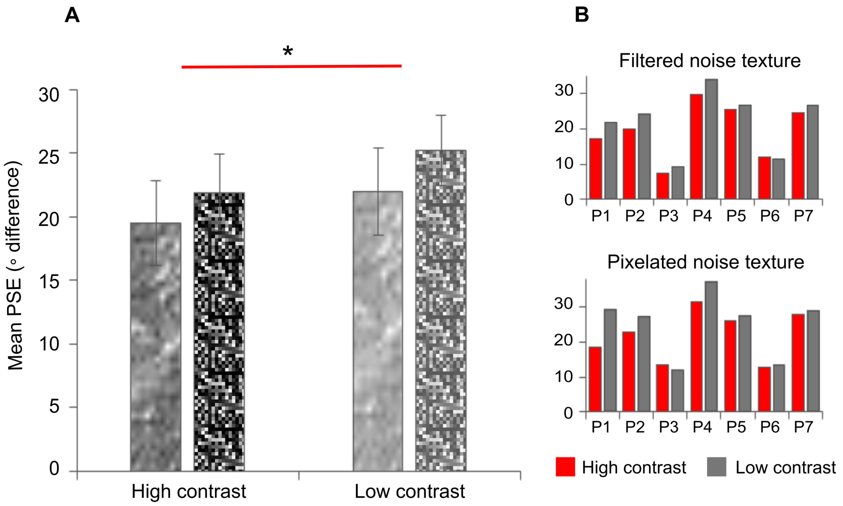

Figure 5. Results of Experiment 2A. For illustrative purposes, the bars contain corresponding textures and contrasts in this figure. (**A**) Mean FLE magnitudes of high and low contrast wedges across texture conditions, pooled over all seven observers. (**B**) Individual observers. Error bars represent standard errors across observers. The asterisk above the bars indicates a statistically significant main effect of contrast level at $p < 0.05$.

3.2.2 Experiment 2B: Effect of contrast on perceived speed

A repeated measures ANOVA revealed a significant main effect of contrast level on perceived speed ($F_{1,6} = 52.85$, $p < 0.001$), with low contrast perceived to move faster than high contrast wedges of the same velocity (**Figure 6A**). There was no significant main effect of texture type ($F_{1,6} = 0.086$, $p = 0.779$) or an interaction effect ($F_{1,6} = 4.68$, $p = 0.074$). The mean PSE was significantly larger for low contrast compared to high contrast wedges in both filtered noise ($t_6 = 5.72$, $p = 0.001$) and pixelated noise ($t_6 = 7.31$, $p < 0.001$) conditions. The effect was highly consistent across individual observers (**Figure 6B**).

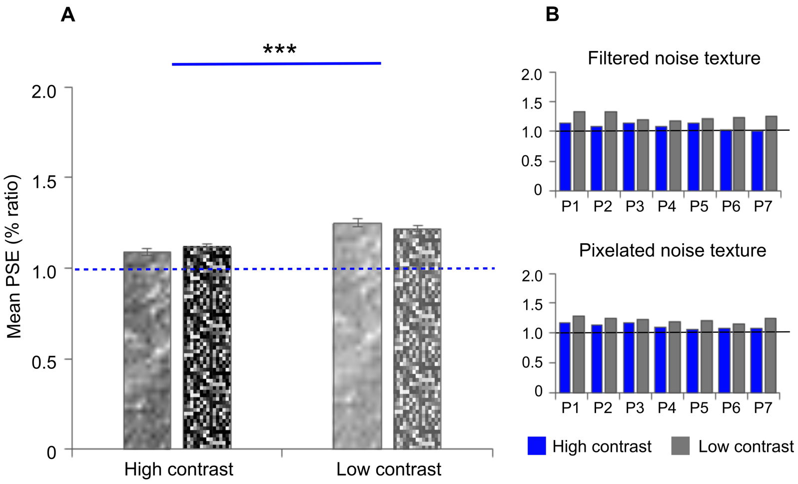

Figure 6. Results of Experiment 2B. For illustrative purposes, the bars contain corresponding textures and contrasts in this figure. (**A**) Mean perceived speeds of high and low contrast wedges across texture conditions, pooled over all seven observers. (**B**) Individual observers. Perceived speed was calculated as a ratio of the comparison relative to the wedge speed, and the baseline speed is indicated as y = 1. Error bars represent standard errors across observers. Asterisks above the bars indicate a statistically significant main effect of contrast level. *** denotes $p < 0.001$.

4. Discussion

The flash-lag effect (FLE) is a classic illusion that has been argued to result from motion extrapolation mechanisms that might be implemented in the visual system to contribute to overcoming its time delays. Although many alternative explanations for the FLE have been presented, the motion extrapolation account uniquely posits a neural computation that requires a representation of velocity. We hypothesized that under this interpretation, experimental manipulations that affect perceived velocity might similarly influence the magnitude of the FLE. Therefore, we examined how illusions of perceived speed could alter the perceived flash-lag in two experiments. We show that dynamically modulating a texture causes an illusory percept of increased speed as well as increased FLE magnitude compared to a static pattern. Similarly, when objects were presented at low contrast, both perceived speed and FLE magnitude also increased relative to high contrast. These findings appear to be consistent with the motion extrapolation account of the FLE.

Experiment 1 shows robust differences between dynamic and static patterns for perceived speed and FLE magnitude. Such differences were qualitatively similar across different textures. Overall, these results closely reproduce the previously reported effects on the perceived speed of gratings, Gabor patches, and in the Floating Square illusion (Carlson et al., 2006; De Valois & De Valois, 1991; Shen et al., 2003; Treue et al., 1993), and are also consistent with the effects of speed on motion extrapolation (Brenner & Smeets, 2000; Freyd & Finke, 1985; Hubbard & Bharucha, 1988; Krekelberg & Lappe, 1999, 2000; López-Moliner & Linares, 2006; Murakami, 2001; Nijhawan, 1994; Nijhawan et al., 2004).

Experiment 2 likewise shows subjective increases in both perceived speed and FLE magnitude at low contrast relative to high contrast. These results diverge from findings of previous motion extrapolation studies on the effects of contrast (e.g., Hubbard & Ruppel, 2014; Maus & Nijhawan, 2006, 2009), in that our motion stimulus was perceived as faster at low contrast than high contrast. The discrepancy with the Thompson effect is likely due to methodological differences because our wedge had revolved at approximately 28.15 dva/s, which is significantly faster than in the literature we outlined here (approximately 3–17 dva/s; Blakemore & Snowden, 1999; Hubbard & Ruppel, 2014; Maus & Nijhawan, 2006, 2009; Stone & Thompson, 1992; Thompson et al., 2006). Consistent with this interpretation, Vaziri-Pashkam and Cavanagh (2011) found that out of a range of velocities, only their highest velocity of 27.36 dva/s was overestimated at low contrasts.

If a neural representation of perceived velocity is used to inform the extrapolation process that underlies the FLE, then we would expect to see a causal relationship between the effects of dynamic pattern on perceived speed and the effects on the FLE. However, we did not observe a correlation between the behavioral responses, similar to Vaziri-Pashkam and Cavanagh (2011), with a portion of observers showing a decrease in FLE magnitude despite clear increases in perceived speed. One possibility is that the experiments in this study illustrate that different processes are involved in making judgments of speed versus relative position. For example, de'Sperati & Thornton (2019) showed that varying the contrast of a moving object influenced extrapolation judgments but not interceptive decisions when observers were required to make saccadic eye movements to avoid colliding with multiple moving targets. Based on these findings, de' Sperati and Thornton suggested these tasks involve intrinsically different processes. An important question arising from our findings is therefore whether the representations of velocity that contribute to explicit reports of motion (i.e. how fast an object is perceived to move) differ from the representations used for implicit processing of position (i.e. where an object is; cf. Brenner & Smeets, 1994; Smeets & Brenner, 1995a) – and involve different cortical areas – or, alternatively, if a shared velocity representation might cause greater differences in perceived speed than in the FLE. The latter explanation could reconcile findings of Vaziri-Pashkam and Cavanagh (2011), as at low luminance the effects of perceived speed did not significantly affect the overall magnitude of the FLE any more than the effects of physical speed.

Furthermore, there was a large variability for individual observers in Experiments 1A and 1B. This could be partly attributed to measurement noise, so we can suspect that the effect of perceived speed (although robust within individuals) might explain only a small proportion of the variance in the illusion across individuals in our study. Other factors might also influence responses in the FLE (thereby contributing to error variance). Kanai et al. (2004) reported FLE biases in the same direction, but they explain their effect of contrast using uncertainty (see also Fu et al., 2001; Maus & Nijhawan, 2006, 2009; Purushothaman et al., 1998; Vreven & Verghese, 2005). According to this interpretation, increases in FLE magnitude are caused by increases in positional uncertainty, due to decreased contrast rather than increased perceived speed. This reasoning is also compatible with Experiment 1 if the dynamic pattern would increase uncertainty; presumably, dynamically patterned or low contrast objects would be expected to elicit weaker or noisier position signals for the visual system to work with. In turn, the perceived position might depend more strongly on predictions originating from extrapolation processes, and be less strongly influenced by (uncertain) sensory information. This would, in principle, be expected to cause increases in FLE magnitude for dynamic pattern or low contrast conditions as we observe here. However, the hypothesized effects of uncertainty have been shown to have little impact on judgments of speed (Stocker & Simoncelli, 2006), and observers can accurately report the extrapolated positions of moving objects despite decreases in visibility (Graf et al., 2005), making this interpretation unlikely.

Instead, we might consider if there is an indirect or additional influence of uncertainty. Uncertainty could predict an increase in error variance, and this would be expected to work against any systemic biases, rather than cause one. This would be compatible with the absence of a cross-effect correlation. Another possible influence on our effects is not in terms of uncertainty, but in the amount of attention allocated to the moving object. We cannot rule this possibility out, as we did not explicitly control or measure attention in our different conditions. Conceivably, an object containing a dynamic pattern might require more attention to localize than an object containing a static texture. If this were true, then we might expect a greater FLE, based on studies showing that the FLE increases when attention is divided (e.g., across concurrent tasks, across multiple stimuli, or for increased speeds; Sarich et al., 2007; Scocchia et al., 2009). This explanation also applies if low contrast requires more attention than high contrast. These alternative interpretations would need to be tested by future experiments.

Overall, the pattern of results is consistent with the notion that the magnitude of the FLE depends on a neural representation of velocity that is sensitive to (some of) the same experimental manipulations that also affect over-reports of perceived speed. This finding is consistent with an explanation of the FLE in terms of visual motion extrapolation. It also corroborates a growing body of evidence supporting the existence of neural mechanisms involved in extrapolation in both animal models (Benvenuti et al., 2020; Berry et al., 1999; Jancke et al., 2004; Palmer et al., 2015; Schwartz et al., 2007; Subramaniyan et al., 2018; Sundberg et al., 2006) and human neuroimaging (Blom et al., 2020; Ekman et al., 2017; Hogendoorn & Burkitt, 2018; Schneider et al., 2019), and is consistent with several decades investigating motion extrapolation in perception (Hogendoorn, 2020; Hubbard, 2005, 2014, 2018; Maus et al., 2010; Nijhawan, 2002, 2008).

References

Anstis, S. (2003). Moving objects appear to slow down at low contrasts. *Neural Networks*, *16*(5–6), 933–938. https://doi.org/10.1016/S0893-6080(03)00111-4

Arnold, D. H., Thompson, M., & Johnston, A. (2007). Motion and position coding. *Vision Research*, *47*(18), 2403–2410. https://doi.org/10.1016/j.visres.2007.04.025

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature*, *378*(6557), Article 6557. https://doi.org/10.1038/378565a0

Benvenuti, G., Chemla, S., Boonman, A., Perrinet, L., Masson, G. S., & Chavane, F. (2020). *Anticipatory responses along motion trajectories in awake monkey area V1* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.03.26.010017

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature*, *398*(6725), 334–338. https://doi.org/10.1038/18678

Blakemore, M. R., & Snowden, R. J. (1999). The Effect of Contrast upon Perceived Speed: A General Phenomenon? *Perception*, *28*(1), 33–48. https://doi.org/10.1068/p2722

Blakemore, M. R., & Snowden, R. J. (2000). Textured backgrounds alter perceived speed. *Vision Research*, *40*(6), 629–638. https://doi.org/10.1016/S0042-6989(99)00214-X

Blom, T., Feuerriegel, D., Johnson, P., Bode, S., & Hogendoorn, H. (2020). Predictions drive neural representations of visual events ahead of incoming sensory information. *Proceedings of the National Academy of Sciences*, *117*(13), 7510–7515. https://doi.org/10.1073/pnas.1917777117

Brenner, E., & Smeets, J. B. J. (1994). Different frames of reference for position and motion. *Naturwissenschaften*, *81*(1), 30–32. https://doi.org/10.1007/BF01138558

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash–lag effect. *Vision Research*, *40*(13), 1645–1648. https://doi.org/10.1016/S0042-6989(00)00067-5

Cantor, C. R. L., & Schor, C. M. (2007). Stimulus dependence of the flash-lag effect. *Vision Research*, *47*(22), 2841–2854. https://doi.org/10.1016/j.visres.2007.06.023

Carlson, T. A., Schrater, P., & He, S. (2006). Floating square illusion: Perceptual uncoupling of static and dynamic objects in motion. *Journal of Vision*, *6*(2), 4. https://doi.org/10.1167/6.2.4

Cavanagh, P. (1997). Visual perception. Predicting the present. *Nature*, *386*(6620), 19, 21. https://doi.org/10.1038/386019a0

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, *375*(6527), 121–123. https://doi.org/10.1038/375121a0

de'Sperati, C., & Thornton, I. M. (2019). Motion prediction at low contrast. *Vision Research*, *154*, 85–96. https://doi.org/10.1016/j.visres.2018.11.004

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research*, *31*(9), 1619–1626. https://doi.org/10.1016/0042-6989(91)90138-U

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion Integration and Postdiction in Visual Awareness. *Science*, *287*(5460), 2036–2038. https://doi.org/10.1126/science.287.5460.2036

Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias localization judgments: A unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. *Journal of Vision*, *7*(4), 3–3. https://doi.org/10.1167/7.4.3

Ekman, M., Kok, P., & De Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*(1), 15276. https://doi.org/10.1038/ncomms15276

Finke, R. A., Freyd, J. J., & Shyi, G. C. (1986). Implied velocity and acceleration induce transformations of visual memory. *Journal of Experimental Psychology: General*, *115*(2), 175–188. https://doi.org/10.1037/0096-3445.115.2.175

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126–132. https://doi.org/10.1037/0278-7393.10.1.126

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society*, *23*(6), 443–446. https://doi.org/10.3758/BF03329847

Fröhlich, F. W. (1924). Über die Messung der Empfindungszeit. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, *202*(1), 566–572. https://doi.org/10.1007/BF01723521

Fu, Y.-X., Shen, Y., & Dan, Y. (2001). Motion-Induced Perceptual Extrapolation of Blurred Visual Targets. *The Journal of Neuroscience*, *21*(20), RC172–RC172. https://doi.org/10.1523/JNEUROSCI.21-20-j0003.2001

Graf, E. W., Warren, P. A., & Maloney, L. T. (2005). Explicit estimation of visual uncertainty in human motion processing. *Vision Research*, *45*(24), 3050–3059. https://doi.org/10.1016/j.visres.2005.08.007

Hawken, M. J., Gegenfurtner, K. R., & Tang, C. (1994). Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature*, *367*(6460), 268–270. https://doi.org/10.1038/367268a0

Hess, C. (1904). Untersuchungenu¨ber den Erregungsvorgang im Sehorgan bei kurz- und bei la¨ngerdauernder Reizung. . . *Pflu¨ger's Arch. Ges. Physiol.*, *101*, 226–262.

Hogendoorn, H. (2020). Motion Extrapolation in Visual Processing: Lessons from 25 Years of Flash-Lag Debate. *The Journal of Neuroscience*, *40*(30), 5698–5705. https://doi.org/10.1523/JNEUROSCI.0275-20.2020

Hogendoorn, H., & Burkitt, A. N. (2018). Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. *NeuroImage*, *171*, 55–61. https://doi.org/10.1016/j.neuroimage.2017.12.063

Hubbard, T. L. (1995). Environmental invariants in the representation of motion: Implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychonomic Bulletin & Review*, *2*(3), 322–338. https://doi.org/10.3758/BF03210971

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, *12*(5), 822–851. https://doi.org/10.3758/BF03196775

Hubbard, T. L. (2014). The flash-lag effect and related mislocalizations: Findings, properties, and theories. *Psychological Bulletin*, *140*(1), 308–338. https://doi.org/10.1037/a0032899

Hubbard, T. L. (2018). Influences on representational momentum. In *Spatial biases in perception and cognition* (pp. 121–138). Cambridge University Press. https://doi.org/10.1017/9781316651247.009

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, *44*(3), 211–221. https://doi.org/10.3758/BF03206290

Hubbard, T. L., & Ruppel, S. E. (2014). An Effect of Contrast and Luminance on Visual Representational Momentum for Location. *Perception*, *43*(8), 754–766. https://doi.org/10.1068/p7714

Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex: Population coding of position in visual cortex. *The Journal of Physiology*, *556*(3), 971–982. https://doi.org/10.1113/jphysiol.2003.058941

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: Spatial uncertainty is the key to perceptual mislocalization. *Vision Research*, *44*(22), 2605–2619. https://doi.org/10.1016/j.visres.2003.10.028

Kerzel, D., & Gegenfurtner, K. R. (2003). Neuronal Processing Delays Are Compensated in the Sensorimotor Branch of the Visual System. *Current Biology*, *13*(22), 1975–1978. https://doi.org/10.1016/j.cub.2003.10.054

Kleiner, M., Brainard, D. H., Pelli, D. G., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1–16.

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research*, *39*(16), 2669–2679. https://doi.org/10.1016/S0042-6989(98)00287-9

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Research*, *40*(2), 201–215. https://doi.org/10.1016/S0042-6989(99)00168-6

Kwon, O.-S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences*, *112*(26), 8142–8147. https://doi.org/10.1073/pnas.1500361112

López-Moliner, J., & Linares, D. (2006). The flash-lag effect is reduced when the flash is perceived as a sensory consequence of our action. *Vision Research*, *46*(13), 2122–2129. https://doi.org/10.1016/j.visres.2005.11.016

Makin, A. D. J., Stewart, A. J., & Poliakoff, E. (2009). Typical object velocity influences motion extrapolation. *Experimental Brain Research*, *193*(1), 137–142. https://doi.org/10.1007/s00221-008-1678-0

Maunsell, J. H., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology*, *68*(4), 1332–1344. https://doi.org/10.1152/jn.1992.68.4.1332

Maus, G. W., Khurana, B., & Nijhawan, R. (2010). History and theory of flash-lag: Past, present, and future. In R. Nijhawan & B. Khurana (Eds.), *Space and Time in Perception and Action* (1st ed., pp. 477–500). Cambridge University Press. https://doi.org/10.1017/CBO9780511750540.027

Maus, G. W., & Nijhawan, R. (2006). Forward displacements of fading objects in motion: The role of transient signals in perceiving position. *Vision Research*, *46*(26), 4375–4381. https://doi.org/10.1016/j.visres.2006.08.028

Maus, G. W., & Nijhawan, R. (2008). Motion Extrapolation Into the Blind Spot. *Psychological Science*, *19*(11), 1087–1091. https://doi.org/10.1111/j.1467-9280.2008.02205.x

Maus, G. W., & Nijhawan, R. (2009). Going, going, gone: Localizing abrupt offsets of moving objects. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 611–626. https://doi.org/10.1037/a0012317

Murakami, I. (2001). The flash-lag effect as a spatiotemporal correlation structure. *Journal of Vision*, *1*(2), 6. https://doi.org/10.1167/1.2.6

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*(6487), 256–256. https://doi.org/10.1038/370256a0

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences*, *6*(9), 387–393. https://doi.org/10.1016/S1364-6613(02)01963-0

Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences*, *31*(2), 179–198. https://doi.org/10.1017/S0140525X08003804

Nijhawan, R., Watanabe, K., Khurana, B., & Shimojo, S. (2004). Compensation of neural delays in visual-motor behaviour: No evidence for shorter afferent delays for visual motion. *Visual Cognition*, *11*(2–3), 275–298. https://doi.org/10.1080/13506280344000347

Nijhawan, R., & Wu, S. (2009). Compensating time delays with neural predictions: Are predictions sensory or motor? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *367*(1891), 1063–1078. https://doi.org/10.1098/rsta.2008.0270

Nover, H., Anderson, C. H., & DeAngelis, G. C. (2005). A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(43), 10049–10060. https://doi.org/10.1523/JNEUROSCI.1661-05.2005

Öğmen, H., Patel, S. S., Bedell, H. E., & Camuz, K. (2004). Differential latencies and the dynamics of the position computation process for moving targets, assessed with the flash-lag effect. *Vision Research*, *44*(18), 2109–2128. https://doi.org/10.1016/j.visres.2004.04.003

Owens, D. A., Wood, J., & Carberry, T. (2010). Effects of Reduced Contrast on the Perception and Control of Speed When Driving. *Perception*, *39*(9), 1199–1215. https://doi.org/10.1068/p6558

Palmer, S. E., Marre, O., Berry, M. J., & Bialek, W. (2015). Predictive information in a sensory population. *Proceedings of the National Academy of Sciences*, *112*(22), 6908–6913. https://doi.org/10.1073/pnas.1506855112

Pollen, D. A. (2003). Explicit Neural Representations, Recursive Neural Networks and Conscious Visual Perception. *Cerebral Cortex*, *13*(8), 807–814. https://doi.org/10.1093/cercor/13.8.807

Pretto, P., Bresciani, J.-P., Rainer, G., & Bülthoff, H. H. (2012). Foggy perception slows us down. *eLife*, *1*, e00031. https://doi.org/10.7554/eLife.00031

Priebe, N. J., & Lisberger, S. G. (2004). Estimating Target Speed from the Population Response in Visual Area MT. *The Journal of Neuroscience*, *24*(8), 1907–1916. https://doi.org/10.1523/JNEUROSCI.4233-03.2004

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature*, *396*(6710), 424–424. https://doi.org/10.1038/24766

Roach, N. W., McGraw, P. V., & Johnston, A. (2011). Visual Motion Induces a Forward Prediction of Spatial Pattern. *Current Biology*, *21*(9), 740–745. https://doi.org/10.1016/j.cub.2011.03.031

Rosenbaum, D. A. (1975). Perception and extrapolation of velocity and acceleration. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(4), 395–403. https://doi.org/10.1037/0096-1523.1.4.395

Sarich, D., Chappell, M., & Burgess, C. (2007). Dividing attention in the flash-lag illusion. *Vision Research*, *47*(4), 544–547. https://doi.org/10.1016/j.visres.2006.09.029

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, *79*(6), 3272–3278. https://doi.org/10.1152/jn.1998.79.6.3272

Schneider, M., Marquardt, I., Sengupta, S., Martino, F. D., & Goebel, R. (2019). *Motion Displaces Population Receptive Fields in the Direction Opposite to Motion* (p. 759183). bioRxiv. https://doi.org/10.1101/759183

Schwartz, G., Taylor, S., Fisher, C., Harris, R., & Berry, M. J. (2007). Synchronized Firing among Retinal Ganglion Cells Signals Motion Reversal. *Neuron*, *55*(6), 958–969. https://doi.org/10.1016/j.neuron.2007.07.042

Scocchia, L., Grosso, R. A., de'Sperati, C., Stucchi, N., & Baud-Bovy, G. (2009). Observer's control of the moving stimulus increases the flash-lag effect. *Vision Research*, *49*(19), 2363–2370. https://doi.org/10.1016/j.visres.2009.06.023

Shen, H., Shimodaira, Y., & Ohashi, G. (2003). Effects of temporal frequency on speed discrimination and perceived speed. *Proceedings of the International Joint Conference on Neural Networks, 2003.*, *1*, 188–193. https://doi.org/10.1109/IJCNN.2003.1223336

Smeets, J. B. J., & Brenner, E. (1995a). Perception and action are based on the same visual information: Distinction between position and velocity. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 19–31. https://doi.org/10.1037/0096-1523.21.1.19

Smeets, J. B. J., & Brenner, E. (1995b). Prediction of a moving target's position in fast goal-directed action. *Biological Cybernetics*, *73*(6), 519–528. https://doi.org/10.1007/BF00199544

Smith, A. T., & Edgar, G. K. (1990). The influence of spatial frequency on perceived temporal frequency and perceived speed. *Vision Research*, *30*(10), 1467–1474. https://doi.org/10.1016/0042-6989(90)90027-I

Snowden, R. J., Stimpson, N., & Ruddle, R. A. (1998). Speed perception fogs up as visibility drops. *Nature*, *392*(6675), 450–450. https://doi.org/10.1038/33049

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, *9*(4), 578–585. https://doi.org/10.1038/nn1669

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, *32*(8), 1535–1549. https://doi.org/10.1016/0042-6989(92)90209-2

Subramaniyan, M., Ecker, A. S., Patel, S. S., Cotton, R. J., Bethge, M., Pitkow, X., Berens, P., & Tolias, A. S. (2018). Faster processing of moving compared with flashed bars in awake macaque V1 provides a neural correlate of the flash lag illusion. *Journal of Neurophysiology*, *120*(5), 2430–2452. https://doi.org/10.1152/jn.00792.2017

Sundberg, K. A., Fallah, M., & Reynolds, J. H. (2006). A motion-dependent distortion of retinotopy in area V4. *Neuron*, *49*(3), 447–457. https://doi.org/10.1016/j.neuron.2005.12.023

The MathWorks Inc. (2017). *MATLAB version R2017b* [Computer software]. The MathWorks Inc. https://mathworks.com/

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, *22*(3), 377–380. https://doi.org/10.1016/0042-6989(82)90153-5

Thompson, P., Brooks, K., & Hammett, S. T. (2006). Speed can go up as well as down at low contrast: Implications for models of motion perception. *Vision Research*, *46*(6–7), 782–786. https://doi.org/10.1016/j.visres.2005.08.005

Treue, S., Snowden, R. J., & Andersen, R. A. (1993). The effect of transiency on perceived velocity of visual patterns: A case of “temporal capture.” *Vision Research*, *33*(5), 791–798. https://doi.org/10.1016/0042-6989(93)90198-6

Vaziri-Pashkam, M., & Cavanagh, P. (2008). Apparent speed increases at low luminance. *Journal of Vision*, *8*(16), 1–12. https://doi.org/10.1167/8.16.9

Vaziri-Pashkam, M., & Cavanagh, P. (2011). Effect of Speed Overestimation on Flash-Lag Effect at Low Luminance. *I-Perception*, *2*(9), 1063–1075. https://doi.org/10.1068/i0435

Vreven, D., & Verghese, P. (2005). Predictability and the Dynamics of Position Processing in the Flash-Lag Effect. *Perception*, *34*(1), 31–44. https://doi.org/10.1068/p5371

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience*, *1*(8), 656–657. https://doi.org/10.1038/3659

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research*, *40*(2), 137–149. https://doi.org/10.1016/S0042-6989(99)00166-2

Wojtach, W. T., Sung, K., Truong, S., & Purves, D. (2008). An empirical explanation of the flash-lag effect. *Proceedings of the National Academy of Sciences*, *105*(42), 16338–16343. https://doi.org/10.1073/pnas.0808916105

CHAPTER 3

When visual attention is divided in the flash-lag effect

This chapter is based on the work submitted to the *Journal of Vision* (Yook, Hogendoorn et al., in press). Formatting has been adapted to suit the thesis.

Abstract

The flash-lag effect (FLE) occurs when a flash's position appears delayed relative to a continuously moving object, even though both are physically aligned. While several studies have demonstrated that reduced attention increases FLE magnitude, the precise mechanism underlying these attention-dependent effects remains elusive. In this study, we investigated the influence of visual attention on the FLE by manipulating the level of attention allocated to multiple stimuli moving simultaneously in different locations. Participants were cued to either focus on one moving stimulus or split their attention among two, three, or four moving stimuli presented in different quadrants. We measured trial-wise FLE to explore potential changes in the magnitude of perceived displacement and its trial-to-trial variability under different attention conditions. Our results reveal that FLE magnitudes were significantly larger when attention was divided among multiple stimuli compared to when attention was focused on a single stimulus, suggesting that divided attention considerably augments the perceptual illusion. However, FLE variability, measured as the coefficient of variation, did not differ between conditions, indicating that the consistency of the illusion is unaffected by divided attention. We discuss the interpretations and implications of our findings in the context of widely accepted explanations of the FLE within a dynamic environment.

1. Introduction

Attention is crucial in how we process and represent information in our visual environment. For dynamic information, attention may be essential for updating and maintaining coherent representations of moving objects (Iordanescu et al., 2009; Kerzel, 2003). When attention is limited, however, perceptual biases and illusions, such as the displacement in an object's position due to its motion or the motion of other objects, can be significantly altered.

One such illusion is the flash-lag effect (FLE). The FLE occurs when a static flash appears next to a continuously moving object, leading to a perceived spatial offset between their positions despite their physical alignment at the moment of the flash (Nijhawan, 1994). Specifically, the flash is misperceived as lagging behind the moving object. Although several studies have demonstrated that the allocation of attention modulates the magnitude of the FLE, the precise mechanism underlying such attention-dependent effects remains elusive. Additionally, the relationship between attention and (in)variability of moving objects' perceived positions is not well understood.

Sarich et al. (2007) employed a dual-task paradigm to compare the magnitude of the FLE in a flash-lag task performed alone or concurrently with a target detection task. They found that the FLE, measured as a point of subjective equality, was smallest when the flash-lag task was performed separately or with a slight interval from the detection task. Notably, when another task required simultaneous attention, the magnitude of the FLE increased, and detection worsened. This complements observations regarding the phenomenon of representational momentum (Freyd & Finke, 1984), where the perceived final position of a moving object is shifted forward in the direction of anticipated motion and increases with divided attention (Hayes and Freyd, 2002).

Shioiri et al. (2010) manipulated attention by varying the number of moving stimuli and presenting the flash next to one of these stimuli. They observed an increase in the FLE when attention was divided among two or six stimuli and a decrease when attention was focused solely on one stimulus. Conversely, when participants were pre-cued to the upcoming locations of either the moving object or the flash, the FLE reduced following valid cues compared to invalid (Baldo & Klein, 2010; Namba & Baldo, 2004; Vreven & Verghese, 2005; but see Khurana et al., 2000) or no cues (Shioiri et al., 2010). This pattern was also reported regarding representational momentum (Hubbard et al., 2009).

However, in tasks involving multiple object tracking (Pylyshyn & Storm, 1988), which requires precise localization of targets among other moving objects, divided attention has been shown to increase errors in extrapolating predictable motion trajectories (Adamian & Andersen, 2022; Howe & Holcombe, 2012; Luu & Howe, 2015; Zhong et al., 2014). These varying effects of reduced attention across different illusions and paradigms underscore the complexity of how attention influences the representation of moving objects and the need to disentangle the specific mechanisms underlying these effects. Therefore, the present study investigated whether attention not only influences the strength of perceptual illusions like the FLE but also affects the quality of representation, such as spatial resolution (Anton-Erxleben & Carrasco, 2013; Barbot & Carrasco, 2017; Yeshurun & Carrasco, 1998, 1999).

To this end, we examined whether divided attention affects the magnitude or the trial-to-trial variability of the FLE using trial-wise spatial cues. Previous studies often used stimuli grouped as a single moving object (e.g., dots arranged in a line; Baldo & Klein, 1995; Krekelberg & Lappe, 1999) or multiple objects following the same motion trajectory (e.g., dots revolving in a circular path; Khurana et al., 2000; Shioiri et al., 2010), which could introduce grouping effects. In contrast, our study employed objects characterized by independent motion trajectories to minimize such effects.

In our novel experimental paradigm, participants viewed an array of identically appearing bars presented in different locations. A cue preceding each trial indicated which bar(s) participants should track. While all bars rotated simultaneously, participants were required to covertly track the moving bar in the cued quadrant(s) of the display. Importantly, the perceptual configuration remained the same whether attention was focused on a single bar or divided across two, three, or four bars. During each trial, a target flash was presented next to one of the bars, and participants were instructed to indicate the position of the corresponding bar at the time of the flash. We compared the effects of divided attention to focused attention on both FLE magnitude and consistency (trial-to-trial variability of the FLE).

2. Materials and Methods

2.1 Participants

Twenty-six healthy adults were recruited from the Forschungszentrum Jülich to participate in the experiment. Two participants were excluded due to insufficient data quality (see **section 2.5** for details), resulting in data from 24 participants (15 female, 9 male, age [M: 28.8 years, SD: 4 years], range 21–40 years) being included in the final analysis. The sample size was determined by a preliminary power calculation for a desired medium effect size (Cohen's *f* = 0.25) with a power of 80% and an alpha level of 0.05 for a repeated-measures analysis of variance (ANOVA). All participants were fluent in spoken and written English, had normal or corrected-to-normal visual acuity, and had no neurological or psychiatric disorders. Handedness was not a selection criterion; participants self-reported their handedness preferences (19 right-handed, 5 left-handed). All participants provided informed written consent and received compensation of 10 Euros per hour for their participation. This study was conducted in accordance with the Declaration of Helsinki and approved by the ethics committee of the German Psychological Society (DGPs) (Ethics ID: 2022-02-03VA).

2.2 Apparatus

Visual stimuli were generated using Psychopy 3.0 (Peirce, 2007, 2009) and presented on a 22-inch Samsung SyncMaster 2233RZ LCD monitor with a resolution of 1680 x 1050 pixels and a refresh rate of 120 Hz (as outlined in Wang & Nikolic, 2011). A chinrest was used to stabilize the head and maintain a viewing distance of 70 cm, with the center of the screen approximately at eye level. A standard QWERTY keyboard was positioned below the chinrest and not visible from the participants' field of view, with the left hand on the space bar and the right hand on the arrow keys. The experiment was conducted in a soundproof, light-attenuated room.

2.3 Stimuli

A white fixation cross, subtending 0.5° x 0.5° of visual angle (dva), was continuously displayed at the center of a black background. The spatial cue consisted of one, two, three, or four yellow dots (radius = 0.15 dva) presented at the corners of the central fixation cross, indicating the quadrant(s) where the target flash could occur. Every possible combination of dots was realized (e.g., for two quadrants, two dots on the left, right, upper, lower, or diagonal quadrants).

The moving stimuli consisted of four gray bars, each rotating smoothly around the center of its respective quadrant at an angular velocity of 240°/s. Each bar measured 0.13 x 6.04 dva, with its inner end positioned at 9.98 dva from the central fixation. On one end of each bar was a small dot (radius = 0.13 dva) used to indicate the bar's position relative to the target flash. The target was a red dot of the same size (radius = 0.13 dva) and was always positioned at 0.8 dva from the bar's dot end.

2.4 Procedure

In each trial, participants were instructed to fixate on the central fixation cross for the duration of the stimulation sequence (**Figure 1A**). Attention cues were displayed for 600 ms, pointing to the quadrant(s) to be attended. These cues pointed to the quadrant but not the specific position within the quadrant where the target was likely to appear in the upcoming display.

After a 500-ms interval, allowing participants to covertly orient their attention to the cued quadrant(s), an array of four rotating bars was presented around the central fixation cross. The bars rotated in either a clockwise or counter-clockwise direction, alternating randomly across trials. Each bar started from a unique orientation, chosen from one of 16 uniformly distributed orientations. This configuration allowed the bars to rotate along independent motion trajectories despite their shared direction of motion.

Periodically after 850, 1000, or 1150 ms (chosen at random), the target briefly flashed (25 ms) next to one of the bars. After the target offset, the bars continued to rotate for an additional 1150, 1000, or 850 ms (for a total duration of 2000 ms). Participants' task was to indicate the instantaneous position of the cued bar at the time of the flash on a circular response probe.

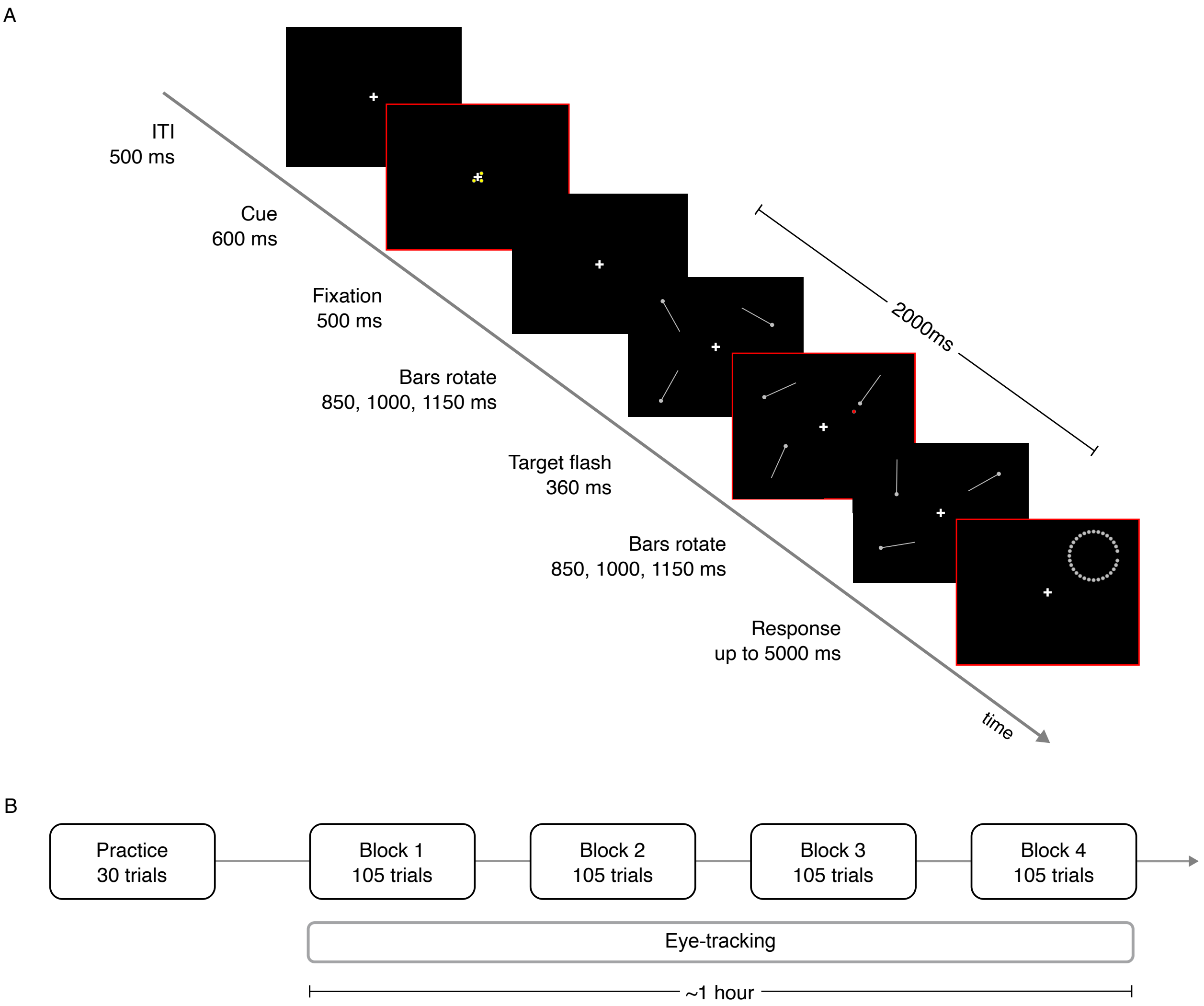

Figure 1. Experimental paradigm. (**A**) Trial sequence. Spatial attention cues were presented at the central fixation cross to indicate that the participants should orient their attention to the upcoming display, featuring an array of four rotating bars (not drawn to scale) in corresponding quadrants. The target flashed next to a bar after half the time of the overall sequence. Here, a target is presented at a cued quadrant, but sometimes the target appeared at a non-cued quadrant (see text for details). Following the sequence, a response ring was presented until participants made a response. The subsequent trial began after an inter-trial interval of 500 ms. Each trial lasted up to 8 s. (**B**) Procedure timeline. Participants performed at least 30 practice trials for task familiarization, followed by four blocks of 105 trials, totaling 420 trials over approximately one hour, including eye-tracking. Each block comprised trials from all four attention conditions and every combination of cue-target configurations with equal probability. The order of trials was randomized within the block.

The response probe comprised a ring of 60 small grey dots (each dot radius = 0.13 dva) – whose diameter was equal to the length of a bar. To perform the task, participants navigated the ring using the left and right arrow keys to select their perceived position of the bar's dot end displayed on the ring.

The starting position on the ring randomly varied from trial to trial, which introduced an inherent variability in how quickly participants could accurately respond, resulting in longer response times (RTs) when the distance from the desired position was greater. While RT is a common measure in similar paradigms (see Khurana et al., 2000), we did not analyze RTs for the current experiment. Instead, participants were instructed to respond within a 5000-ms window, focusing on accuracy rather than speed. A response was registered when the enter key was pressed. The subsequent trial began after an inter-trial interval (ITI) of 500 ms.

Importantly, since the target and its corresponding bar were presented in alignment, the perceived displacement of the stimulus positions in the direction of motion, or the lack thereof, directly determined the FLE for each trial. This method contrasts with the conventional approach of averaging offsets based on the point of subjective equality across blocks within a specific condition (cf. Yook et al., 2022).

The probe location corresponded to the quadrant where the target had been presented in the previous display, eliminating the need for participants to additionally report the target's quadrant in each trial. However, the cue-target validity varied across trials. Specifically, the target appeared in the cued quadrant(s) 90% of the time. In the remaining 10% of trials, termed "catch trials", the target appeared in a non-cued quadrant, and participants were instructed to press the space bar without indicating the bar's position with the response probe. Catch trials ensured participants' attention to the cues. Incorrect responses, such as responding using the probe in catch trials or pressing the space bar in non-catch trials, prompted feedback, "Please ATTEND TO THE CUES as they are helpful", at fixation for 1000 ms. No feedback on behavioral performance was given otherwise.

The task was organized into four blocks, each consisting of 105 trials (96 non-catch trials and 9 catch trials; **Figure 1B**), totaling 420 trials conducted over approximately one hour. Each trial lasted up to 8 s. Every possible combination of cue-target configurations appeared within each block with equal probability. Participants underwent four attention conditions, and the order of trials was randomized within each block. To minimize fatigue, participants were instructed to take regular self-paced breaks between blocks and between every ~5–10 mins within a given block, as indicated by a break screen.

Before the main experiment, participants completed at least 30 practice trials to familiarize themselves with the task. Upon completion of the experiment, participants were debriefed about the study's purpose related to the FLE. None of the participants reported being aware that the positions of the attended bar and target were aligned.

2.5 Data quality assessment

Data from participants who failed to identify at least 75% of catch trials were excluded from the analysis to mitigate potential non-compliance with task instructions and task artifacts. This criterion ensured participants' adherence to attending to the cued bar(s), as instructed. Data from two participants who detected an average of 2.5 catch trials (SD: 3.5) out of 36 throughout the experiment were excluded. Additionally, incorrectly responded non-catch trials were excluded. Among the remaining 24 participants, who on average detected 33.3 catch trials (SD: 2.4), data from an average of 379 correct non-catch trials (min: 371, max: 384, SD: 3.8) were available for analysis.

2.6 Behavioral analysis

Behavioral data were pre-processed using Python 3.5 (Van Rossum & Drake, 2009) within the Anaconda environment (Anaconda Inc., 2016) and analyzed in R Statistical Software v4.3.1 (R Core Team, 2023) with custom scripts.

During each trial, participants were instructed to direct their attention towards a single moving bar (focused attention condition) or spread their attention between two, three, or four bars (divided attention conditions: attend-to-two, -three, or -four). The FLE for each trial was determined by indexing the magnitude of the difference between the actual position of the target and the participant's chosen position of the bar. A value of zero indicated no perceived difference between the positions of the two stimuli. Positive values indicated a flash-lag effect, where the moving bar was perceived as ahead of the flash. Conversely, negative values indicated a flash-lead effect, where the flash was perceived as ahead of the moving bar.

2.6.1 Assessment of magnitude

The FLE values were aggregated to compute the median FLE magnitudes for each condition of each participant. We then compared these medians between the focused and divided attention conditions using a repeated-measures ANOVA using the four-level factor *attention*, with the ez package (Lawrence, 2016) in R. Following the ANOVA, we performed paired-sample *t*-tests to test the hypotheses that there would be significant differences in FLE magnitudes between focused attention and divided attention conditions, as well as between the various divided attention conditions

To account for multiple comparisons, statistically significant differences were determined based on a threshold of $p = \frac{0.05}{\mathrm{n - rank + 1}}$ (one-tailed), adjusting for familywise error rate (FWE) using the Holm-Bonferroni correction (Holm, 1979). Greenhouse-Geisser corrections were applied when Mauchly's test indicated a violation of sphericity.

To handle potential outliers, FLE values exceeding 3 standard deviations from the mean FLE values were identified and excluded. Analyses were performed both including and excluding outliers; however, no significant differences were observed when outliers were excluded. Therefore, the reported results include outliers, with descriptive statistics shown in **Supplementary Materials, *Figure S1.1*** to provide a more comprehensive insight into the distribution of the FLE values. To validate our findings against non-normality assumptions, we employed a Friedman's test as a non-parametric alternative. The results of the Friedman's test aligned with those of the repeated-measures ANOVA.

2.6.2 Assessment of variability

In order to examine whether the distribution of FLE magnitudes became more variable with divided attention, the same procedure was repeated using the coefficient of variation (CV), calculated by dividing the standard deviation by the absolute mean of each condition of each participant. This allowed for comparing trial-to-trial variability across participants that vary widely in magnitude, by offering a relative measure of variability for comparing individuals.

2.7 Eye-tracking recording and analysis

We monitored eye movements using an infrared EyeLink 1000 Plus (SR Research Ltd., Mississauga, Canada) system at a sampling rate of 1000 Hz to evaluate how well participants had maintained fixation during the stimulation sequence of each trial. At the beginning of the experiment, the eye tracker was calibrated with a 5-point calibration procedure to establish an accurate gaze position of the left eye. Acceptable calibration values had to meet the validation criterion of < 0.5 dva and maximum error of < 1.5 dva. Due to technical issues, eye-tracking data was not recorded for 12 participants, and two additional datasets were incomplete, yielding 10 analyzable datasets.

Eye-tracking data were pre-processed and analyzed in R using the eyelinker package (Barthelme, 2021) and custom scripts. Fixations and saccades were determined from the raw gaze position data using Eyelink's default event parser. A fixation was defined as an event lasting at least 100 ms, allowing deviation of 1-dva radius from the central fixation cross. Saccades were identified using a velocity threshold of 30°/s and acceleration threshold of $8000°/s^2$. Eye-blink events were excluded.

Across attention conditions, we examined the proportion of fixation time participants spent on the central fixation cross relative to the overall time during each of three distinct phases of a trial, wherein participants were instructed to maintain fixation:

1. During the cue presentation (cue) lasting 600 ms,
2. Immediately after the cue offset (orient) lasting 500 ms, and
3. Before and after the target presentation (target), with a combined duration of 2000 ms.

Differences in the proportion of fixation between the different attention conditions were analyzed with a repeated-measures ANOVA using the four-level factor *attention*.

2.8 Data and code availability

All experiment and analysis scripts are publicly available at https://osf.io/5pn9s/.

3. Results

3.1 Fixation controlled under all attention conditions

The proportions of fixation time across various phases of the experiment (cue, orient, and target) are presented in **Figure 2**.

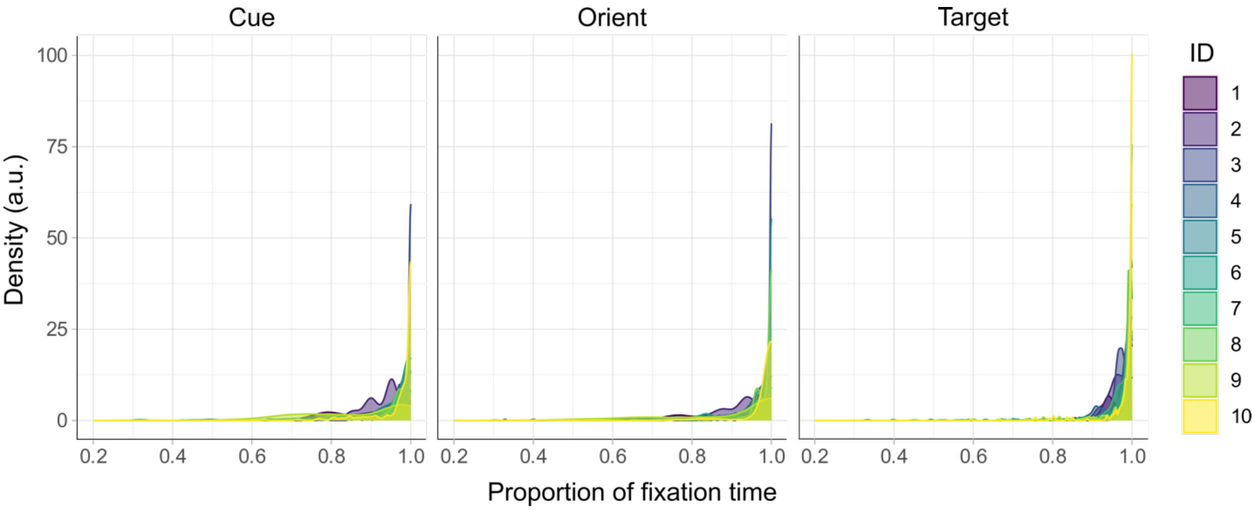

Figure 2. Individual proportions of fixation time across various phases of the experiment (cue, orient, and target), collapsed across attention conditions. Here shown are 10 participants whose eye-tracking data met the criteria for data quality. Note the fixation duration may be influenced by interruptions such as blinks, signal fallout, or saccades. These events were initially identified and removed ahead of the fixation control analysis.

We performed a repeated-measures ANOVA to compare the proportion of fixation between different attention conditions within each phase. There were no significant differences in fixation across cue (one-way ANOVA, *attention*: $F(3, 27) = 0.7$, $p = 0.5$), orient (one-way ANOVA, *attention*: $F(3, 27) = 0.5$, $p = 0.7$), or target phases (one-way ANOVA, *attention*: $F(3, 27) = 0.4$, $p = 0.8$). Overall, these results indicate that participants consistently maintained a high degree of central fixation, regardless of whether their attention was focused on a single quadrant or divided across multiple quadrants.

3.2 Attentional modulation of FLE magnitude, but not trial-to-trial variability

Participants reported the perceived position of a moving bar, which was compared to its physical position. This task assessed the flash-lag effect (FLE), where the difference between the physical and perceived (reported) position of the bar at the time of the flash indicates the magnitude of the illusion. We analyzed FLE magnitudes, calculated as the average of individual medians, across different attention conditions, as shown in **Figure 3A**.

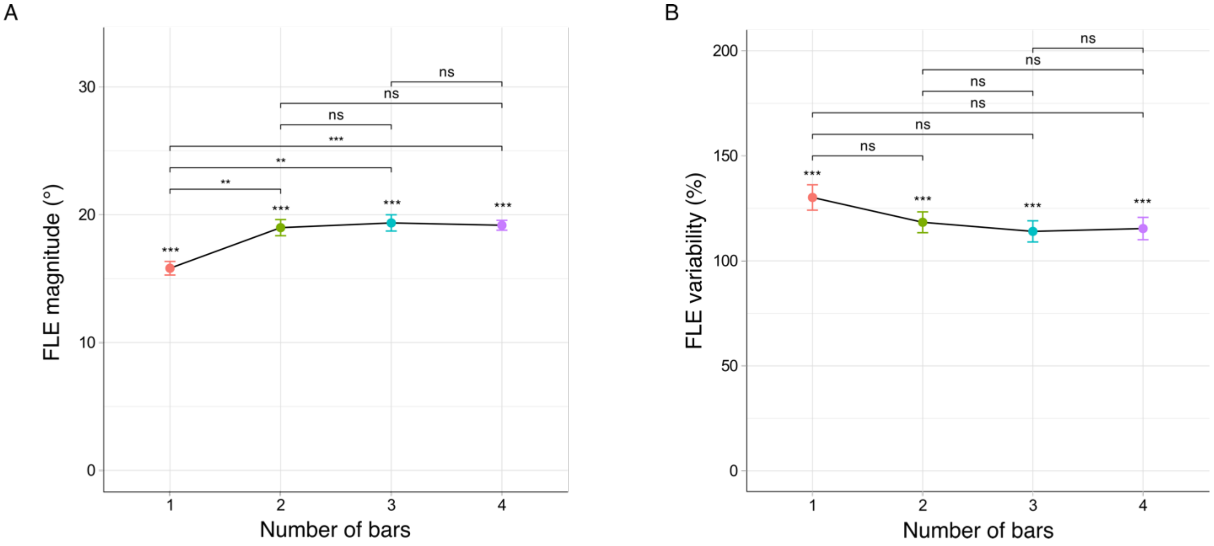

Figure 3. FLE results. (**A**) FLE magnitudes as a function of number of attended moving bars. All values are reported as means ± within-subject standard error of the mean (S.E.M., error bars). ns (not significant) denotes $p > 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$. (**B**) The same described in **A** but for FLE variability (coefficient of variation).

The illusion was prominent across all attention conditions, averaging between 15-20° (see **Supplementary Materials, *Table S1.2***). Overall, one-way ANOVA revealed a significant effect of attention on the FLE magnitude ($F(3, 69) = 9.2$, $p < 0.001$). Specifically, the mean FLE magnitudes were significantly larger in divided attention conditions (two to four bars) compared to the focused attention condition (one bar), as detailed in **Table 1**.

Table 1.

Paired-sample *t*-tests (one-tailed) results with correction for multiple comparisons for FLE magnitude.

Pair	FLE Magnitude	
	t_{23}	p
1 vs 2	-3.9	< 0.01

1 vs 3	-4.3	< 0.01
1 vs 4	-5.2	< 0.001
2 vs 3	-0.4	1
2 vs 4	-0.3	1
3 vs 4	0.3	1

Figure 3B (see also **Supplementary Materials, S2**) depicts a contrasting pattern in trial-to-trial FLE variability, as measured as the coefficient of variation (CV), with a larger CV in focused attention than in divided attention conditions. However, no significant differences emerged across conditions (one-way ANOVA, *attention*: $F(3, 69) = 1.9$, $p = 0.1$). Notably, the CVs exceeding 100% indicate a high degree of variability of FLE values relative to their mean, suggesting that the FLE values varied widely within participants (see also **Supplementary Materials, Figure *S1.1***).

4. Discussion

In this study, we aimed to investigate how divided attention affects the perceived position of a moving object compared to a physically aligned static flash within the context of the FLE. The FLE is a well-documented illusion where a flash, when presented next to a moving object, appears to lag behind the moving object despite both being physically aligned. We found that FLE magnitude was augmented under divided attention conditions compared to focused attention. However, FLE variability across trials did not differ significantly between any of the attention conditions.

Participants viewed an array of four independently moving objects in distinct quadrants of the display during each trial. Attention was manipulated using spatial cues that directed participants to either focus on a single quadrant or to distribute attention across two, three, or four quadrants. This design allowed us to systematically compare FLE magnitude and trial-to-trial variability under different attentional loads. To ensure accurate allocation of attention, we included catch trials to verify that participants adhered to the provided cues. Subsequently, excluding catch responses from the analysis was crucial for differentiating between genuine attentional effects and potential confounding effects from incorrect cue compliance (see details below).

Our results revealed a clear effect of attention on the magnitude of the FLE. When attention was divided among multiple moving objects in different quadrants, FLE magnitudes increased significantly compared to when attention was focused on a single object. This finding aligns with previous studies (Namba & Baldo, 2004; Sarich et al., 2007; Shioiri et al., 2010) and is strengthened by our larger sample size (N = 24 compared to 15, 14, and 5, respectively) and a trial-wise measure of the FLE, which provides a more precise and nuanced understanding of the illusory effect. Interestingly, FLE variability did not show significant differences across the various attention conditions. These results suggest that while divided attention in the FLE amplifies the FLE magnitude, it does not necessarily lead to greater variability in representations from trial to trial compared to focused attention. However, it is important to note that the individual distributions of the FLE values contributed to large coefficients of variation (CV).

Moreover, the increase in FLE magnitude was significant only when shifting from one object to multiple objects, with no further increases when attention was distributed among two, three, or four objects. This pattern is consistent with Shioiri et al.'s (2010) findings, which involved up to six dots arranged in a circle. By contrast, Hogendoorn et al. (2010) observed a gradual increase in FLE magnitude, measured as time, with a step-wise increase in FLE variability from one to two clocks, with no further increases beyond that. This suggests a saturation effect at the condition where attention was focused on two objects, indicating that further dividing attention did not significantly impact the FLE under highly divided-attention conditions (Howe & Holcombe, 2012; Luu & Howe, 2015).

Two possible explanations arise from these observations. Firstly, as proposed by Shiori et al. (2010), the saturation effect may relate more to the spatial spread of attention than to the number of attended objects. For instance, Khurana et al. (2000) observed no changes in FLE magnitudes when one of two possible flash locations (either above or below the fixation point) was cued, similar to our study. This suggests that the effect of attention may not follow a straightforward pattern but is highly dependent on task complexity (Khurana et al., 2000; Shioiri et al., 2010). Secondly, attentional resources might be constrained by hemifield (Alvarez & Cavanagh, 2005; Luck et al., 1989) or even by quadrants (Carlson et al., 2007, 2011). If each quadrant has a limited attentional capacity, further dividing attention might not affect the representation of additional moving objects in extra quadrants. This could account for the deviations noted in our analysis of anisotropies (see **Supplementary Materials, *S3***) and the discrepancies observed in Khurana et al. (2000), where no additional attentional modulation was found compared to other FLE studies considered here. Future research could explore these mechanisms of attentional saturation by manipulating perceptual load, such as varying the number of stimuli (> 1) within each quadrant or adjusting their eccentricities (Baldo et al., 2002). Additionally, exploring the impact of motion speed (Shioiri et al., 2010) could further provide insights into whether increased attention is needed to track faster motion (cf. Yook et al., 2022), potentially influencing FLE magnitude to a greater extent than observed here.

Our findings align with several prominent theoretical accounts of the FLE, including theories of sequential processing including temporal integration (Krekelberg & Lappe, 1999, 2000), discrete sampling perception (Schneider, 2018), attention shifting (Baldo & Klein, 1995), and postdiction (Brenner & Smeets, 2000; Eagleman & Sejnowski, 2000). For specific details of each of these theories, readers are referred to Hogendoorn (2020), Holcombe and Corbett (2023), Hubbard (2014), and Schneider (2018). Arguably, when attention is divided, the brain has fewer resources to process each object’s motion trajectory, reducing information processing speed. This may contribute to increased delays in the visual system’s ability to extract relevant motion and position information for each object, as the brain needs to switch between multiple objects, including the flash. This delay in processing time results in a more considerable latency difference between the moving objects and the flash, augmenting the FLE.

Under the predictive motion explanation theory (Nijhawan, 1994), these compounded delays across multiple objects may necessitate more compensation, thereby increasing the FLE magnitude. The brain might overcompensate if it prioritizes efficiency over accuracy when processing multiple moving objects under dynamic conditions, such as in this experiment. In principle, this would allow the brain to simultaneously monitor and update representations of multiple objects, even when sensory information is delayed. According to this interpretation, the attention-dependent effects of perceptual biases and illusions outlined in the **Introduction** can be reconciled. When attention is reduced in representational momentum and the FLE, the visual system may rely more heavily on existing predictions rather than on slowly-arriving sensory input, enhancing the perceived forward displacement in moving objects. In contrast, reduced attention results in decreased extrapolation during multiple object tracking. This occurs because the task requires continuous and precise tracking of several objects, and the brain cannot allocate enough resources to accurately predict each object's motion trajectory, leading to less effective extrapolation. Hence, the impact of reduced or divided attention varies depending on the nature of the task. While our findings are open to interpretation, we believe our findings contribute to this body of evidence by revealing the role of attention in efficiently processing multiple moving objects in dynamic environments. By contrast, our findings do not seem to align with the differential processing latencies theory (Whitney et al., 2000; Whitney & Murakami, 1998), which would predict that as moving objects are processed faster than static objects, divided attention would lead to a smaller processing latency difference between the two and a smaller FLE magnitude.

Taken together, these results suggest that, in the context of the FLE, attention primarily influences the rate at which events are processed within the visual system, rather than the quality of processing. While overall perception is certainly impacted by the quality of representation, the FLE appears to be more a result of how the visual system effectively manages highly attention-demanding and dynamic environments. This suggests that when attentional resources are divided among multiple moving objects, delays in accessing relevant information may accumulate (Carrasco & McElree, 2001; Giordano et al., 2009), leading to slower recognition of motion and delayed availability of position information for each object.

An alternative interpretation is that attention affects the anticipatory processing of the flash (Baldo et al., 2002; Vreven & Verghese, 2005), potentially reducing the latency difference between the flash and the moving object when attention is focused. Although it is challenging to rule out an effect of divided attention on the flash, this alone is unlikely to account for our results. This interpretation would also predict increased FLE magnitudes with divided attention, but it assumes that participants might anticipate the flash even without explicit cues. Previous studies (Baldo et al., 2002; Sarich et al., 2007; Shioiri et al., 2010; Vreven & Verghese, 2005) may have been confounded by this. However, our study’s use of catch trials reduces the likelihood that participants were simply waiting for the flash. If participants had been passively anticipating the flash, then they would have reported the FLE regardless of the attentional cues. Our data show that participants did notice the flash when it appeared next to a non-cued bar, suggesting that they did not merely wait for the flash on every trial. While our analyses concentrated on correctly responded non-catch trials, future research could include a separate condition for catch trials. A lack of discernable difference in FLE magnitudes between catch and non-catch conditions would strongly suggest that the changes observed were primarily driven by the flash itself. This would offer further insights into how attention affects the representation of moving objects rather than the detection of the flash.

In conclusion, the experiment reported here is the first to manipulate attention in the flash-lag paradigm combining conventional attentional cueing with divided attention procedures, alongside trial-wise FLE readouts. Our results demonstrate that the magnitude of motion-induced illusory perceptual effects varies with the level of attention in dynamic environments.

Supplementary Materials for When visual attention is divided in the flash-lag effect

Jane Yook *et al*.

Supplementary Results

S1. Descriptive statistics of FLE magnitude

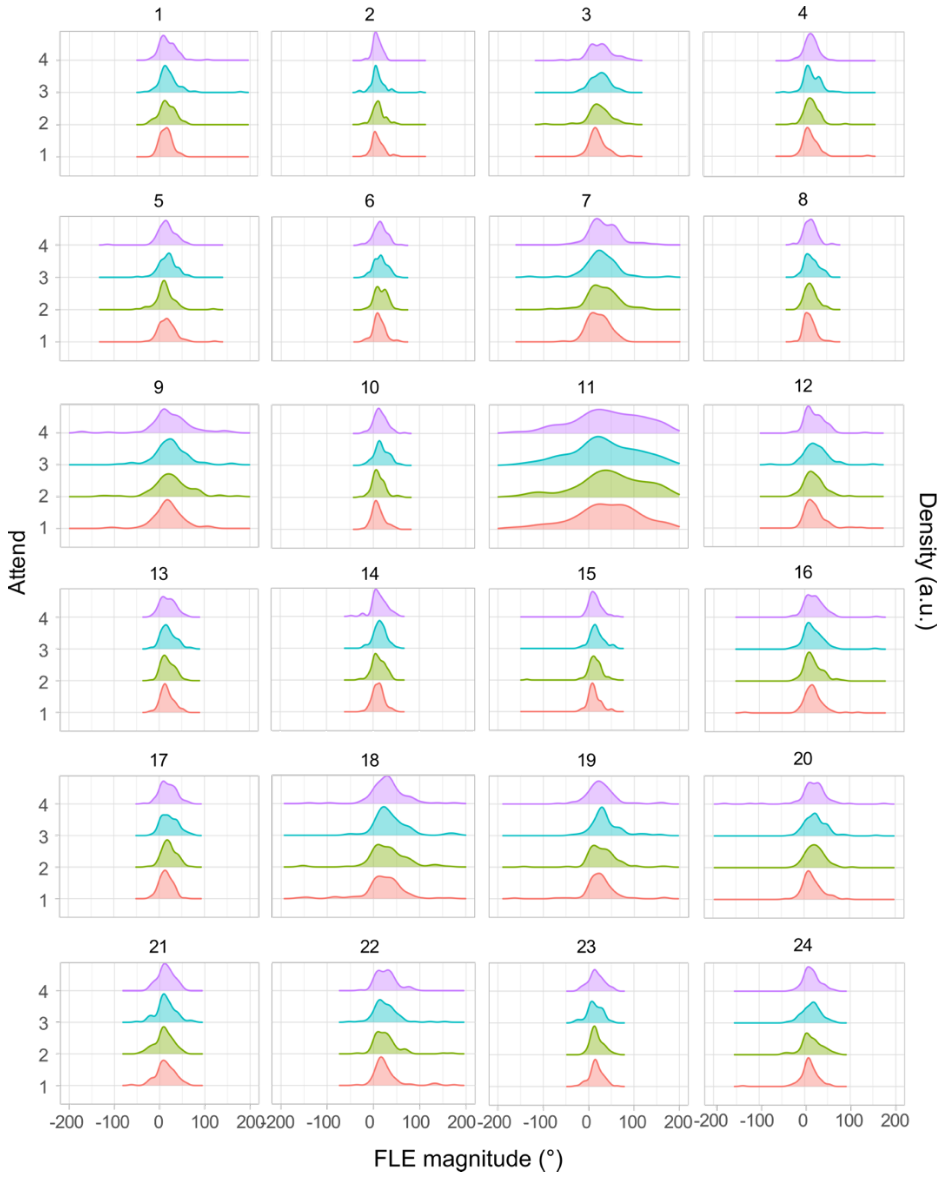

Figure S1.1. Individual distributions of FLE magnitudes across attention conditions.

Table S1.2.
One-sample *t*-test (one-tailed) results against zero.

Bars	*M*	*SD*	t_{23}	*p*
1	15.8°	6.4°	12.1	< 0.001
2	19.0°	9.0°	10.3	< 0.001
3	19.4°	6.5°	14.6	< 0.001
4	19.2°	7.2°	13.0	< 0.001

S2. Descriptive statistics of FLE variability

Table S2.

One-sample *t*-test (one-tailed) results against zero.

Bars	*M*	*SD*	t_{23}	*p*
1	130.2%	45.3%	14.1	< 0.001
2	118.4%	36.2%	16.0	< 0.001
3	114.0%	32.2%	17.3	< 0.001
4	115.4%	32.2%	17.6	< 0.001

S3. Hemifield effects in the FLE

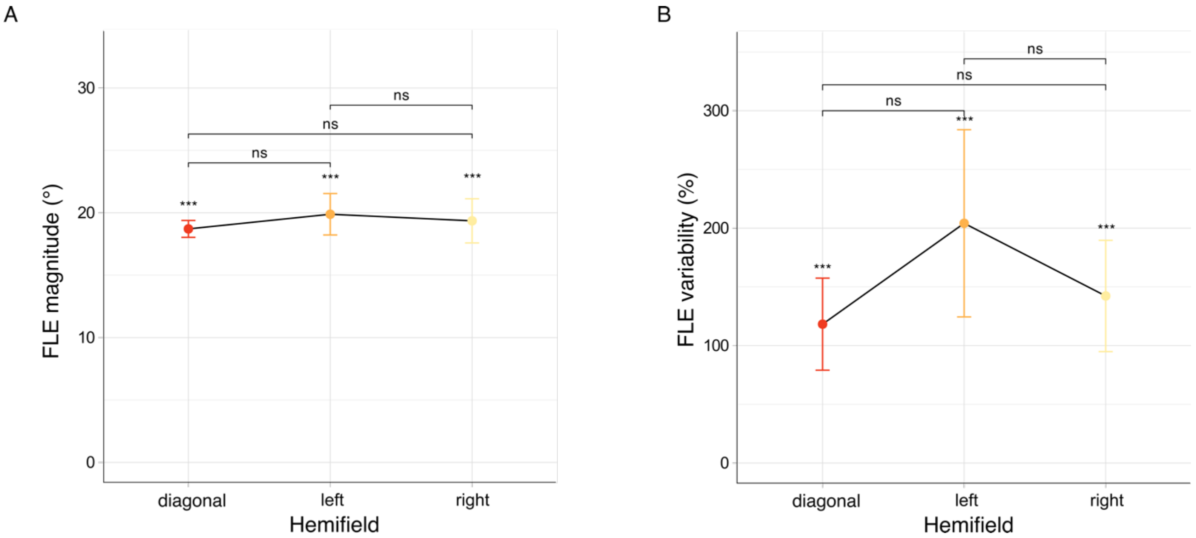

Figure S3. Hemifield results. Previous studies using linear motion have demonstrated anisotropies in the left and right sides of the visual field (Kanai et al., 2004; Shi & Nijhawan, 2008; Suzuki et al., 2023). Notably, the FLE tends to be more substantial when stimuli are presented in the left hemifield. Here, we aimed to test whether the effect of divided attention could be pronounced with these anisotropies, particularly within the attend-to-two condition. To achieve this, we examined the magnitude and variability of the FLE on trials where participants were cued to two quadrants within their left, right, or diagonally across the upper and lower visual hemifields. (**A**) The FLE magnitudes are plotted as a function of the to-be-attended hemifields. All values are reported as means ± within-subject standard error of the mean (S.E.M., error bars). ns (not significant) denotes $p > 0.05$. Although the FLE magnitude in the left hemifield was marginally higher than in both the right or diagonal hemifields, no significant modulation of attention was detected (one-way ANOVA, *hemiField:* $F(2,46) = 0.2$, $p = 0.7$). (**B**) The same as described in **A** but for FLE variability (one-way ANOVA, *hemiField:* $F(2,46) = 0.6$, $p = 0.5$).

References

Adamian, N., & Andersen, S. K. (2022). Attentional Enhancement of Tracked Stimuli in Early Visual Cortex Has Limited Capacity. *The Journal of Neuroscience*, *42*(46), 8709–8715. https://doi.org/10.1523/JNEUROSCI.0605-22.2022

Alvarez, G. A., & Cavanagh, P. (2005). Independent Resources for Attentional Tracking in the Left and Right Visual Hemifields. *Psychological Science*, *16*(8), 637–643. https://doi.org/10.1111/j.1467-9280.2005.01587.x

Anaconda Software Distribution (Version 2-2.4.0). (2016). [Computer software]. Anaconda, Inc. https://www.anaconda.com/

Anton-Erxleben, K., & Carrasco, M. (2013). Attentional enhancement of spatial resolution: Linking behavioural and neurophysiological evidence. *Nature Reviews. Neuroscience*, *14*(3), 188–200. https://doi.org/10.1038/nrn3443

Baldo, M. V. C., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an Attentional Component of the Perceptual Misalignment between Moving and Flashing Stimuli. *Perception*, *31*(1), 17–30. https://doi.org/10.1068/p3302

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature*, *378*(6557), Article 6557. https://doi.org/10.1038/378565a0

Baldo, M. V. C., & Klein, S. A. (2010). Paying attention to the flash-lag effect. In R. Nijhawan & B. Khurana (Eds.), *Space and Time in Perception and Action* (1st ed., pp. 396–407). Cambridge University Press. https://doi.org/10.1017/CBO9780511750540.023

Barbot, A., & Carrasco, M. (2017). Attention Modifies Spatial Resolution According to Task Demands. *Psychological Science*, *28*(3), 285–296. https://doi.org/10.1177/0956797616679634

Barthelme, S. (2021). *eyelinker: Import ASC Files from EyeLink Eye Trackers* (Version 0.2.1) [Computer software]. https://cran.r-project.org/web/packages/eyelinker/index.html

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash–lag effect. *Vision Research*, *40*(13), 1645–1648. https://doi.org/10.1016/S0042-6989(00)00067-5

Carlson, T. A., Alvarez, G. A., & Cavanagh, P. (2007). Quadrantic deficit reveals anatomical constraints on selection. *Proceedings of the National Academy of Sciences*, *104*(33), 13496–13500. https://doi.org/10.1073/pnas.0702685104

Carlson, T. A., Cho, H., Turret, J., & Dakin, S. (2011). Psychoanatomy of visual attention: A unified account of quadrant and hemifield effects. *Journal of Vision*, *11*(11), 101–101. https://doi.org/10.1167/11.11.101

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences*, *98*(9), 5363–5367. https://doi.org/10.1073/pnas.081074098

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion Integration and Postdiction in Visual Awareness. *Science*, *287*(5460), 2036–2038. https://doi.org/10.1126/science.287.5460.2036

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126–132. https://doi.org/10.1037/0278-7393.10.1.126

Giordano, A. M., McElree, B., & Carrasco, M. (2009). On the automaticity and flexibility of covert attention: A speed-accuracy trade-off analysis. *Journal of Vision*, *9*(3), 30. https://doi.org/10.1167/9.3.30

Hayes, A. E., & Freyd, J. J. (2002). Representational momentum when attention is divided. *Visual Cognition*, *9*(1–2), 8–27. https://doi.org/10.1080/13506280143000296

Hogendoorn, H. (2020). Motion Extrapolation in Visual Processing: Lessons from 25 Years of Flash-Lag Debate. *The Journal of Neuroscience*, *40*(30), 5698–5705. https://doi.org/10.1523/JNEUROSCI.0275-20.2020

Hogendoorn, H., Carlson, T. A., Vanrullen, R., & Verstraten, F. A. J. (2010). Timing divided attention. *Attention, Perception, & Psychophysics*, *72*(8), 2059–2068. https://doi.org/10.3758/BF03196683

Holcombe, A. O., & Corbett, J. (2023). *Temporal errors: Researchers should stop studying the flash-lag effect*. https://doi.org/10.31234/osf.io/swzr7

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Howe, P. D. L., & Holcombe, A. O. (2012). Motion information is sometimes used as an aid to the visual tracking of objects. *Journal of Vision*, *12*(13), 10–10. https://doi.org/10.1167/12.13.10

Hubbard, T. L. (2014). The flash-lag effect and related mislocalizations: Findings, properties, and theories. *Psychological Bulletin*, *140*(1), 308–338. https://doi.org/10.1037/a0032899

Hubbard, T. L., Kumar, A. M., & Carp, C. L. (2009). Effects of spatial cueing on representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 666–677. https://doi.org/10.1037/a0014870

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of velocities during multiple-object tracking. *Journal of Vision*, *9*(4), 1–1. https://doi.org/10.1167/9.4.1

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: Spatial uncertainty is the key to perceptual mislocalization. *Vision Research*, *44*(22), 2605–2619. https://doi.org/10.1016/j.visres.2003.10.028

Kerzel, D. (2003). Attention maintains mental extrapolation of target position: Irrelevant distractors eliminate forward displacement after implied motion. *Cognition*, *88*(1), 109–131. https://doi.org/10.1016/S0010-0277(03)00018-0

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The Role of Attention in Motion Extrapolation: Are Moving Objects 'Corrected' or Flashed Objects Attentionally Delayed? *Perception*, *29*(6), 675–692. https://doi.org/10.1068/p3066

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research*, *39*(16), 2669–2679. https://doi.org/10.1016/S0042-6989(98)00287-9

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Research*, *40*(2), 201–215. https://doi.org/10.1016/S0042-6989(99)00168-6

Lawrence, M. (2016). *ez: Easy Analysis and Visualization of Factorial Experiments* (Version 4.4-0) [R]. https://CRAN.R-project.org/package=ez

Luck, S. J., Hillyard, S. A., Mangun, G. R., & Gazzaniga, M. S. (1989). Independent hemispheric attentional systems mediate visual search in split-brain patients. *Nature*, *342*(6249), 543–545. https://doi.org/10.1038/342543a0

Luu, T., & Howe, P. D. L. (2015). Extrapolation occurs in multiple object tracking when eye movements are controlled. *Attention, Perception, & Psychophysics*, *77*(6), 1919–1929. https://doi.org/10.3758/s13414-015-0891-8

Namba, J., & Baldo, M. V. C. (2004). The Modulation of the Flash-Lag Effect by Voluntary Attention. *Perception*, *33*(5), 621–631. https://doi.org/10.1068/p5212

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*(6487), 256–256. https://doi.org/10.1038/370256a0

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*. https://www.frontiersin.org/articles/10.3389/neuro.11.010.2008

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197. https://doi.org/10.1163/156856888x00122

R Core Team. (2023). *R: A Language and Environment for Statistical Computing* (Version 4.3.1) [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org/

Sarich, D., Chappell, M., & Burgess, C. (2007). Dividing attention in the flash-lag illusion. *Vision Research*, *47*(4), 544–547. https://doi.org/10.1016/j.visres.2006.09.029

Schneider, K. A. (2018). The Flash-Lag, Fröhlich and Related Motion Illusions Are Natural Consequences of Discrete Sampling in the Visual System. *Frontiers in Psychology*, *9*, 1227. https://doi.org/10.3389/fpsyg.2018.01227

Shi, Z., & Nijhawan, R. (2008). Behavioral significance of motion direction causes anisotropic flash-lag, flash-drag, flash-repulsion, and movement-mislocalization effects. *Journal of Vision*, *8*(7), 24. https://doi.org/10.1167/8.7.24

Shioiri, S., Yamamoto, K., Oshida, H., Matsubara, K., & Yaguchi, H. (2010). Measuring attention using flash-lag effect. *Journal of Vision*, *10*(10), 10–10. https://doi.org/10.1167/10.10.10

Suzuki, Y., Atmaca, S., & Laeng, B. (2023). The lateralized flash-lag illusion: A psychophysical and pupillometry study. *Brain and Cognition*, *166*, 105956. https://doi.org/10.1016/j.bandc.2023.105956

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Vreven, D., & Verghese, P. (2005). Predictability and the Dynamics of Position Processing in the Flash-Lag Effect. *Perception*, *34*(1), 31–44. https://doi.org/10.1068/p5371

Wang, P., & Nikolic, D. (2011). An LCD Monitor with Sufficiently Precise Timing for Research in Vision. *Frontiers in Human Neuroscience*, *5*. https://www.frontiersin.org/articles/10.3389/fnhum.2011.00085

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience*, *1*(8), 656–657. https://doi.org/10.1038/3659

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research*, *40*(2), 137–149. https://doi.org/10.1016/S0042-6989(99)00166-2

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*(6706), Article 6706. https://doi.org/10.1038/23936

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks1Parts of this study were presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology (May 1997) and at the Annual Meeting of the Psychonomics Society (November 1997) and published in Abstract format (Yeshurun and Carrasco, 1997and Carrasco and Yeshurun, 1997, respectively).1. *Vision Research*, *39*(2), 293–306. https://doi.org/10.1016/S0042-6989(98)00114-X

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2022). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. *Vision Research*, *193*, 107978. https://doi.org/10.1016/j.visres.2021.107978

Zhong, S. -h., Ma, Z., Wilson, C., Liu, Y., & Flombaum, J. I. (2014). Why do people appear not to extrapolate trajectories during multiple object tracking? A computational investigation. *Journal of Vision*, *14*(12), 12–12. https://doi.org/10.1167/14.12.12

CHAPTER 4

Meaningful objects: Selective visual attention guides predictive motion extrapolation

Abstract

Predictive motion extrapolation allows observers to predictively encode the position of moving objects along their trajectories, aligning perception with real-time and enhancing our interaction with dynamic environments. An intriguing question is how motion extrapolation relates to selective attention, which is crucial for prioritizing sensory information. Previous studies have demonstrated that the neural representation of an object's position is influenced by predictions about its future position. Thus far, evidence for extrapolation has only been reported for explicitly tracked stimuli. It is unknown whether extrapolation involves attention or extends to visible but unattended moving stimuli, i.e., not explicitly tracked. To investigate this possibility, we used a novel paradigm that could enable the spatial positions of two competing stimulus streams to be independently decoded from their evoked electroencephalographic (EEG) activity. We recorded EEG while human observers performed a color detection task that required attending to a task-relevant stream and ignoring a concurrently present distractor stream. In each stream, an apparent motion stimulus appeared every 100 ms, each following an unrelated circular trajectory. This led to a superposition of neural activity from both streams, containing information about their current and prior positions and putatively their future stimulus positions at each time-point. We performed time-resolved pattern classification to track these representations during ongoing motion sequences in each stream. Here we show that attention modulates neural representations yet the positions of an ignored stimulus remain represented. When the sequence stops, the representations are more selective, with information about its future position significantly fading for the ignored stimulus. This selectivity appears to bias competition to prioritize perceiving and predicting motion for the task-relevant stimulus.

1. Introduction

The progression of visual information takes time, introducing an inherent latency of approximately 100 ms from retinal stimulation to the emergence of perception (Maunsell & Gibson, 1992; Schmolesky et al., 1998). While neural processing delays of this length would not be problematic for static stimuli, they pose a challenge for dynamic stimuli as they constantly change during this interval. This is especially true when interacting with moving objects since the perception would lag behind an object's veridical position.

However, convergent evidence from animals and humans demonstrate that predictive mechanisms allow, at multiple levels of the processing hierarchy, for the neural coding of objects in or near real-time positions despite neural delays (Berry et al., 1999; Blom et al., 2021; Burkitt & Hogendoorn, 2021; Cheong et al., 2012; Hogendoorn & Burkitt, 2019; Jancke et al., 2004; Johnson et al., 2023; Kaplan et al., 2013; Maus et al., 2013; Palmer et al., 2015; Schneider et al., 2019; Subramaniyan et al., 2018; Sundberg et al., 2006). These predictions are reflected in patterns of neural activity containing predictive information about an object's current or subsequent position during apparent motion (Agostino et al., 2023; Hogendoorn & Burkitt, 2018; Robinson et al., 2020; Teichmann et al., 2021, 2022). Remarkably, these anticipatory patterns persist even in the absence of afference sensory information (Blom et al., 2020), indicating that extrapolation is an obligatory consequence of motion processing.

An important question is how motion extrapolation relates to attention. Our environment inundates us with streams of visual input, yet only a subset can be processed due to the capacity limitations on attention and perception. While some studies suggest that unattended information can still be represented (Mangun & Hillyard, 1990; Sahan et al., 2019; Volpe et al., 1979), other studies using fMRI and EEG decoding show that attended information undergoes more extensive processing than unattended or ignored information (Cohen et al., 2015; Noah et al., 2023; Yi et al., 2004). These findings could imply that visual information that is not attentionally selected might be lost before extrapolation. Moreover, the notion that attention may be involved in motion extrapolation is supported by observations that spatial attention itself moves predictively (Hogendoorn et al., 2007; Shioiri et al., 2002). In prior studies involving apparent motion, there was an implicit assumption that the stimulus was attentionally tracked (Agostino et al., 2023; Blom et al., 2020; Hogendoorn & Burkitt, 2018; Robinson et al., 2020; Teichmann et al., 2021). This raises the question of whether extrapolation entails attention to motion, or if it can occur even when motion receives less or no attention.

In a recent study by Ekman et al. (2017), observers were presented with sequences of stimuli along a horizontal path. Interestingly, the activation observed in response to the first stimulus presentation (out of 4) triggered a similar activity wave for the entire motion sequence, irrespective of whether observers were attending to the stimulus or not. This raises the possibility that, to some extent, extrapolation could operate pre-attentively (although it could also be argued that this finding is more in line with course anticipation related to memory than prediction). However, as with prior studies, it is difficult to determine the role of attention with a single stimulus.

Therefore, we investigated the neural measures of motion extrapolation, as explored by Blom et al. (2020), using an apparent motion paradigm with two stimulus streams. We employed time-resolved multivariate pattern analyses (MVPA) to analyze EEG data obtained from 25 human observers engaged in a color detection task. In this task, apparent motion and color changes were independently present within both the task-relevant (to-be-attended) and distractor (to-be-ignored) streams. Therefore, perceptual load was high, and selective attention was required for successful behavioral performance. Our goal was to assess motion extrapolation in each stream by examining whether the apparent motion of a stimulus triggered predictions about its future position, thereby activating sensory-like representations. We hypothesized that if attention contributes to motion extrapolation, then evidence of extrapolation would only manifest for the attended stimulus and remain absent for the ignored stimulus. If attention is not involved in motion extrapolation, then we expected to observe evidence for both stimuli. Alternatively, considering the potential depletion of processing resources under demanding task conditions (Cosman & Vecera, 2010), we expected a third possibility where neither stimulus could demonstrate extrapolation.

2. Materials and Methods

2.1 Participants

25 younger healthy individuals (17 female, 8 male, age [M: 26.4 years, SD: 4 years]) completed all experimental requirements and were included in the final sample. To be included in the study, observers had to be 18–40 years, right-handed, fluent in spoken and written English, have normal or corrected-to-normal vision, and no neurological or psychiatric disorders. Observers were invited to complete two experimental sessions on two separate days, with at least one day between sessions, and were reimbursed 15 AUD per hour for their participation. 21 additional observers were excluded for failing to return for a second session (n = 13, during the SARS-CoV-2 (COVID-19) pandemic), or insufficient data quality (n = 8; see **section 2.8** for details). All observers provided written informed consent prior to the start of the experiment. This study was approved by the Human Research Ethics Committee of Melbourne School of Psychological Sciences of the University of Melbourne (Ethics ID: 1954628.4) and was conducted in accordance with the Declaration of Helsinki.

2.2 Apparatus and stimuli

All stimuli were generated and displayed using Psychophysics Toolbox 3 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) with MATLAB R2021a (The MathWorks Inc., 2021), on a HP EliteDesk 800 G3 TWR with an LCD screen (24.5-inch ASUS ROG PG258; resolution: 1920 × 1080 pixels; refresh rate: 120 Hz). Observers were positioned approximately 60 cm from the screen and viewed the stimuli from a chinrest in a dimly lit EEG chamber.

On a gray background, trapezoid-shaped stimuli (i.e., wedges) were displayed on a notional circle (radius = 7° of visual angle (dva)) centered at the white fixation dot (radius = 0.27 dva). Each wedge had an outer width of 1.4 dva, and its inner and outer edges were 6.3 and 7.7 dva respectively from the fixation point. Wedges were presented in streams of either black or white stimuli, and each stream contained infrequent targets in either purple (instead of black) or green (instead of white). Up to two wedges could be displayed simultaneously (one from each stream). Stimuli were shown in 16 possible angular positions on the circle, separated by 22.5°, with white wedges shown in odd positions and black wedges shown in even positions to ensure that simultaneously displayed wedges would not have a spatial overlap. Relative to the vertical axis, the white wedges could be displayed at 0°, 45°, 90°, 135°, 180°, 225°, 270°, or 315°, and black wedges could be displayed at 22.5°, 67.5°, 112.5°, 157.5°, 202.5°, 247.5°, 292.5°, or 337.5° (**Figure 1A**).

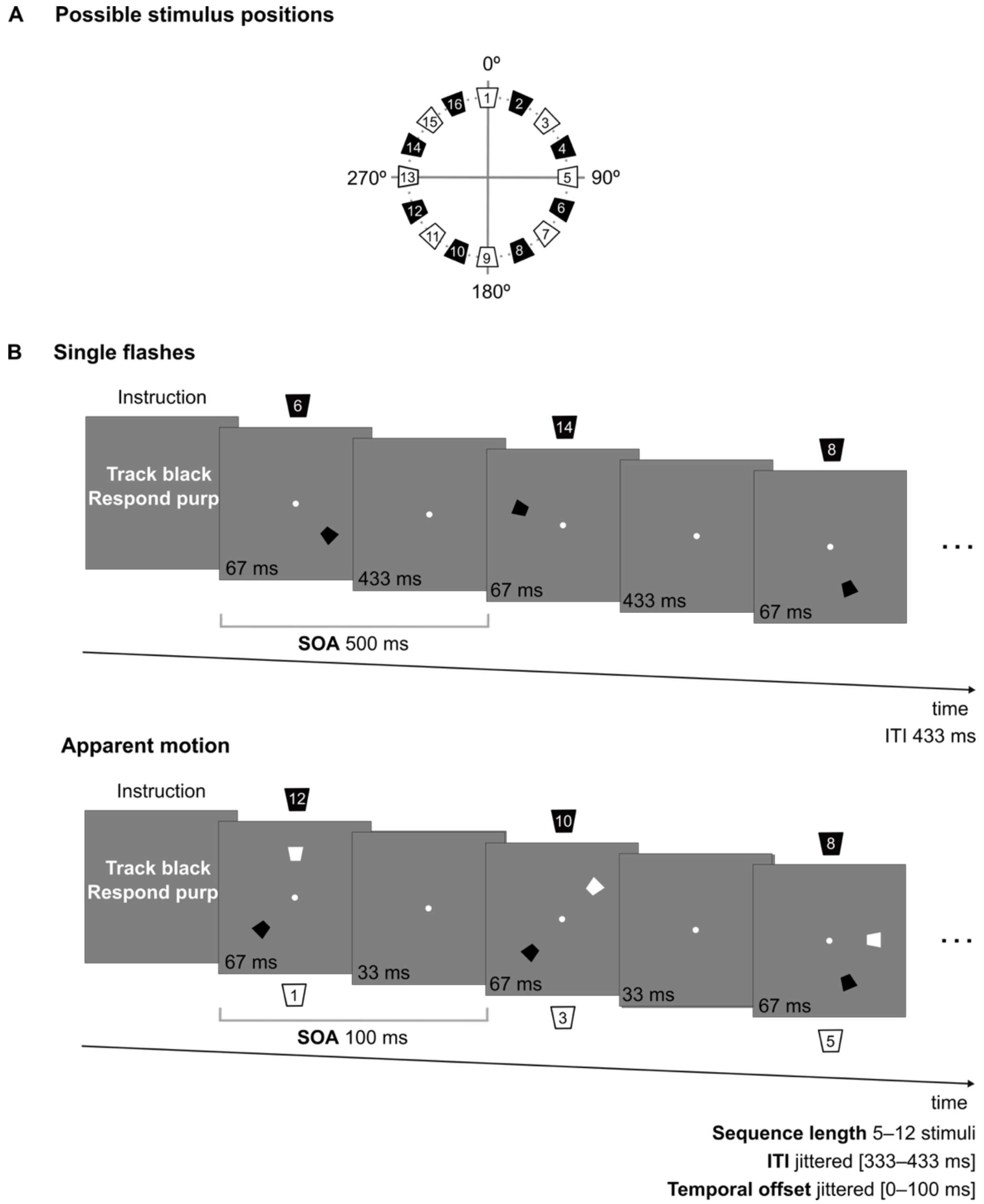

Figure 1. Stimuli and experimental paradigm. (**A**) Wedges were displayed on a notional circle around central fixation at a radius of 7 dva. The positions of each wedge were sampled from a set of 16 possible angular positions based on stimulus color white (odd positions) and black (even positions). (**B**) Stimulus timing for the Single flashes (*top*) and Apparent motion (*bottom*) conditions. In the Single flashes condition, a single wedge was displayed at eight positions in a random order. In the Apparent motion condition, two independent streams of wedges were presented concurrently. Within each stream, each wedge was displayed at successive positions in a predictable sequence order (see text for details). Each motion sequence lasted for 5–12 stimuli (unknown to the observer). Here shown are simultaneously displayed wedges, but the wedges

were displayed with variable temporal offset between [0–100 ms] across sequences. In both conditions, observers were instructed to track a wedge and report any target color change such as green while tracking white, or purple while tracking black.

2.3 Experimental paradigm

Observers performed a color detection task within a dynamic visual environment, where wedges were displayed at variable and rapidly changing positions but with an implicit regularity that was of consequence to the task.

The task was organized into blocks (**Figure 1B**). Each block began with a screen instructing observers to track a wedge of a specific color (e.g., black) across its variable positions over time and detect when its color changed to the target color (e.g., purple), for which they had 1 s to respond. Wedges with other colors (e.g., white, green), if present, were task-irrelevant distractors to be ignored. The offset of the instruction screen triggered by a keypress was followed by a continuous stream of briefly visible wedges, with variable colors and positions. Observers were instructed to maintain fixation throughout and report whenever a color change occurred by pressing the spacebar on a keyboard with their right hand. The block ended with a screen requesting observers to take a short break before they initiated the next block. No feedback on behavioral performance was provided.

The blocks differed in (1) the proportion of wedges with task-relevant colors (100% versus 50%) and (2) the variability of wedge positions (random positions versus along a motion trajectory). These factors varied the demands on selective attention and the predictability of wedge positions. The two tested scenarios are referred to as the Single flashes and Apparent motion conditions. While the Apparent motion condition was the primary focus of the experiment, the Single flashes condition served as a stimulus localizer to train multivariate pattern classifiers to recognize patterns of neural activity associated with stimuli in each position. These classifiers were subsequently tested on the Apparent motion condition (see **section 2.6**). Each condition was separated into two blocks based on the instructed color pair to be tracked/detected (black-purple and white-green). All four blocks were presented once in a randomized order in each of the two experimental sessions. Each session lasted a total of 1.5 hours.

2.3.1 Single flashes condition

A single wedge was displayed for 67 ms with a fixed stimulus onset asynchrony (SOA) of 500 ms. The wedge's position was selected randomly (uniformly) from the eight possible positions for the instructed color pair (**Figure 1A**). Only one stream of stimuli was presented in a given block (i.e., either black or white, with their corresponding targets). Within a block (~8 min), 800 stimuli were presented (8 positions x 100 repetitions), and 50 of these stimuli were randomly selected to be the target stimuli.

2.3.2 Apparent motion condition

Two independent stimulus streams were presented concurrently, one black and one white. Within each stream, a single wedge was displayed for 67 ms with an SOA of 100 ms. The positions of the wedge were ordered into sequences so that it would be perceived as being in apparent motion along a circular trajectory. A motion sequence began with a wedge's starting position (uniformly) randomly selected from the eight possible color-specific positions. Each successive wedge position was shown 45° clockwise or counterclockwise from the prior position, which corresponded to an angular speed of 450°/s (i.e., 1.25 rotations/s) in the sequence-specific direction. The motion sequence ended after a variable length of 5–12 stimuli (unknown to the observer), followed by a variable blank interval of [333–433 ms] before the start of the next sequence. Target colors were presented in both streams, but observers had to respond only to the instructed target color (i.e., green while tracking white, purple while tracking black). The other target color, serving as a distractor, did not require any response and was to be ignored. The Apparent motion condition therefore contained motion sequences differing only in the instructed attended color, which was counterbalanced across blocks.

The two stimulus streams had the same SOA but were otherwise independent as the motion sequences (defined by starting position, motion direction, and sequence length) were independently and randomly selected for each stream in a block. Moreover, the relative onset of the two streams was jittered by [0–100 ms], so depending on this jitter, the two wedges of different colors could be displayed simultaneously for a variable duration of up to 66 ms. Note each sequence could also be considered equivalent to a statistically independent "trial" in a typical trial-based experimental design. Within a block, 1200 sequences were presented for each stream (8 starting positions x 2 directions x 8 lengths x 8-10 randomly sampled repetitions), and 50 of these sequences were randomly selected to contain a single target stimulus. Observers viewed a total of ≈ 20350 wedges (≈ 10175 wedges per stream) spanned over ~30 min (three sets of ~10 min with a short break between each set).

2.4 Data acquisition and pre-processing

2.4.1 EEG acquisition

EEG activity was recorded using a standard 64-electrode BioSemi ActiveTwo system (BioSemi, Amsterdam, The Netherlands) and digitized at a sampling rate of 2048 Hz. Electrodes were arranged in a spherical array layout according to the International 10–20 system. Two external electrodes were placed on the mastoids behind the ears for offline re-referencing. Electrooculographic activity (EOG) was recorded with six ocular electrodes (three electrodes placed above, below, and at the outer cathus of each eye). During the recording, all data were referenced to the electrode Cz and filtered using a bandpass [0.16–100 Hz] filter. All electrode impedances were kept below 20 kΩ.

2.4.2 EEG pre-processing

EEG data were pre-processed in MATLAB R2021a using EEGLAB v14.1.1b (Delorme & Makeig, 2004) and custom scripts. Continuous EEG data were down-sampled to 512 Hz (one data point per ~2 ms) and re-referenced to the average activity of the mastoid electrodes. The data were then visually inspected to reject time periods with artifacts related to abnormal voltage fluctuations, muscle activity, signal noise, and artifacts noted during the recording. Bad channels were identified by visual inspection, and their signals were replaced with signals from neighboring channels using spherical spline interpolation, as implemented by EEGLAB. All datasets had less than 6 bad channels per session, which was our pre-defined quality threshold.

Because our analysis employed time-resolved pattern analysis methods to make time-related inferences (**section 2.6**), we avoided using artifact correction methods that might distort the temporal structure of the EEG signals. Specifically, no further filtering and standard eye-blink correction, such as Independent Component Analysis (ICA) (Makeig et al., 2004) or regression-based methods (Croft & Barry, 2000), were applied to the continuous data. This decision was made to prevent introducing phase distortions to the data, a concern noted by Van Driel et al. (2021) and Widmann et al. (2015). We instead chose to reject epochs with large voltage changes (including eye-blinks), having the least possible impact on the signals even if they were potentially correctable (see **section 2.4.4** for details).

2.4.3 Epoch definition

The continuous data from the Single flashes condition was segmented into epochs from [-200 ms to +500 ms] relative to the onset of each wedge stimulus. This time range was chosen based on the condition’s SOA of 500 ms, such that epochs would not contain additional stimulus events apart from the reference stimulus. The baseline for the Single flashes epochs was the interval [-200 to 0 ms] prior to stimulus onset. The continuous data from the Apparent motion condition was also segmented into epochs from [-200 ms to +500 ms] relative to the onset of each wedge stimulus of each stimulus stream. This means each epoch contained activity related to a motion sequence, rather than a single reference stimulus. Due to the short SOA of 100 ms, an epoch could contain up to eight stimulus events from the same sequence, including the reference stimulus. To retain sequence-specific information, the baseline for each Apparent motion epoch was the 200-ms interval immediately prior to the first stimulus onset of the corresponding sequence. Note that an epoch could also capture up to eight events from the other, concurrently presented, stimulus stream; however, the other-stream events were not systematically related to the reference stimulus across epochs as the two streams were statistically independent. Epochs of both conditions that contained target stimuli (purple or green, irrespective of instruction) and responses were excluded. All epochs were baseline corrected by subtracting the mean voltages over the corresponding baseline from the voltages at all other time-points of that epoch.

2.4.4 Epoch rejection

Epochs with signal standard deviations (estimated across all channels and time-points) greater than 80 μV were rejected. The 80 μV threshold was heuristically selected to be conservative while avoiding a large data loss (also see Blom et al., 2020; Hogendoorn & Burkitt, 2018). With this threshold, an average of 8.8% (SD: 11.36%) of Single flashes epochs were rejected per observer. This rejection threshold was not applied to Apparent motion epochs to prevent discarding complex but relevant activity patterns. Finally, all remaining epochs from both conditions were visually inspected to confirm the exclusion of epochs with atypical voltage changes.

2.5 Analysis rationale

In each Apparent motion block, observers would have had to relate the successive positions of wedges of the same color over time, in order to (1) perceive two separate streams with distinctive motion-based regularities and (2) allocate more processing resources to the task-relevant stream over the distractor stream. Therefore, the two streams were analyzed separately based on their instructed relevance, referred to as the Attended and Ignored streams.

We evaluated the evidence for motion extrapolation from each stream based on whether the apparent motion of a wedge triggered predictions about its future position. Since observers were not instructed to make explicit position predictions, this had to be indirectly inferred from the measured neural activity. In each epoch, we tested whether the neural activity following the reference stimulus at a position (P) contained information about the position (P_{+1}) of the next, as yet unseen stimulus in the motion sequence, i.e., a putative representation of the wedge's future position. The analysis was conducted separately for the Attended and Ignored streams using the spatiotemporal decoding approach described by Blom et al. (2020) (see **Supplementary Materials, *S1***), following these steps:

1. Neural responses to single flashes were characterized, separately for black and white Single flashes epochs. This allowed us to train sets of pairwise classifiers to distinguish the activity trained evoked by each pair of the color-matched positions separated by 45°, 90°, 135°, or 180°.

2. Motion sequences were extracted, again separately for black and white Apparent motion epochs, time-locked to stimulus presentations within ongoing sequences (i.e., subsequent stimuli were presented immediately after, referred to as ONGOING) and to the final stimulus presentation in each sequence (i.e., a subsequent stimulus was expected but not presented, referred to as STOP).

3. Classifiers trained on Single flashes in the first step were tested on all Apparent motion (ONGOING and STOP) epochs. As per Blom et al., each epoch corresponding to P was tested using classifiers trained to discriminate between its preceding position (P_{-1}) and its subsequent position $P_{+1.}$ This was done separately for each combination of training time-point and testing time-point, generating a Temporal Generalization Matrix (TGM) (Carlson et al., 2011; King & Dehaene, 2014).

4. Apparent motion epochs were separated into Attended and Ignored streams and collapsed across colors, yielding a 2 x 2 set of conditions ({ONGOING, STOP} x {Attended, Ignored}).

These steps were performed for individual observers (first-level analysis, **section 2.6**), and the per-observer outcomes were aggregated for group-level statistical testing (second-level analysis, **section 2.7**). Due to the many assumptions involved with the spatiotemporal decoding approach, our analyses included tests to validate these assumptions.

2.6 First-level analysis

A schematic of the first-level analysis pipeline is shown in **Figure 2**.

2.6.1 Data selection

We selected relevant epochs of the appropriate stream and positions from the Single flashes condition and the Apparent motion condition (**Figure 2A**).

2.6.1.1 Single flashes condition

At each time-point, the neural activity of any Apparent motion epoch comprised a superposed activity evoked by stimuli at multiple positions (see **Supplementary Materials, *S1***). Therefore, it was essential to distinguish activity specifically related to P_{+1} from activity related to P and other preceding positions in the sequence, such as $P_{-1.}$ Therefore, we chose P_{-1} as a discriminative position for $P_{+1,}$ as the features that discriminate P_{+1} from P_{-1} should also discriminate P_{+1} from P in the Apparent motion condition (P_{+1} and P_{-1} are 90° apart; Blom et al., 2020). All epochs corresponding to P_{+1} and P_{-1} as the reference stimuli were used for subsequent analyses (**Figure 2A, *top***).

2.6.1.2 Apparent motion condition

Our analysis focused on the stimulus transition from one position to its subsequent position in a motion sequence (i.e., $P \rightarrow P_{+1}$). To ensure that observers had sufficient context to perceive P in apparent motion, we specifically examined transitions where P was preceded by at least four stimuli, meaning P was at least the 5th presented stimuli in the sequence (i.e., $P_{-1} \rightarrow P \rightarrow P_{+1}$). Another important configuration included transitions where P did not proceed to P_{+1} at the end of the sequence (i.e., $P_{-1} \rightarrow P \rightarrow [P_{+1}]$). Here, $[P_{+1}]$ signifies that P_{+1} was predictable from the motion but was *not* physically presented. These configurations correspond to ONGOING (i.e., P_{+1} present) and STOP (i.e., P_{+1} absent), respectively. All epochs corresponding to P as the reference stimulus in the ONGOING and STOP configurations of each stream were used for subsequent analyses (**Figure 2A, *bottom***).

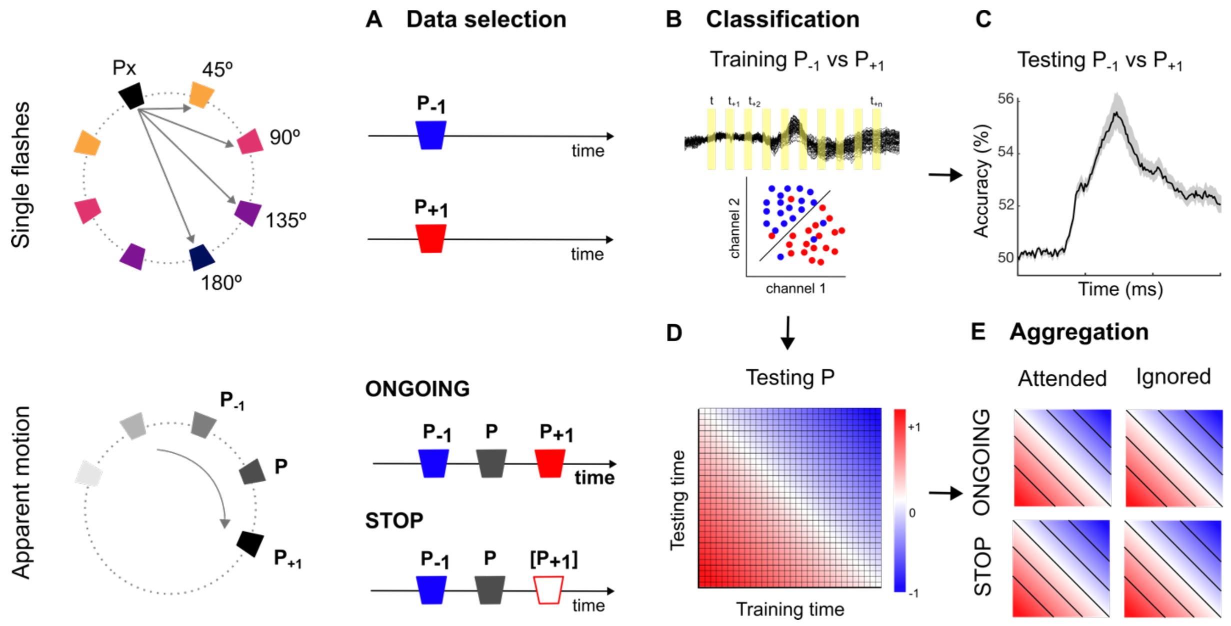

Figure 2. First-level analysis pipeline. (**A**) (*top*) From the Single flashes condition, epochs corresponding to the positions P_{+1} and P_{-1} were selected. (*bottom*) From the Apparent motion condition, ONGOING epochs corresponding to the midpoint of the motion sequence (i.e., $P_{-1} \rightarrow P \rightarrow P_{+1}$) and STOP epochs corresponding to the end of the motion sequence (i.e., $P_{-1} \rightarrow P \rightarrow [P_{+1}]$) were selected. (**B**) At each time-point within an epoch, samples of activity across all EEG channels were extracted (only two channels shown for simplicity) to train a classifier to distinguish between samples of P_{+1} and P_{-1}. Two ensembles of classifiers were used for testing the classifier's decision of independent testing samples for the (**C**) validation analysis and (**D**) main analysis). Classification assignments were evaluated as classification accuracy or classification bias (for P_{+1} or P_{-1}) in a temporal generalization matrix (TGM). (**E**) The classification procedure was repeated for all pairwise combinations of positions and conditions. The individual classification results were aggregated for second-level analyses (see text).

2.6.2 Classification

We then applied pattern classification to the selected epochs (**Figure 2B**).

2.6.2.1 Training

The Single flashes epochs for P_{+1} and P_{-1} were used to train an ensemble of time-resolved binary classifiers. These classifiers were trained using the Linear Discriminant Analysis (LDA) machine learning algorithm (Fisher, 1936), as implemented in the Statistics and Machine Learning Toolbox (The MathWorks Inc., 2023) in MATLAB. LDA was chosen for its simple algorithm and common use in decoding research (Grootswagers et al., 2017).

At each time-point within an epoch, the instantaneous activity across all 64 channels was extracted. These time-specific activity samples, constituting one 64-feature vector per epoch, were used to train a linear binary classifier to discriminate between two classes, P_{+1} and P_{-1}. To ensure that the classifier was not influenced by the over-representation of any specific class (position), the training sets from the two classes were always balanced by randomly excluding samples from the larger class. No normalization was applied to the samples. The outcome of the training was a classifier (or decision-rule) capable of assigning whether a particular activity sample was evoked by P_{+1} or P_{-1}, even if the sample was not seen during training.

One such classifier was generated per time-point across [0–500 ms]. We generated two different ensembles of classifiers depending on the sampling rate of the epochs. For validation checks, we used the original timeline with samples every ~2 ms, resulting in an ensemble of 251 distinct classifiers. For our main analyses, a more fine-grained sampling was obtained using a smoothed timeline. The sample at time-point t_{train} (in 1 ms steps) from an epoch represented the channel-wise mean of all samples over every [t_{train} ± 10 ms] of that epoch. This produced an ensemble of 501 distinct classifiers.

2.6.2.2 Testing (validation)

Before applying the trained classifiers to the Apparent motion condition, each classifier within the ensemble underwent a five-fold cross-validation (CV) procedure (Blum et al., 1999). In brief, the training samples for a specific time-point were partitioned into five equal sets. Within each of the five folds, samples were averaged across sets to improve the signal-to-noise ratio (Bae & Luck, 2018). Each of the 251 classifiers was then trained on pooled samples from four sets and tested on the fifth left-out set, to determine whether the sample should be classified as P_{+1} or P_{-1}. This training and testing procedure was iterated until each set was tested once, and for each iteration, the classification accuracy (i.e., proportion of correctly classified samples) was averaged across folds, with a chance level of 50%. The mean CV accuracy was computed across observers at each time-point (**Figure 2C**) and compared against random chance using a permutation test (see below). Above-chance accuracy at any given time-point indicates that the neural activity contains position-specific information from a single flash.

2.6.2.3 Testing (main)

The trained classifiers were then used to evaluate the ONGOING and STOP epochs. Each of the 501 classifiers was trained using the activity sample corresponding to a specific time-point t_{train} and then tested at different time-points t_{test} within that epoch. The resulting t_{test} x t_{train}

classification assignments are summarized as a TGM, shown in **Figure 2D**. Since the training and testing samples were derived from distinct conditions, cross-validation was not necessary for assessing the classifier's performance. However, it is crucial to note that there is no single true class (position) label (corresponding to P) for the ONGOING and STOP epochs. This is due to each sample encompassing lingering activity of preceding stimuli presented at P_{-1} and earlier, and, for ONGOING epochs, including activity corresponding to future presentations at P_{+1} or beyond. Therefore, the classification assignments at each point in the TGM were characterized as a *bias* towards P_{-1} (i.e., position preceding P) or P_{+1} (i.e., position following P) rather than classification "accuracy".

2.6.3 Aggregation

TGMs were aggregated separately for ONGOING and STOP configurations, as well as for Attended and Ignored streams. Subsequently, these matrices were averaged across clockwise and counter-clockwise motion directions and across all 16 stimulus positions. This resulted in four TGMs per observer, representing the factorial combination of {ONGOING, STOP} x {Attended, Ignored} conditions (**Figure 2E**). Group-level statistical testing was conducted after aggregating across observers.

2.7 Second-level analysis

2.7.1 Single flashes

For every CV accuracy, a corresponding permuted accuracy was generated. This involved training and testing each classifier on a permuted dataset where the labels associated with each position were randomly scrambled. This eliminated any systematic position information in the samples at each time-point. Consequently, the classifier's assignments were entirely random. This permutation procedure was repeated 1000 times for each observer to create a null distribution of permuted accuracy values for classification at that time-point.

To assess group-level significance, the null distributions were averaged to obtain a single null distribution with 1000 values for each time-point. To determine whether the observed CV accuracy was greater than random chance, each accuracy value was compared against the 95 percentiles of this null distribution. To correct for multiple comparisons across time-points, statistically significant accuracies are reported based on a threshold of $p = (0.05/251)$ (one-tailed), adjusted for familywise error rate (FEW) using Bonferroni correction.

2.7.2 Apparent motion

To improve comparability among observers, a "baseline" correction was applied to each TGM. Classifiers trained at time-points within [0–50 ms] (i.e., immediately after stimulus onset) were expected to perform near chance levels. Therefore, the mean value of these "baseline" classifiers on each row of the TGM was subtracted from all values of that row. Following this correction, a smoothing process was implemented using a boxcar kernel with an 8 ms width to reduce variability. The corrected and smoothed TGMs from all observers were then submitted for group-level statistical tests.

2.7.2.1 Overall effects

For each condition, we tested whether each value, indicating a classification bias towards P_{+1} or $P_{-1,}$ of the TGM significantly deviated from zero. A one-sample *t*-test was used for at each time-point to evaluate whether the bias was greater than zero (or less than zero). An image-based significance assessment, commonly used in EEG time-frequency analyses (Litvak et al., 2011) and fMRI analyses (Friston, Holmes, et al., 1994; Friston, Jezzard, et al., 1994), corrected for multiple comparisons across time-points, revealing statistically significant clusters of classification bias. This correction accounted for the non-independence of values in adjacent pixels (i.e., time-points) but treated all pixels the same in terms of their temporal relevance (see below). An FWE *t*-value threshold was estimated using the Random Gaussian Field approach (Worsley & Friston, 1995), as implemented in SPM 12 (www.fil.ion.ucl.ac.uk/spm, Welcome Department of Imaging Neuroscience, London, UK). The cluster-forming threshold was set at $p < 0.001$ (uncorrected), and the cluster-level FWE-correction was applied at $p < 0.05$ (one-tailed). Differences between conditions were assessed with the same procedure.

2.7.2.2 Temporally paired effects

A classifier trained at a given time-point in the Single flashes condition was assumed to be more sensitive to decoding the neural state at that time-point following a (predicted or actual) stimulus onset in the Apparent motion condition (see **Supplementary Methods, *S1***). This assumption guided a complementary analysis of the TGM where, at each time point *t* (row), each position was "paired" with specific classifiers. In this context, the fixed 100 ms SOA in the Apparent motion condition corresponded to specific temporal pairings: P paired with the classifier trained at time *t* (column), P_{-1} with the classifier trained at *t*-100 ms, and P_{+1} with the classifier trained at *t*+100 ms, represented as diagonal lines in **Figure 2E**.

P_{+1}-paired classifiers were expected to have a positive bias (i.e., towards P_{+1}) and P_{-1}-paired classifiers a negative bias (i.e., towards P_{-1}) at different time-points *t* (row) of the TGM. Due to the predicted time-invariance of these classifiers, the classification assignments of a particular position were averaged (i.e., by row along the diagonal line) to obtain a single bias value for that position. The diagonal-averaged bias values from each observer were used to test whether the group mean was consistent with predictions (i.e., greater than zero for P_{+1}). Differences between conditions were evaluated by a factorial repeated-measures ANOVA using the simple_mixed_anova function (Caplette, 2017) in MATLAB.

2.8 Data quality assessment

A set of exclusion criteria was implemented to ensure data quality and participant compliance with task instructions. Firstly, observers who failed to identify at least 80% of color changes in the instructed color pair (excluding false alarms related to the distractor color pair in the Apparent motion blocks) were excluded. This criterion aimed to confirm that observers attended to the task-relevant stream as instructed. Consequently, three observers who detected fewer than 70% of color changes throughout the experiment were excluded (Single flashes [M: 66.5%, SD: 17.07%], Apparent motion [M: 69%, SD: 3.9%]). Secondly, data of insufficient quality were also excluded. All classification analyses were performed on minimally pre-processed data, requiring a minimum of 128 artifact-free epochs per position. Two observers with fewer epochs (M: 102.9, SD: 19.95) were excluded. Among the remaining 28 observers, an average of ~175 Single flashes epochs (min: 150, max: 187, SD: 10.58) and ~2255 Apparent motion epochs (min: 2001, max: 2292, SD: 65.19) were included for analysis. Thirdly, data with CV accuracy exceeding -3 SD from the group mean across time were excluded to prevent any undue impact on the overall results. Three datasets were excluded based on this threshold. The final dataset comprised data from 25 observers.

3. Results

3.1 Behavioral performance

Observers performed a difficult color detection task in both Single flashes and Apparent motion conditions. The task involved reporting the target color while tracking a paired color (i.e., green while tracking white, purple while tracking black). Although task difficulty was not matched between the conditions (see **Materials and Methods**), observers detected task-relevant color changes with a high hit rate (> 90%) in both conditions.

The mean hit rate in the Single flashes condition (white/green: [M: 97.52%, SD: 2.96%]; black/purple: [M: 96.36%, SD: 4.7%]) was marginally higher than in the Apparent motion condition (white/green M: 93.56%, SD: 6.8%; black/purple M: 94.48%, SD: 5.21%). A repeated-measures ANOVA demonstrated no significant effect of color (two-way ANOVA, *motionType*color*: $F(1, 24) = 1.52$, $p = 0.2$; *motionType*: $F(1, 24) = 22.16$, $p < 0.001$; *color*: $F(1, 24) = 0.03$, $p = 0.8$).

Response times were notably faster in the Single flashes condition by ~60 ms (white/green M: 442.45 ms, SD: 52.9 ms; black/purple M: 440 ms, SD: 69.88 ms) than in the Apparent motion condition (white/green M: 507 ms, SD: 73.54 ms; black/purple M: 504.28 ms, SD: 65.11 ms), also without a detectable influence of color (two-way ANOVA, *motionType*color*: $F(1, 24) = 0.0$, $p = 0.9$; *motionType*: $F(1, 24) = 86.43$, $p < 0.001$; *color*: $F(1, 24) = 0.192$, $p = 0.6$). Overall, these results indicate that observers were able to perform the same task comparably in both conditions.

3.2 Decoding position

Following Blom et al. (2020), we assumed that a stimulus at a particular position should evoke a similar position-specific neural response in both the Single flashes and Apparent motion conditions. This allowed us to train classifiers on the Single flashes condition to identify corresponding patterns in the Apparent motion condition. Therefore, we first wanted to verify the validity of this key assumption.

3.2.1 Qualitative differences in spatially and temporally specific responses

If a stimulus evokes a position-specific neural response, then stimuli presented at different positions should be distinguishable (i.e., classifiable) based on their evoked activity. Since our classification analyses rely on differences in 2-ms snapshots of evoked potentials across the scalp, we examined whether these small-scale activity patterns could differentiate positions within the Single flashes condition. A typical signature of such activity is a lateralized neural response to peripherally presented stimuli (**Figure 3**). Consistent with this signature, stimuli presented to the left (or right) hemifield evoked a qualitatively higher mean positivity at 100 ms over the posterior channels contralateral to the stimulus than the corresponding ipsilateral channels. However, differences between neighboring positions in the same hemifield were subtler, with a crucial role for time. For instance, the pattern at one time-point (e.g., 100 ms) was not a simple variant of the patterns at other time points (e.g., 150 or 250 ms). Therefore, we anticipated that inter-stimulus activity differences would dynamically evolve over time.

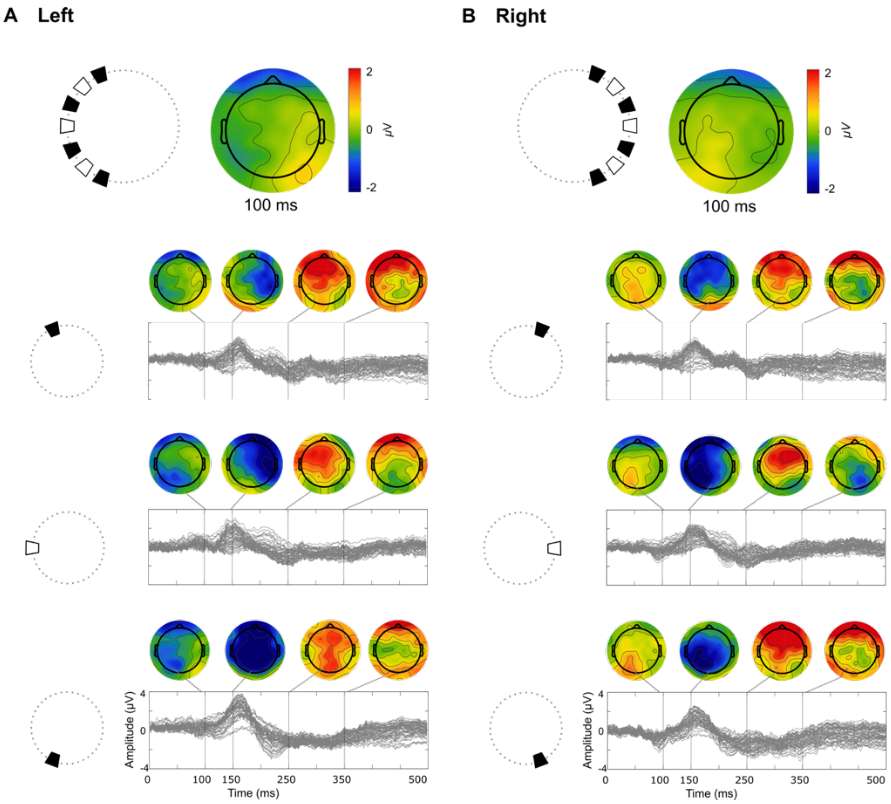

Figure 3. Position-specific neural responses to single flashes. Topography and event-related potentials (ERPs) were computed relative to the pre-stimulus baseline period over [-200–0 ms] for stimuli displayed to the (**A**) left and (**B**) right hemifield, averaged across observers. For positions on the most left (or most right), scalp maps show activity patterns at 100, 150, 250, and 300 ms following stimulus onset.

3.2.2 Position-classification accuracy varies with distance and time

We tested *whether* and *when* stimulus positions were classifiable within the Single flashes condition. A classifier was trained to discriminate activity patterns between two color-matched positions at each time-point over [0–500 ms] following stimulus onset and tested on activity patterns from the same time-points not used for training.

We first assessed whether two positions could be classified *at all* at any time-point. The peak classification accuracy across the entire time range (computed per individual observer and subsequently averaged across observers) for the position pairs is plotted in **Figure 4A** based on their spatial distance. The peak accuracy was well above the 50% chance level for all position pairs (**Supplementary Materials, *Table S2***). As expected, peak accuracy increased as positions with spatial separation, ranging from 59.83% (SD: 1.52%) for positions with an angular distance of 45° to 64.29% (SD: 3.71%) with an angular distance of 180° (one-way ANOVA, $F(3, 72) = 70.39$, $p < 0.001$). We also observed a linear relationship between accuracy and the linear distance between stimulus positions (i.e., chord length), confirming that *spatial position* drove the classification over and above other co-varying features, such as the orientation of the wedge.

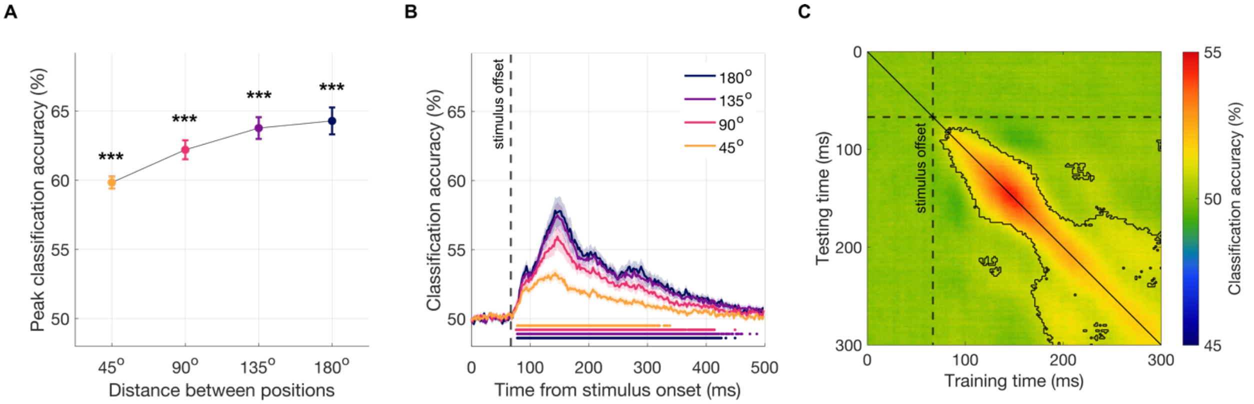

Figure 4. Position classification results of single flashes. (**A**) An ensemble of classifiers was trained and tested on independent, non-trained samples between [0–500 ms]. The peak classification accuracy across the entire time range (computed per individual, then averaged across observers) is plotted as a function of angular distance between position pairs. All values are reported as means ± standard error of the mean (S.E.M., error bars). *** denotes $p < 0.001$. (**B**) The cross-validated (CV) accuracy is plotted for all position pairs, averaged across positions. All values are reported as means ± S.E.M. (shading). The dashed vertical line at 67 ms corresponds to the offset time of the stimulus. Colored dots below the accuracy mark the time-points at which the positions could be classified above 50% chance level, compared to a permuted null distribution ($p < 0.002$, FWE-corrected). (**C**) An ensemble of trained classifiers (x-axis) was tested at different time-points (y-axis) between [0–300 ms]. The classification assignments are summarized as a TGM, averaged across positions and observers. The dashed lines at 67 ms correspond to the offset time of the stimulus. The diagonal line identifies where training and testing time-points are equal. Solid black contours indicate accuracy greater than 50% chance (one-sample t-test, $p < 0.05$, FWE-corrected).

However, the peak accuracy described above was agnostic about *when* the maximum accuracy occurred, which could differ across observers. Therefore, we next evaluated *when* in time classification was possible. The time-resolved classification (cross-validated; CV) accuracy is plotted for each position pair in **Figure 4B**. For all pairs, classification was significantly above 50% chance level at multiple time-points concentrated in [100–300 ms] (permutation test, *ps* < 0.002, FWE-corrected). CV accuracy increased significantly above chance from 78 ms, reaching a peak around 146 ms, before gradually reducing back to chance levels at ~450 ms. Peak accuracy varied from 53.14% (SD: 1.86%) for positions with an angular distance of 45° to 57.71% (SD: 5.33%) for positions with an angular distance of 180° (one-way ANOVA, $F(3, 72) = 35.11$, $p < 0.001$). These time-resolved peak accuracies were notably lower by ~7% than the above time-agnostic peak accuracies (**Figure 4A**), which indicates a large variability in the timing of maximum accuracy across individuals.

3.2.3 Low temporal generalization of high accuracy time-resolved classifiers

The [80–200 ms] interval of significant classification corresponded qualitatively to the period during which the ERPs showed rapid changes in magnitude and polarity (**Figure 3**). This similarity raised the possibility that classification over this interval might be driven by a common multivariate pattern that varied in strength over time (i.e., first increasing and then decreasing). If this were the case, then the position-specific information underlying classification at one time-point could enable classification at other time-points within [0–300 ms]. To test this possibility, we evaluated the classification time course. The TGM in **Figure 4C** summarizes how well an ensemble of classifiers trained at time-points t_{train} (x-axis) could classify the same positions at different time-points t_{test} (y-axis), averaged across all position pairs. Mean accuracy along the diagonal was significantly greater than 50% chance level between [~100–300 ms] (one-sample *t*-test, $p < 0.05$, FWE-corrected, cluster threshold = 10 pixels), as expected from **Figure 4B**. Interestingly, above-chance accuracy at the off-diagonal pixels between [100–200 ms] was also evident in a narrow band around the diagonal. These results of high accuracy indicate that classifiers were temporally specific within this narrow band and could not share position information with neighboring classifiers, including classifiers at future time-points. Cross-time classification was broadly possible around [200–300 ms] (***lower right corner***), suggesting a classifier trained at 200 ms (x-axis) could generalize to time-points over the entire range [~180–300 ms] (y-axis). Therefore, the position information in the evoked response was truly changing rapidly over time, rather than reflecting the changing strength of a common activity pattern.

3.3 Generalizing from Single flashes to Apparent motion

3.3.1 Cross-condition position-classification accuracy is modulated by attention

Following the confirmation in the Single Flashes condition, we tested whether classifiers could generalize to the Apparent motion condition given the superposed activity evoked by multiple stimuli presented at different positions and times (see **Supplementary Materials, *S1***). Thus, a trained classifier was tested on positions from any order (1st – 12th stimulus) within a given motion sequence. This evaluation was conducted at each time-point over [0–300 ms] following stimulus onset. The classification accuracy for the Attended and Ignored conditions is plotted for each position pair in **Figure 5A**.

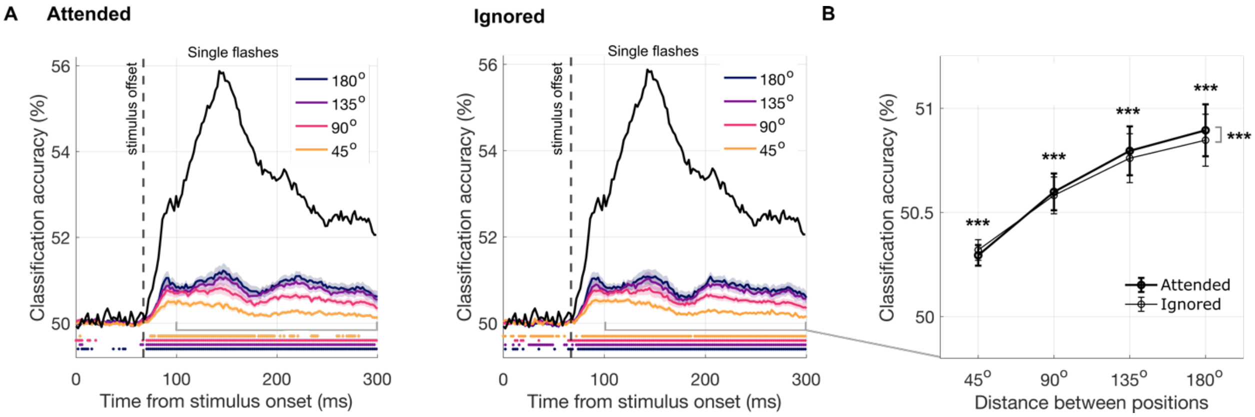

Figure 5. Cross-condition classification results of apparent motion. (**A**) An ensemble of classifiers was trained in the Single flashes condition and then applied to the Apparent motion condition between [0–300 ms]. The time-resolved classification accuracy is plotted for all position pairs, averaged across positions, in the Attended (*left*) and Ignored (*right*) conditions. All values are reported as means ± S.E.M. (shading). The dashed vertical line at 67 ms corresponds to the offset time of the stimulus. Colored dots below the accuracy mark the time-points at which the positions could be classified above 50% chance level from the superposed activity evoked by multiple stimuli presented at different positions and times, compared to a permuted null distribution (p < 0.002, FEW-corrected). The black line labeled Single flashes corresponds to the CV accuracy in the Single flashes condition for comparison. (**B**) Above-chance accuracy values were averaged over [100–300 ms] (indicated in **A**) and plotted as a function of angular distance between position pairs. All values are reported as means ± S.E.M. (error bars). *** denotes p < 0.001.

In both conditions, accuracy values ranged between 50% and 51%. These accuracies were ~5% lower than the CV accuracy associated with the Single flashes. Despite the low values, classification was possible at multiple time-points concentrated in [100–300 ms] (p < 0.002, FEW-corrected) in the Attended condition. Mean accuracy rose significantly above chance from 74 ms, first peaking at around 146 ms and second between [205–220 ms]. A similar pattern was observed in the Ignored condition. To examine whether classification strength differed between the two conditions, we compared the mean accuracy averaged over [100–300 ms] (**Figure 5B**). In the Attended condition, mean accuracy increased with the distance between positions, ranging from 50.29% (SD: 0.12%) for positions with an angular distance of 45° to 50.89% (SD: 0.31%) for positions with an angular distance of 180° (**Supplementary Materials, *Table S4***). This was slightly lower in the Ignored condition, measuring 50.32% (SD: 0.11%) for positions with an angular distance of 45° to 50.85% (SD: 0.28%) for positions with an angular distance of 180° (two-way ANOVA, *attention*distance:* $F(3, 72) = 3.81$, $p = 0.013$; *distance:* $F(3, 72) = 143.18$, $p < 0.001$; *attention:* $F(1, 24) = 0.41$, $p = 0.528$).

These results indicated that attention modulated position classification within motion sequences. However, this analysis did not explicitly address whether the neural activity following a stimulus at a position (P) conveyed information about its subsequent position (P_{+1}). We therefore examined the ONGOING sequence configuration.

3.3.2 Stimulus-consistent patterns represent both task-relevant and distractor positions

Figure 6A presents TGMs of the ONGOING conditions over [0–300 ms] following the stimulus onset of P. Each time-point was evaluated by an ensemble of classifiers (x-axis, in 1 ms steps) trained to distinguish the activity evoked by P_{+1} and P_{-1} in the Single flashes condition. TGM values indicate classification bias for P_{+1} or P_{-1}. Despite the low values ranging from -0.01% to +0.01%, these biases showed statistically significant deviations from zero across a large number of classification assignments, even at stricter thresholds used to correct for multiple comparisons.

The Attended and Ignored TGMs were qualitatively similar. With P_{-1} (and P_{-2}) presented before P, negative bias was concentrated from [0–150 ms] (y-axis) and absent after ~200 ms. This was particularly prominent for classifiers between [~100–300 ms] (x-axis), with a concentration along the upper diagonal. Bias for P_{-1} was also evident along the uppermost diagonal, as expected from the high spatial generalization of classifiers (see **Supplementary Materials, *Figure S3***).

Importantly, the horizontal line at 100 ms (y-axis) indicates the onset time of the stimulus at P_{+1}. Positive bias commenced at ~75 ms following this onset marker (~175 ms after P). Unlike the bias for P_{-1}, the bias for P_{+1} was not as widespread and was concentrated along the lower diagonal over [~80–200 ms] (x-axis).

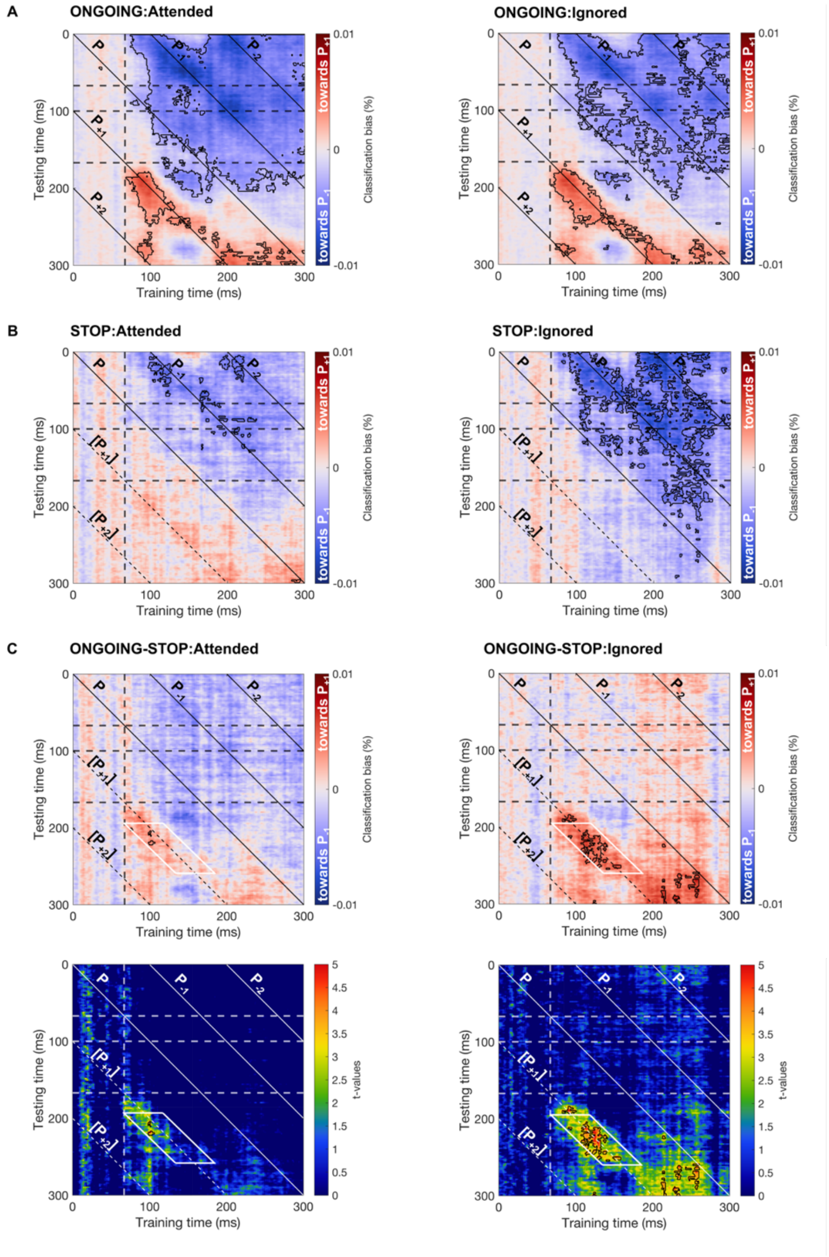

Figure 6. Main results. (**A**) An ensemble of classifiers was trained to distinguish P_{+1} and P_{-1} (90° apart) over [0–300 ms] in the Single flashes condition were tested on P in the ONGOING:Attended (*left*) and ONGOING:Ignored (*right*) conditions. The classification assignments are summarized as a TGM, relative to the onset time of the stimulus at P (y-axis). The dashed lines at 67 ms correspond to the offset time of the stimulus at P, and the dashed horizontal lines at 100 ms and 167 ms (y-axis) indicate the onset and offset times at P_{+1}, respectively. The diagonal lines represent the position-paired classifiers for P_{-2} and P_{-1} (upper diagonal lines), P (diagonal line) and P_{+1} and P_{+2} (lower diagonal line). Solid black contours indicate clusters of classification bias that differ significantly from zero (one-sample *t*-test, $p < 0.05$, FEW-corrected). Negative bias is illustrated in dark blue and positive bias in dark red. (**B**) Same described in **A** but for the STOP

conditions. The dashed horizontal lines at 100 ms and 167 ms (y-axis) indicate the predicted (but not actual) onset and offset time at P_{+1}, respectively. Similarly, the dotted diagonal lines represent the classification decisions by the position-paired classifiers for the predicted P_{+1} and P_{+2} (lower diagonal lines). (**C**) The subtraction of Attended (*left*) and Ignored (*right*) conditions and corresponding t-values for the contrast compared to zero (one-sample *t*-test, $p < 0.05$, FEW-corrected). The white box signify a narrow band of P_{+1}-paired classifiers ± 30 ms around the diagonal.

In sum, the TGMs revealed unambiguously strong on-off bias patterns consistent with a stimulus appearing at every 100 ms within ongoing motion sequences of each stream. However, it was ambiguous whether the information about P_{+1} was veridically evoked by the stimulus at P_{+1} or if it also included a motion-based prediction of P_{+1} when viewing P. To disentangle these alternatives, we next evaluated the STOP sequence configuration.

Figure 6B presents TGMs of the STOP conditions. As in the ONGOING conditions, negative bias was also concentrated from [0–150 ms] (y-axis). However, there was a striking absence of statistically significant positive bias, consistent with the absence of a stimulus presented at P_{+1}. To confirm whether this was indeed the case, we tested whether the bias related to P_{+1} was comparable in the ONGOING and STOP conditions (**Figure 6C**).

The ONGOING-STOP contrast for the Attended condition was statistically significant at small clusters around ~200 ms following P (or ~100 ms after P_{+1}) for P_{+1}-paired classifiers, confirmed by the distribution of positive t-values along the lower diagonal. This indicated that the bias for P_{+1} was comparable within the Attended stream. However, the ONGOING-STOP contrast of the Ignored condition was significant at more prominent clusters for the same classifiers over the range [200–250 ms] (y-axis), indicating that the bias for P_{+1} was considerably different for the Ignored stream.

3.3.3 Evidence of motion extrapolation shows selectivity for task-relevant compared to distractor stream

The results of the ONGOING-STOP:Attended and ONGOING-STOP:Ignored contrasts were in the desired direction, suggesting that the observed differences might be attributed to prediction. However, it could be argued that these differences were marginal and potentially influenced by attention. To address this concern, we conducted a more detailed examination of how changes in attentional state could impact the relationship between the ONGOING and STOP conditions.

We computed an average bias value for P_{+1} for a narrow band of P_{+1}-paired classifiers ± 30 ms around the lower diagonal depicted in **Figure 6C** (white box). The diagonal-averaged bias values are plotted in **Figure 7** to match the TGMs above. The bias values were qualitatively similar for the Attended conditions (**Figure 7A**), but a significant divergence was observed for the Ignored conditions (**Figure 7B**). Specifically, the diagonal-averaged bias values for STOP:Ignored condition were close to zero (i.e., chance level) without a notable peak across time, whereas the bias for the STOP:Attended condition was relatively higher, with a subtle peak, albeit lower than either ONGOING conditions.

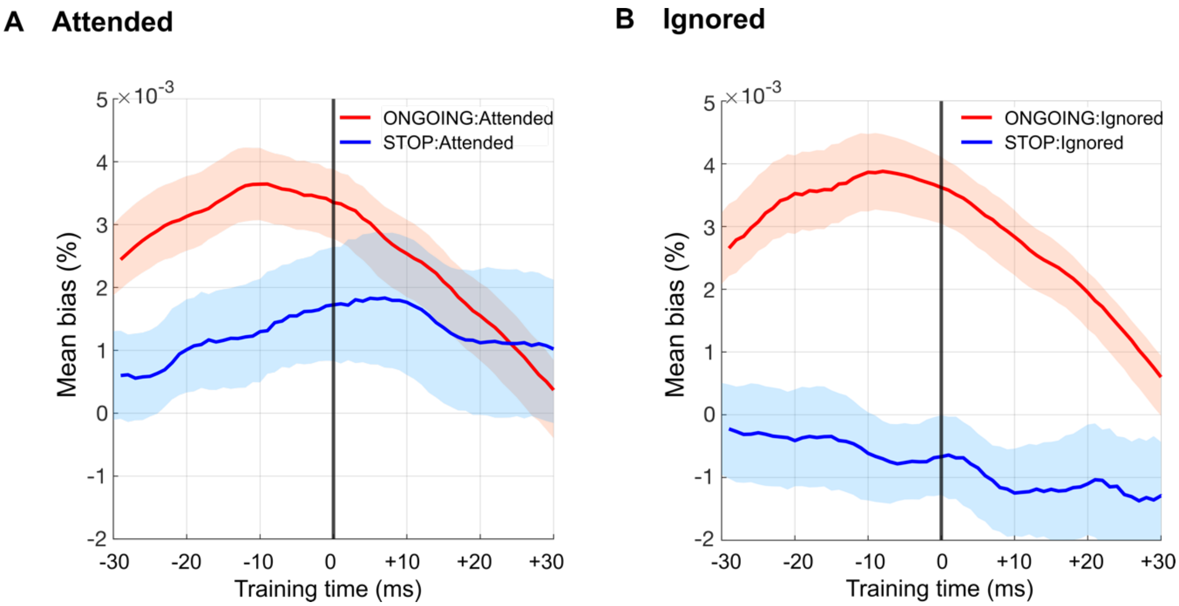

Figure 7. Temporal pairing results. (**A**) Diagonal-averaged bias values for the Attended condition. All values corresponding to the ONGOING:Attended (orange line) and STOP:Attended (blue line) conditions are reported as means ± S.E.M. (shading). These colors are not associated with the colors from the TGMs. The vertical line at 0 ms represents the lower diagonal line in the TGM shown in **Figure 6**. (**B**) Same described in **A** but for the Ignored condition.

To formally assess these differences, we conducted a 3-way mixed ANOVA with the factors of *time* ([-30 ms to +30 ms]), *attention* (Attended, Ignored), and *sequenceType* (ONGOING, STOP). **Table 1** (***column 90°***) indicates a significant main effect for sequence type and a marginal two-way interaction with attention. The three-way interaction did not reach the threshold for statistical significance, and this effect was also not detectable by direct comparisons of the TGMs (see **Supplementary Materials, *Figure S5***). The direction of the two-way interaction was further evaluated in a post-hoc, two-way ANOVA involving the factors of *time* and *sequenceType*, revealing a notable effect for sequence type in the Ignored condition (two-way ANOVA, *time*sequenceType*: $F(60, 1440) = 3.22$, $p < 0.001$; *sequenceType*: $F(1, 24) = 16.38$, $p < 0.001$; *time*: $F(60, 1440) = 6.08$, $p < 0.001$), which was not detectable in the Attended condition (two-way

ANOVA, *time*sequenceType*: *F*(60, 1440) = 4.98, $p < 0.001$; *sequenceType*: F(1, 24) = 2.26, $p = 0.146$; *time*: *F*(60, 1440) = 4.45, $p < 0.001$).

Table 1.
Results of a three-way mixed ANOVA for diagonal-averaged bias values.

Effects	90°				135°			
	df1	*df2*	*F*	*p*	*df1*	*df2*	*F*	*p*
time*attention*sequenceType	60	1440	1.13	0.238	60	1440	1.06	0.364
attention*sequenceType	1	24	3.23	0.085	1	24	2	0.170
time* sequenceType	60	1440	7.23	< 0.001	60	1440	3.72	< 0.001
time*attention	60	1440	1.01	0.345	60	1440	0.24	1
sequenceType	1	24	17.76	< 0.001	1	24	15.62	< 0.001
attention	1	24	2.18	0.153	1	24	5.02	0.035
time	60	1440	7.15	< 0.001	60	1440	1.66	0.210

Taken together, the dissimilarity observed in the Ignored conditions demonstrated that the information about P_{+1} had a solely stimulus-evoked origin – present in ONGOING but absent in STOP. In contrast, the Attended conditions showed a different relationship. Despite the absence of P_{+1} in STOP, there was a weak bias for P_{+1} that was statistically indistinguishable from the corresponding bias in ONGOING, where P_{+1} was present. Therefore, when a stimulus was attended, the information about P_{+1} seemed to have additional, predictive, origins.

3.3. Confirmation with 135°

To eliminate the possibility that our findings were influenced by the specifics of the classifiers, we replicated our main analyses employing a different ensemble of classifiers to train on P_{+1} and P_{-2}, which were 135° apart and shared similar characteristics to the 90° classifiers (see **Supplementary Materials, *Figure S3***). Although the 135° classifiers were trained on different positions from the Single flashes conditions compared to the classifiers above, the epochs from the Apparent motion condition were unchanged.

All corresponding TGMs are provided in **Supplementary Materials, *Figure S6.1***. As expected, comparable bias patterns were recovered from the activity related to the motion sequences. There was a notable difference in the bias for P_{+1} was present in the ONGOING:Ignored condition, which was absent in the STOP:Ignored condition. Additionally, there were no marked differences in the bias for P_{+1} between the ONGOING:Attended and STOP:Attended conditions. These results were reflected in the diagonal-averaged bias values (**Figure S6.2**).

Moreover, the three-way ANOVA revealed a significant effect of attention (**Table 1*, column 135°***). This was accompanied by an effect of sequence type for the Ignored condition (two-way ANOVA, *time*sequenceType*: $F(60, 1440) = 2.63$, $p < 0.001$; *sequenceType*: $F(1, 24) = 21.13$, $p < 0.001$; *time*: $F(60, 1440) = 1.87$, $p < 0.001$) that was not detectable for the Attended condition (two-way ANOVA, *time*sequenceType*: $F(60, 1440) = 2.76$, $p < 0.001$; *sequenceType*: $F(1, 24) = 1.59$, $p = 0.22$; *time*: $F(60, 1440) = 1.01$, $p = 0.471$).

All key findings were replicated using the 135° classifiers. This provided a huge confirmation that the findings cannot be solely due to the classifier choice.

4. Discussion

In this study, we investigated whether attention was necessary for motion extrapolation. Observers viewed two streams of apparent motion stimuli, with one stream relevant to the task. Using MVPA, we aimed to track in fine temporal resolution the neural representation of stimulus position throughout the motion sequences in each stream. Specifically, we probed whether the neural representation of a stimulus contained sensory-like representations driven by motion-based prediction (Blom et al., 2020). Our findings revealed evidence of motion extrapolation in the task-relevant stream and a concluding absence of this effect in the distractor stream.

At any given moment, the concurrently presented stimulus streams vie for processing in the visual processing hierarchy. Despite the significant role of top-down selective attention in early selection, our classification results uncovered significant responses evoked by all detectable positions of both attended and ignored stimuli during motion sequences. We observed some modulations of attention in enhanced representation strength of the attended stimulus throughout the response. However, there was a shift in selectivity as the motion sequence unfolded. A dissociable representation of the stimulus' future position emerged when the motion sequence stopped (unknown to the observer), coinciding with their respective attentional states. Notably, we discerned no predictive information about the future position of the ignored stimulus. These findings echo the conclusions drawn by Shioiri et al. (2002), suggesting that attention not only extends to newly occupied positions but also anticipates future positions to facilitate the tracking of a moving object. Our results imply that while the brain possesses the capacity to represent the location of all presented stimuli that overlapped in time (Mangun & Hillyard, 1990; Sahan et al., 2019; Volpe et al., 1979), it seems to absorb little information from motion that is ignored.

One possible mechanism is an implicit selection mechanism guided by behavior relevance, as observers needed to focus on specific stimuli to perform the color detection task successfully. This involved detecting color changes within the task-relevant stream while ignoring any color changes occurring within the distractor stream. Despite having independent moving trajectories, the two streams were presented within the same visual field, enabling observers to perceive both stimuli in different positions simultaneously (or with a slight jitter). This setup created a competitive interaction between the two streams. A bias in this competition could emerge as attentional resources were predominantly allocated to the positions of the task-relevant stimulus, anticipating the possibility of the stimulus appearing in the target color. This selective allocation could impose a suppressive effect on the distractor stream (Desimone & Duncan, 1995).

This mechanism bears resemblance to the well-studied “Cocktail Party” effect observed in auditory perception (Cherry, 1953), wherein attentional resources selectively anticipate events within the task-relevant speech stream (Zion Golumbic et al., 2013). By virtue of this suppression, events presented within the task-relevant stream can be effectively perceived and predicted.

Observers might have been unable to extrapolate the motion presented within the distractor stream because, although attention may not be necessary for detecting features such as position or color, it is required for relating these features to the same object (i.e., wedges). This binding process would be essential for addressing the superposition problem, which demonstrates that overlapping and conflicting signals can create ambiguous representations. To accurately perceive the two separate streams, the brain must resolve these ambiguities in neural coding by allocating resources to the task-relevant stream. This would enhance its neural representation and facilitate the perception of the attended stimulus (Luck & Ford, 1998). Even though motion was not explicitly instructed, this implies that attention is involved in binding the successive positions and colors within the task-relevant stream, enabling extrapolation (Albright & Stoner, 1995; cf. feature integration theory; O’Craven et al., 1997; Treisman & Gelade, 1980). This mechanism provides ‘what’ in ‘where’ information early on, helping to navigate dynamic environments (Holcombe, 2023; Reynolds & Desimone, 1999). Conversely, the distractor stream may fail to provide a meaningful, coherent representation of motion, leading to the perception of a collection of colored flashes without motion extrapolation – similar to how listeners might struggle to understand snippets of speech without paying attention to the conversation.

Alternatively, motion might be perceived for both attended and ignored stimuli, but predictive mechanisms may be reserved for attended stimuli. Such selectivity could stem from the behavior irrelevance of extrapolation for the ignored stimulus, contributing to a more efficient allocation of cognitive resources. While the current study cannot definitively conclude whether observers did perceive apparent motion in both streams or if there was a difference in the perception of apparent motion, it is nonetheless suggestive. Since this was not explicitly measured, it remains open to interpretation, and additional research may be needed to disentangle these possibilities.

Our results illuminate the predictive encoding of the position of a moving object, particularly when attention is directed towards it. However, our results diverge from prior studies such as Blom et al. (2020), which associated pre-activation with the compensatory role of predictive motion extrapolation. While Blom et al. observed pre-activation as early activation of representation of the stimulus’ next position [70–90 ms] before its expected onset, even before information about its current position arrived, we did not observe such latency advantages associated with prediction-driven representations. We observed pre-activation that emerged from 30 ms before and peaking precisely at the expected arrival of information. This discrepancy could be attributed to factors specific to our paradigm – since the classification results of the static stimuli (Single flashes) replicated Blom et al. (2020), differences in the apparent motion conditions likely contribute to a weaker signal-to-noise ratio, potentially impacting the temporal dynamics of the classifications observed here. One key difference lies in our study involving two stimulus streams. The consequence is a weaker signal in the neural response to any one stimulus, or a superposed mixture of activity from the current, prior, and future positions from both streams. As seen in the low classification values in **Figure 6**, the vertical streaks of positive or negative biases (x-axis), denoting noise, are as prominent as the actual position-specific information decoded. In principle, this weaker signal could contribute to the delayed onset of significant classification, as noted in Grootswagers et al. (2017). Furthermore, our sequences were notably shorter, ranging from 5-12 stimulus presentations, compared to their varying lengths up to 44 stimulus presentations. As latency tends to decrease with increasing trajectory lengths (Benvenuti et al., 2020), the discrepancy in sequence lengths could also influence the amount or quality of motion signals available for extrapolation in both studies.

While our present results cannot definitively establish pre-activation ahead of neural delays as observed in Blom et al. (2020), it is clear that: (1) motion extrapolation aligns the future position with the expected moment of the attended stimulus despite neural delays, and (2) extrapolation did not occur for the ignored stimulus. A more comprehensive investigation of the predictive mechanisms and attentional resources in a dynamic environment could involve merging the two experiments, in addition to a one-Apparent motion condition with a single task-relevant stream, as used in Blom et al. (2020) and a mixed condition with one stream of Single flashes in random positions and another Apparent motion stream in successive positions. This investigation would provide further insights into the impact of the strategic allocation of attention in different contexts of stimulus complexity and task demands.

One critique of the present study is the absence of false alarms in the task. While the hit rate indicated that observers responded appropriately to the task-relevant stream, the inclusion of false alarms would have provided insights into whether the observers also responded to the distractor stream. Despite demonstrating dissociable evidence of motion extrapolation for attended and ignored stimuli, incorporating a behavioral measure that aligns with the neural evidence could offer a more complete understanding of how attention shapes the coding of task-irrelevant, ignored stimuli. This presents a potential avenue for future research.

Another limitation worth noting is the discrepancy in the number of epochs satisfying the STOP configuration compared to the ONGOING configuration. This discrepancy could account for the numerical difference in classification strength between the ONGOING and STOP conditions. However, it is important to highlight that both STOP:Attended and STOP:Ignored conditions had a comparable number of epochs. The setup for the Attended and Ignored conditions was identical, except for the instructions for the instructed attended colors. While we note the discrepancy in testing data, the training data and STOP conditions were matched, so it does not pose a major concern to the overall validity of the results.

When perceiving a moving object, motion extrapolation relies on its past trajectory to generate predictions about its future positions along the expected path. This study employed pattern classification analyses of EEG data to investigate the position representations of attended and ignored apparent motion stimuli. Our findings demonstrate that top-down attention plays a crucial role in guiding motion extrapolation to be selective, and the effect of this selection is that a meaningful representation of motion may not be constructed for the ignored stimulus. While subject to interpretation, our data strongly support that attention may be essential to piece together a tacit representation of motion.

Supplementary Materials for Meaningful objects: Selective visual attention guides predictive motion extrapolation

Jane Yook *et al.*

Supplementary Methods

S1. Spatiotemporal decoding approach and the superposition problem

The spatiotemporal decoding approach (Blom et al., 2020) was used to address a key difficulty posed by the motion sequence. In each sequence, successive stimuli were presented with a short constant SOA of 100ms to induce the percept of apparent motion. With this short SOA, the activity at a time-point in the sequence was effectively an overlapping (superposed) mixture of activity evoked by multiple stimuli presented at different positions and times. Furthermore, the fixed (non-jittered) SOA limited the use of deconvolution methods (e.g., Temporal Response Function (Crosse et al., 2016) to isolate the neural responses to individual stimuli. This presents a difficulty in making inferences about a specific predicted position from superposed neural activity.

The solution afforded by the spatiotemporal decoding approach was based on the following rationale. A stimulus seen at a particular position Px is assumed to evoke a characteristic position-specific activity pattern across the scalp that evolves in a distinctive manner over time. This neurodynamic activity pattern is assumed to be similar both for stimuli and predicted stimuli (Blom et al., 2020). However, as described above, this position-specific pattern was not directly measurable in the Apparent motion condition due to the superposed activity from stimuli presented at other positions. Nevertheless, the characteristic pattern for position Px is assumed to persist in the superposed activity even if in a degraded form.

When considering the transition from a position P to a predicted position P_{+1}, the strategy is to decode P_{+1} from superposed activity by using a pre-constructed template that can sensitively discriminate the pattern for P_{+1} (even if degraded) from “competing” superposed patterns from P and other preceding positions of the sequence. In the current study, this discriminative template for P_{+1} was constructed (i.e., trained) using machine learning from pattern examples obtained in the Single flashes condition where superposition was minimal (SOA = 500 ms). Blom et al. used a different set of examples for training. Importantly, this discriminative template was temporal as it consisted of an ensemble of multiple time-specific binary classifiers.

Supplementary Results

S2. Above-chance time-agnostic peak cross-validated (CV) accuracy

Table S2.

Results of one-sample *t*-tests (one-tailed) for the time-agnostic peak CV accuracy values.

Distance	*M*	*SD*	t_{27}	*p*
45°	59.83%	1.52%	32.3	< 0.001
90°	62.2%	2.68%	22.76	< 0.001
135°	63.77%	3.33%	20.71	< 0.001
180°	64.29%	3.71%	19.24	< 0.001

S3. High spatial generalization of position classifiers

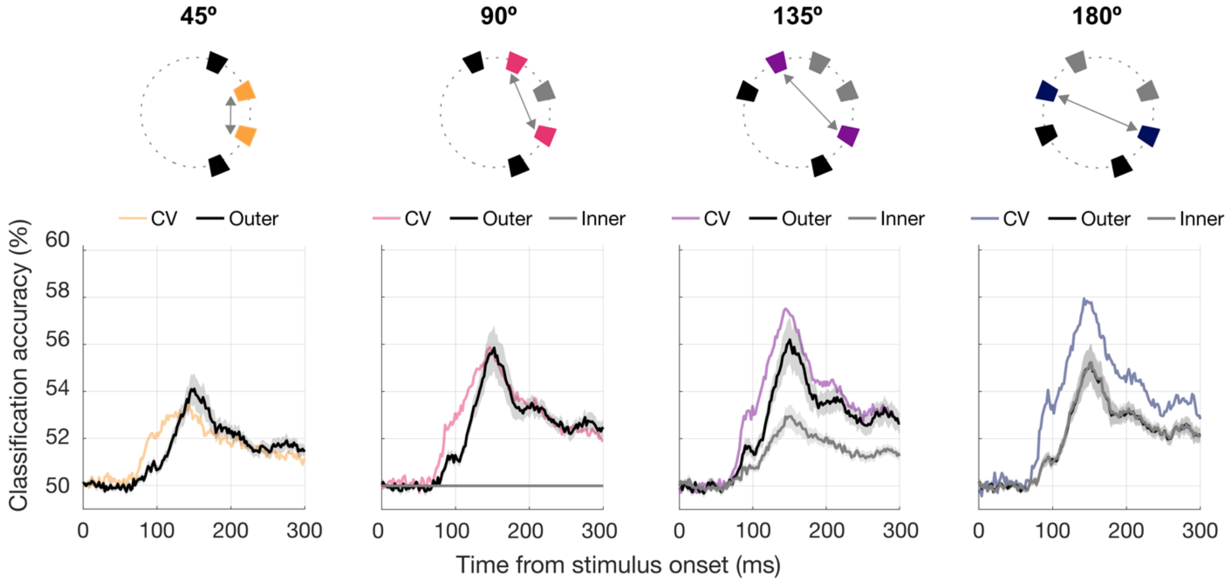

Figure S3. Spatial generalization results. We evaluated whether classifiers could generalize to nearby spatial positions in the Single flashes condition. An ensemble of classifiers was trained to distinguish two positions between [0–300 ms] and tested on positions that not used for training, namely, a position in between (inner) or outside (outer) each position pair. These positions are graphically depicted in the corresponding Inset (position pairs shown clockwise for simplicity). Note that there is no inner position (gray) for the position pair with the closest angular distance of 45°. For the position pair with the furthest angular distance of 180°, the inner and outer positions are on opposite sides of the circle and are functionally identical. The classification accuracy is plotted for the inner (gray) and outer positions (black). All values are reported as means ± S.E.M. (shading). The colored line corresponds to the CV accuracy for comparison. In line with the CV results, mean accuracy for the inner positions is qualitatively distinct from that for the outer positions. Notably, mean accuracy for the inner position of 90° is at 50% chance level across the entire period following the onset time of the stimulus. Moreover, the classification accuracy improves with increasing distance between positions, such that the accuracy values for the inner and outer positions of 180° are equal.

S4. Above-chance cross-condition classification accuracy

Table S4.

Results of one-sample *t*-tests (one-tailed) for the cross-condition classification accuracy values.

Distance	Attended				Ignored			
	M	*SD*	t_{27}	*p*	*M*	*SD*	t_{27}	*p*
45°	50.29%	0.12%	12.11	< 0.001	50.32%	0.11%	14.69	< 0.001
90°	50.6%	0.22%	13.51	< 0.001	50.58%	0.19%	15.6	< 0.001
135°	50.8%	0.29%	13.53	< 0.001	50.76%	0.25%	15.42	< 0.001
180°	50.89%	0.32%	14.34	< 0.001	50.85%	0.28%	15.22	< 0.001

S5. Interaction effects for 90°

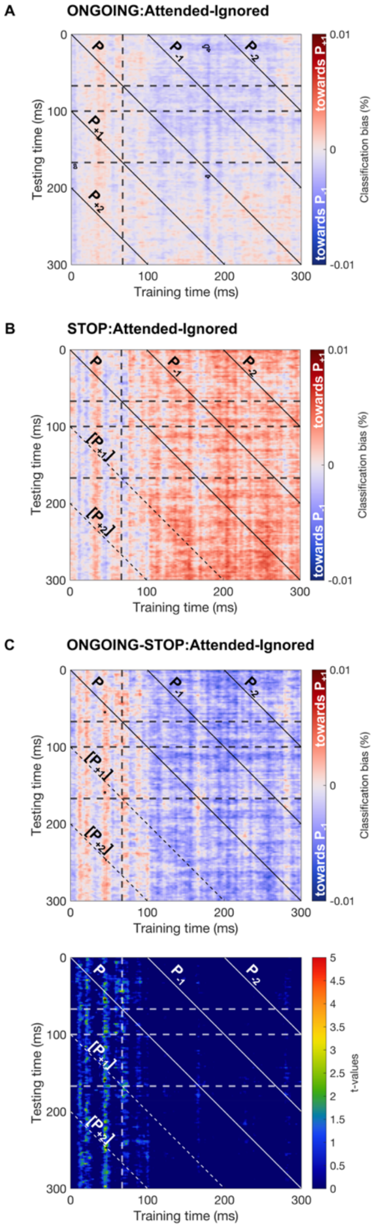

Figure S5. Interaction effects. (**A**) An ensemble of classifiers was trained to distinguish P_{+1} and P_{-1} (90° apart) over [0–300 ms] in the Single flashes condition were tested on P in the ONGOING:Attended and ONGOING:Ignored conditions. The subtraction of the two ONGOING conditions are plotted in ONGOING:Attended-Ignored**.** The classification assignments are summarized as a TGM, relative to the onset time of the stimulus at P (y-axis). The dashed lines at 67 ms correspond to the offset time of the stimulus at P, and the dashed horizontal lines at 100 ms and 167 ms (y-axis) indicate the onset and offset times at P_{+1}, respectively. The diagonal lines represent the position-paired classifiers for P_{-2} and P_{-1} (upper diagonal lines), P (diagonal line) and

P_{+1} and P_{+2} (lower diagonal line). Negative bias is illustrated in dark blue and positive bias in dark red. (**B**) Same described in **A** but for the STOP conditions.

The dashed horizontal lines at 100 ms and 167 ms (y-axis) indicate the predicted (but not actual) onset and offset time at P_{+1}, respectively. Similarly, the dotted diagonal lines represent the classification decisions by the position-paired classifiers for the predicted P_{+1} and P_{+2} (lower diagonal lines). (**C**) The subtraction of ONGOING:Attended-Ignored and STOP:Attended-Ignored conditions and corresponding t-values for the contrast compared to zero (one-sample *t*-test, $p < 0.05$, FWE-corrected).

S6. Key findings replicated using 135° classifiers

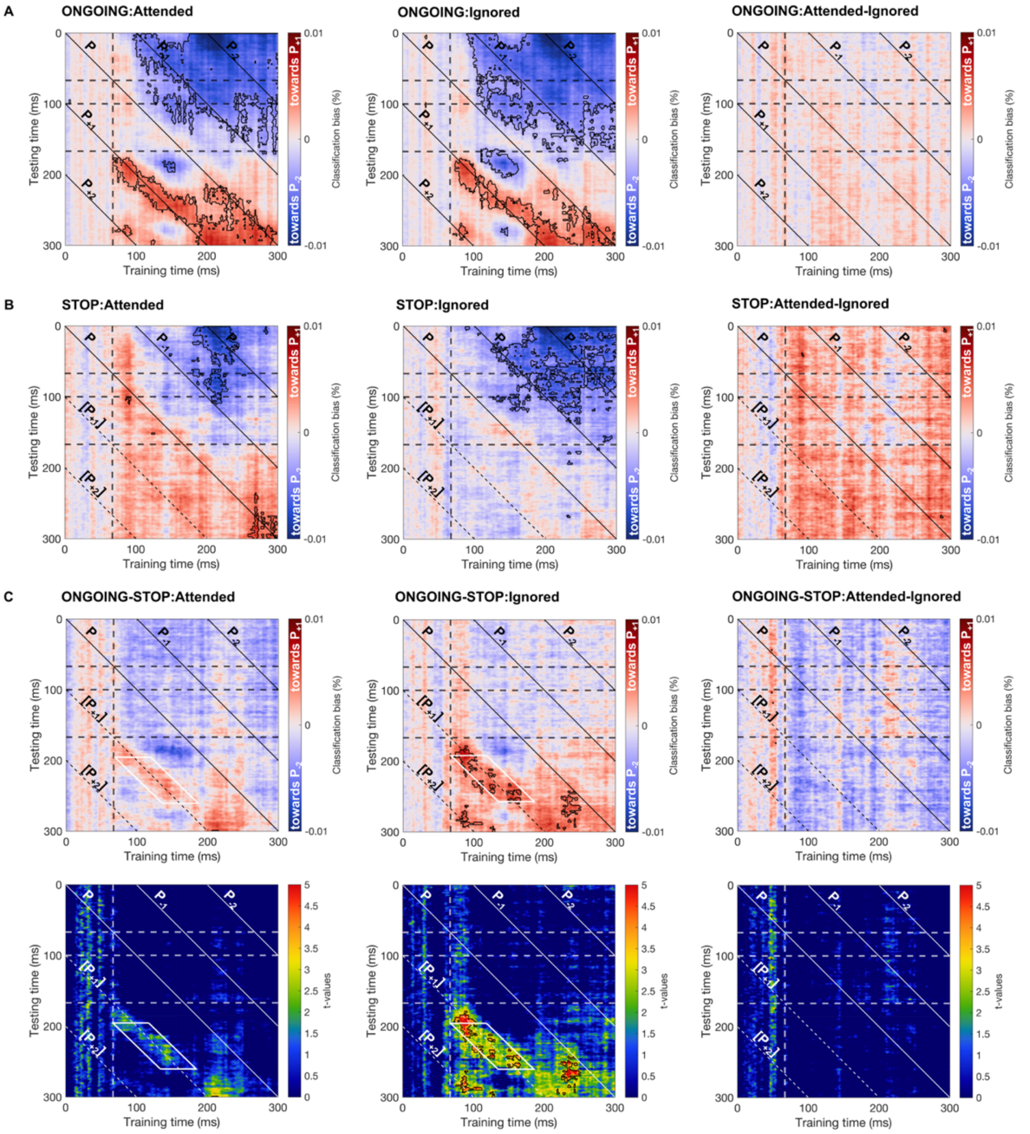

Figure S6.1. Replicated main results. (**A**) An ensemble of classifiers was trained to distinguish P_{+1} and P_{-2} (135° apart) over [0–300 ms] in the Single flashes condition were tested on P in the ONGOING:Attended (*left*) and ONGOING:Ignored (*right*) conditions. The classification assignments are summarized as a TGM, relative to the onset time of the stimulus at P (y-axis). The dashed lines at 67 ms correspond to the offset time of the stimulus at P, and the dashed horizontal lines at 100 ms and 167 ms (y-axis) indicate the onset and offset times at P_{+1}, respectively. The diagonal lines represent the position-paired classifiers for P_{-2} and P_{-1} (upper diagonal lines), P (diagonal line) and P_{+1} and P_{+2} (lower diagonal line). Solid black contours indicate clusters of classification bias that differ significantly from zero (one-sample *t*-test, $p < 0.05$, FWE-corrected). Negative bias is illustrated in dark blue and positive bias in dark red. (**B**) Same described in **A** but for the STOP conditions. The dashed horizontal lines at 100 ms and 167 ms (y-axis) indicate the predicted (but not actual) onset and offset time at P_{+1}, respectively. Similarly, the dotted diagonal lines represent the classification decisions by the position-paired classifiers for the predicted P_{+1} and P_{+2} (lower diagonal lines). (**C**) The subtraction of Attended (*left*) and Ignored (*right*) conditions and corresponding t-values for the contrast compared to zero (one-sample *t*-test, $p < 0.05$, FWE-corrected). The white box signify a narrow band of P_{+1}-paired classifiers ± 30 ms around the diagonal.

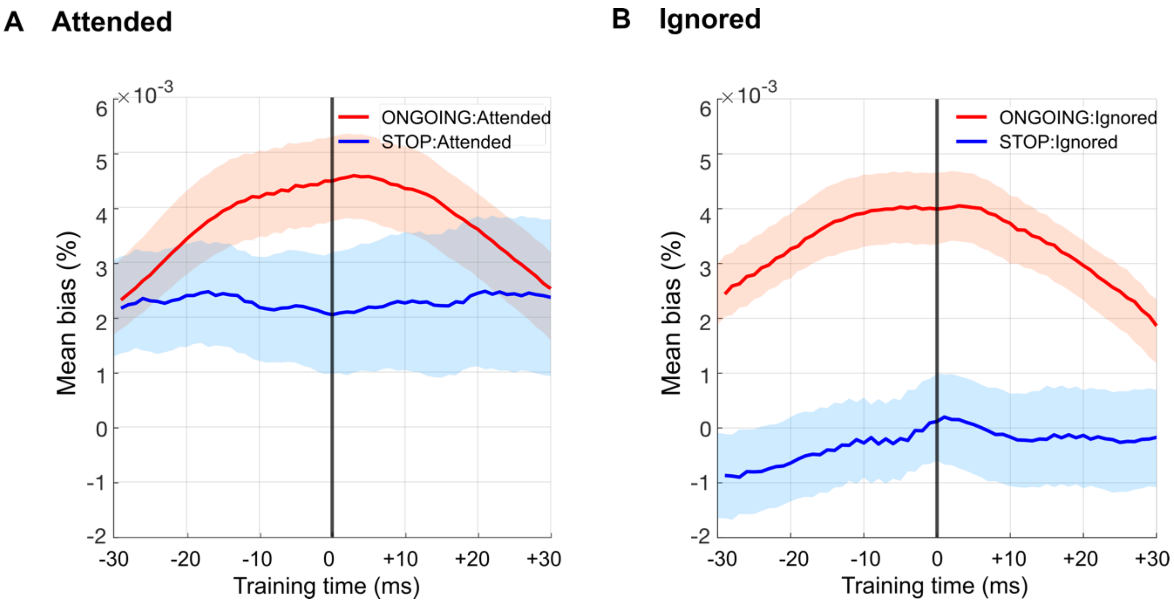

Figure S6.2. Replicated temporal pairing results. (**A**) Diagonal-averaged bias values for the Attended condition. All values corresponding to the ONGOING:Attended (orange line) and STOP:Attended (blue line) conditions are reported as means ± S.E.M. (shading). The vertical line at 0 ms represents the lower diagonal line in the TGM shown in **Figure S6.1**. (**B**) Same described in **A** but for the Ignored condition.

References

Agostino, C. S., Merkel, C., Ball, F., Vavra, P., Hinrichs, H., & Noesselt, T. (2023). Seeing and extrapolating motion trajectories share common informative activation patterns in primary visual cortex. *Human Brain Mapping*, *44*(4), 1389–1406. https://doi.org/10.1002/hbm.26123

Albright, T. D., & Stoner, G. R. (1995). Visual motion perception. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(7), 2433–2440. https://doi.org/10.1073/pnas.92.7.2433

Bae, G.-Y., & Luck, S. J. (2018). Dissociable Decoding of Spatial Attention and Working Memory from EEG Oscillations and Sustained Potentials. *The Journal of Neuroscience*, *38*(2), 409–422. https://doi.org/10.1523/JNEUROSCI.2860-17.2017

Benvenuti, G., Chemla, S., Boonman, A., Perrinet, L., Masson, G. S., & Chavane, F. (2020). *Anticipatory responses along motion trajectories in awake monkey area V1* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.03.26.010017

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature*, *398*(6725), 334–338. https://doi.org/10.1038/18678

Blom, T., Bode, S., & Hogendoorn, H. (2021). The time-course of prediction formation and revision in human visual motion processing. *Cortex*, *138*, 191–202. https://doi.org/10.1016/j.cortex.2021.02.008

Blom, T., Feuerriegel, D., Johnson, P., Bode, S., & Hogendoorn, H. (2020). Predictions drive neural representations of visual events ahead of incoming sensory information. *Proceedings of the National Academy of Sciences*, *117*(13), 7510–7515. https://doi.org/10.1073/pnas.1917777117

Blum, A., Kalai, A., & Langford, J. (1999). Beating the hold-out: Bounds for K-fold and progressive cross-validation. *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, 203–208. https://doi.org/10.1145/307400.307439

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Burkitt, A. N., & Hogendoorn, H. (2021). Predictive Visual Motion Extrapolation Emerges Spontaneously and without Supervision at Each Layer of a Hierarchical Neural Network with Spike-Timing-Dependent Plasticity. *Journal of Neuroscience*, *41*(20), 4428–4438. https://doi.org/10.1523/JNEUROSCI.2017-20.2021

Caplette, L. (2017). *Simple RM/Mixed ANOVA for any design*. MATLAB Central File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/64980-simple-rm-mixed-anova-for-any-design

Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. *Journal of Vision*, *11*(10), 9–9. https://doi.org/10.1167/11.10.9

Cheong, D., Zubieta, J.-K., & Liu, J. (2012). Neural Correlates of Visual Motion Prediction. *PLoS ONE*, *7*(6), e39854. https://doi.org/10.1371/journal.pone.0039854

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975–979. https://doi.org/10.1121/1.1907229

Cohen, M. A., Nakayama, K., Konkle, T., Stantić, M., & Alvarez, G. A. (2015). Visual Awareness Is Limited by the Representational Architecture of the Visual System. *Journal of Cognitive Neuroscience*, *27*(11), 2240–2252. https://doi.org/10.1162/jocn_a_00855

Cosman, J. D., & Vecera, S. P. (2010). Attentional capture by motion onsets is modulated by perceptual load. *Attention, Perception, & Psychophysics*, *72*(8), 2096–2105. https://doi.org/10.3758/BF03196686

Croft, R. J., & Barry, R. J. (2000). Removal of ocular artifact from the EEG: A review. *Neurophysiologie Clinique/Clinical Neurophysiology*, *30*(1), 5–19. https://doi.org/10.1016/S0987-7053(00)00055-1

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Frontiers in Human Neuroscience*, *10*, 219245. https://doi.org/10.3389/fnhum.2016.00604

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Ekman, M., Kok, P., & De Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*(1), 15276. https://doi.org/10.1038/ncomms15276

Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, *7*(2), 179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210. https://doi.org/10.1002/hbm.460020402

Friston, K. J., Jezzard, P., & Turner, R. (1994). Analysis of functional MRI time-series. *Human Brain Mapping*, *1*(2), 153–171. Scopus. https://doi.org/10.1002/hbm.460010207

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience*, *29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Hogendoorn, H., & Burkitt, A. N. (2018). Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. *NeuroImage*, *171*, 55–61. https://doi.org/10.1016/j.neuroimage.2017.12.063

Hogendoorn, H., & Burkitt, A. N. (2019). Predictive Coding with Neural Transmission Delays: A Real-Time Temporal Alignment Hypothesis. *Eneuro*, *6*(2), ENEURO.0412-18.2019. https://doi.org/10.1523/ENEURO.0412-18.2019

Hogendoorn, H., Carlson, T. A., & Verstraten, F. A. J. (2007). The time course of attentive tracking. *Journal of Vision*, *7*(14), 2. https://doi.org/10.1167/7.14.2

Holcombe, A. O. (2023). *Attending to Moving Objects* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009003414

Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex: Population coding of position in visual cortex. *The Journal of Physiology*, *556*(3), 971–982. https://doi.org/10.1113/jphysiol.2003.058941

Johnson, P. A., Blom, T., Van Gaal, S., Feuerriegel, D., Bode, S., & Hogendoorn, H. (2023). Position representations of moving objects align with real-time position in the early visual response. *eLife*, *12*, e82424. https://doi.org/10.7554/eLife.82424

Kaplan, B. A., Lansner, A., Masson, G. S., & Perrinet, L. U. (2013). Anisotropic connectivity implements motion-based prediction in a spiking neural network. *Frontiers in Computational Neuroscience*, *7*. https://doi.org/10.3389/fncom.2013.00112

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210. https://doi.org/10.1016/j.tics.2014.01.002

Kleiner, M., Brainard, D. H., Pelli, D. G., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1–16.

Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., Barnes, G., Oostenveld, R., Daunizeau, J., Flandin, G., Penny, W., & Friston, K. (2011). EEG and MEG data analysis in SPM8. *Computational Intelligence and Neuroscience*, *2011*, 852961. https://doi.org/10.1155/2011/852961

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences*, *95*(3), 825–830. https://doi.org/10.1073/pnas.95.3.825

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210. https://doi.org/10.1016/j.tics.2004.03.008

Mangun, G. R., & Hillyard, S. A. (1990). Allocation of visual attention to spatial locations: Tradeoff functions for event-related brain potentials and detection performance. *Perception & Psychophysics*, *47*(6), 532–550. https://doi.org/10.3758/bf03203106

Maunsell, J. H., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology*, *68*(4), 1332–1344. https://doi.org/10.1152/jn.1992.68.4.1332

Maus, G. W., Ward, J., Nijhawan, R., & Whitney, D. (2013). The Perceived Position of Moving Objects: Transcranial Magnetic Stimulation of Area MT+ Reduces the Flash-Lag Effect. *Cerebral Cortex*, *23*(1), 241–247. https://doi.org/10.1093/cercor/bhs021

Noah, S., Meyyappan, S., Ding, M., & Mangun, G. R. (2023). Time Courses of Attended and Ignored Object Representations. *Journal of Cognitive Neuroscience*, *35*(4), 645–658. https://doi.org/10.1162/jocn_a_01972

O'Craven, K. M., Rosen, B. R., Kwong, K. K., Treisman, A., & Savoy, R. L. (1997). Voluntary Attention Modulates fMRI Activity in Human MT–MST. *Neuron*, *18*(4), 591–598. https://doi.org/10.1016/S0896-6273(00)80300-1

Palmer, S. E., Marre, O., Berry, M. J., & Bialek, W. (2015). Predictive information in a sensory population. *Proceedings of the National Academy of Sciences*, *112*(22), 6908–6913. https://doi.org/10.1073/pnas.1506855112

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Reynolds, J. H., & Desimone, R. (1999). The Role of Neural Mechanisms of Attention in Solving the Binding Problem. *Neuron*, *24*(1), 19–29. https://doi.org/10.1016/S0896-6273(00)80819-3

Robinson, A. K., Grootswagers, T., Shatek, S. M., Gerboni, J., Holcombe, A., & Carlson, T. A. (2020). *Overlapping neural representations for the position of visible and imagined objects* (arXiv:2010.09932). arXiv. http://arxiv.org/abs/2010.09932

Sahan, M. I., Dalmaijer, E. S., Verguts, T., Husain, M., & Fias, W. (2019). The Graded Fate of Unattended Stimulus Representations in Visuospatial Working Memory. *Frontiers in Psychology*, *10*, 374. https://doi.org/10.3389/fpsyg.2019.00374

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, *79*(6), 3272–3278. https://doi.org/10.1152/jn.1998.79.6.3272

Schneider, M., Marquardt, I., Sengupta, S., Martino, F. D., & Goebel, R. (2019). *Motion Displaces Population Receptive Fields in the Direction Opposite to Motion* (p. 759183). bioRxiv. https://doi.org/10.1101/759183

Shioiri, S., Yamamoto, K., Kageyama, Y., & Yaguchi, H. (2002). Smooth shifts of visual attention. *Vision Research*, *42*(26), 2811–2816. https://doi.org/10.1016/S0042-6989(02)00405-4

Subramaniyan, M., Ecker, A. S., Patel, S. S., Cotton, R. J., Bethge, M., Pitkow, X., Berens, P., & Tolias, A. S. (2018). Faster processing of moving compared with flashed bars in awake macaque V1 provides a neural correlate of the flash lag illusion. *Journal of Neurophysiology*, *120*(5), 2430–2452. https://doi.org/10.1152/jn.00792.2017

Sundberg, K. A., Fallah, M., & Reynolds, J. H. (2006). A motion-dependent distortion of retinotopy in area V4. *Neuron*, *49*(3), 447–457. https://doi.org/10.1016/j.neuron.2005.12.023

Teichmann, L., Edwards, G., & Baker, C. I. (2021). Resolving visual motion through perceptual gaps. *Trends in Cognitive Sciences*, *25*(11), 978–991. https://doi.org/10.1016/j.tics.2021.07.017

Teichmann, L., Moerel, D., Rich, A. N., & Baker, C. I. (2022). The nature of neural object representations during dynamic occlusion. *Cortex*, *153*, 66–86. https://doi.org/10.1016/j.cortex.2022.04.009

The MathWorks Inc. (2021). *MATLAB version R2021a*. Natick, MA, USA. The MathWorks Inc. https://mathworks.com/.

The MathWorks Inc. (2023). *Statistics and Machine Learning Toolbox* (version: 12.1). Natick, MA, USA. The MathWorks Inc. https://www.mathworks.com.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Van Driel, J., Olivers, C. N. L., & Fahrenfort, J. J. (2021). High-pass filtering artifacts in multivariate classification of neural time series data. *Journal of Neuroscience Methods*, *352*, 109080. https://doi.org/10.1016/j.jneumeth.2021.109080

Volpe, B. T., Ledoux, J. E., & Gazzaniga, M. S. (1979). Information processing of visual stimuli in an ‘extinguished’ field. *Nature*, *282*(5740), Article 5740. https://doi.org/10.1038/282722a0

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data – a practical approach. *Journal of Neuroscience Methods*, *250*, 34–46. https://doi.org/10.1016/j.jneumeth.2014.08.002

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI Time-Series Revisited—Again. *NeuroImage*, *2*(3), 173–181. https://doi.org/10.1006/nimg.1995.1023

Yi, D.-J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience*, *7*(9), 992–996. https://doi.org/10.1038/nn1294

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., Poeppel, D., & Schroeder, C. E. (2013). Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a 'Cocktail Party.' *Neuron*, *77*(5), 980–991. https://doi.org/10.1016/j.neuron.2012.12.037

CHAPTER 5

General Discussion

The research presented in this thesis aimed to extend our understanding of the role of attention in predictive visual motion processing. In this final chapter, I will provide an overview of the key findings from the empirical studies in **Chapters 2, 3, and 4**. Additionally, I will address the apparent contradictions among these findings and explore further questions raised by their implications. Furthermore, I will propose potential directions for future research and highlight several limitations inherent within the studies. The chapter will conclude with brief closing remarks.

Summary of empirical findings

In **Chapter 2**, I presented the results of two experiments using the flash-lag effect (FLE) to investigate how the brain compensates for neural delays in processing moving objects. This compensation relies on prior position and motion (velocity) information to predict future positions along an expected motion trajectory (Nijhawan, 1994, 2008), allowing an extrapolated position to be perceived in the present (White, 2018). To explore these mechanisms, we manipulated perceived speed by introducing temporal noise (static vs dynamic flicker noise) and varying the luminance contrast (high vs low contrast) of the moving stimulus, while keeping its physical speed constant. Using an adaptive staircase method, we varied the positions between the moving and flash stimuli. When the flash was aligned with the moving stimulus, a perceived offset of approximately 20° was observed, representing the point of subjective equality. Our findings revealed a consistent relationship between the magnitude of the FLE and observers' assessments of motion speed compared to a reference stimulus across two textures: filtered noise and pixelated noise. FLE magnitude increased with the perception of faster motion and decreased when motion was perceived as slower, with the latter aligning more closely with the stimulus' veridical position. While changes in FLE magnitude were generally similar across both texture types, the pixelated noise condition was less sensitive to perceived speed manipulations in the FLE than filtered noise. This could be due to pixelated noise containing more high-frequency variations at the pixel level, which could be perceived as more chaotic or disruptive than the smoothed appearance of filtered noise. This increased noise might affect the FLE differently in strength. Although a trend was observed, the study may lack sufficient power to detect a statistically significant effect. The lack of a significant cross-effect correlations between FLE magnitude and perceived speed suggests that this relationship is complex and may not generalize across different textures with varying complexities. Nevertheless, this study underscored how predictive encoding of moving objects relies on perceived speed, supporting the compensatory role of motion extrapolation in addressing neural processing delays within the visual system.

In **Chapter 3**, I presented the results of one experiment extending the findings of **Chapter 2** by examining attentional resources in the context of multiple moving objects. We used a combination of spatial cueing and divided attention to explore how attention affects the FLE with multiple stimuli. Observers viewed an array of four moving stimuli distributed across four quadrants of the display. Attentional cues were used to reduce attention on each individual stimulus. Therefore, observers were required to attend to one, two, three, or all four stimuli, which were moving along independent motion trajectories, while a flash was presented adjacent to one of these stimuli. They then reported the position of the moving stimulus at the time of the flash, providing a trial-wise measure of the FLE. We found that the FLE was consistently present across all attention conditions, with magnitudes similar to those reported in **Chapter 2**, indicating that extrapolation occurs even under dynamic conditions. There was a significant effect of divided attention, resulting in an increase in median FLE magnitudes compared to when attention was focused to a single stimulus. However, there was no difference in the trial-to-trial variability of FLE magnitude across attention conditions, suggesting that divided attention augments FLE magnitude without affecting the consistency of the illusion. These findings imply that divided attention leads to increased delays in processing motion and position information for each object. As these delays accumulate across multiple objects, greater compensation may be required to monitor and update representations of the moving stimuli. This study illuminated the brain's strategy for efficiently navigating dynamic visual environments.

In **Chapter 4**, I reported on an experiment to determine whether attention is necessary for motion extrapolation. Shifting focus to apparent motion, we aimed to explore whether neural representations could be driven by prediction and whether these representations persist when moving objects are selectively attended or ignored. We presented two concurrent stimulus streams, each containing independent sequences of stimuli displayed in apparent motion. Instead of asking observers to report the perceived position of the stimulus, as in **Chapters 2 and 3**, we recorded brain activity using electroencephalography (EEG) while observers performed a color-based task that required attending to one relevant stream while ignoring the other, distractor stream. We applied multivariate pattern analyses (MVPA) to decode the neural representations of attended and ignored stimulus positions. Classifiers were trained to distinguish between neural activity patterns corresponding to different spatial positions within motion sequences: a given position (P), its previous position (P_{-1}), and its subsequent position (P_{+1}). We then tested the trained classifiers on position P during the ongoing sequence and after it stopped, tracking how position-specific information evolved within each stream. Importantly, P_{+1} was predictable from the apparent motion of each stimulus. As the sequence unfolded, we found that representations became increasingly selective. When the stream was attended, the neural activity at P contained predictive information regarding its future position P_{+1}, even when the stimulus was not physically present at that position (P being the final position of the sequence). This predictive information was comparable to stimulus-evoked representations of P_{+1} when the stimulus was actually presented. By contrast, when the stream was ignored, predictive information about P_{+1} was consistent only with actual stimulus presentation, showing no anticipation of the future position. This study demonstrated that attention biases processing toward the stimulus stream with task-relevant features, guiding motion extrapolation in conditions where streams of visual information compete for processing resources.

Theoretical implications

While some of the findings presented in this thesis may align with theories of selective attention that do not explicitly specify a role for motion-based prediction, the predictive coding framework provides a comprehensive explanation for all observed results. According to this framework, the brain continually generates prediction signals regarding incoming sensory input, which are transmitted from higher levels of the processing hierarchy to lower levels through feedback connections. These signals contain both the anticipated information (e.g., what it expects to perceive) and the precision or reliability of these predictions (Hohwy, 2012). When there is a mismatch between the predicted and observed sensory input, prediction errors occur. Attention can modulate the gain of these prediction error signals, either increasing or reducing their impact on predictions (Feldman & Friston, 2010). This modulation may be influenced by the goals of the observer or predictability of the input (Kok et al., 2012; Marzecová et al., 2017; Smout et al., 2019). In this section, I will elucidate how the results provide empirical support for the interaction with top-down predictions and bottom-up input, in accord with the predictive coding framework.

Chapters 2 and 3 illustrate that the FLE arises from the brain's ongoing effort to anticipate and predict the future positions of moving objects based on sensory evidence and prior expectations of continuous motion. When these expectations encounter sensory inputs that are not informative, such as the sudden appearance of a flash, the brain continues to anticipate motion. This can lead to perceptual errors when observers are instructed to report their perceived position of the moving object and/or flash. Findings from **Chapter 2** indicate that the extent of extrapolation increases with the perceived speed of the moving stimulus, suggesting that the model accounts for speed probability in compensating for neural delays and predicting motion trajectories. Computational studies of the FLE have shown a relationship between FLE magnitude and *physical* speed of the moving object. Wojtach et al. (2008) observed a linear increase in the FLE at slower speeds, but a non-linear trend emerged when speeds exceeded 15 dva/s. This non-linear relationship might indicate a saturation effect, where the brain reaches its maximum compensatory capacity for fixed delays in processing dynamic stimuli (Khoei et al., 2017). In line with this, we did not find a significant linear relationship between FLE magnitude and perceived speed (with a physical speed of 28.15 dva/s) in **Chapter 2**. Moreover, at this high physical speed, low contrast stimuli were perceived to move faster, resulting in a larger FLE. Although low contrast is generally associated with reduced perceived speed and processing due to reduced saliency, our study supports findings that this effect can be reversed at high physical speeds (Blakemore & Snowden, 1999; Stone & Thompson, 1992; Thompson, 1982). While attention has been shown to alter judgments of perceived speed (Turatto et al., 2007), **Chapter 2** did not explore attention effects. Given that we examined only one physical speed, it would be valuable to investigate whether the non-linear relationship varies or follows different patterns with illusory changes in perceived speed at various physical speeds.

From **Chapter 2**, it is clear that predictions use prior information to guide the processing of new sensory information. In the cricket player example illustrated in **Chapter 1**, as the cricket ball is bowled at high speed, the batter must not only track the trajectory of the ball but also monitor their limb movements, anticipate motor preparation, and identify any unexpected obstacles before playing a shot. However, due to the constraints of neural transmission and cognitive resources, the feedforward signals involved in the inference process are likely to arrive late, introducing noise or gaps in sensory data. The brain's internal generative model can address these problems within both motor (Smeets & Brenner, 1995) and perceptual systems by integrating feedback prediction signals, which are processed more quickly to update the model compared to newly arriving sensory information. When new sensory data does arrive, the brain can integrate its predictions into perception, taking into account the noise and uncertainty associated with the sensory information. **Chapter 3** demonstrates that in dynamic environments, the model accounts for increased processing delays when managing multiple moving objects with limited attentional resources. Consequently, predictions can be adjusted to the extent of extrapolation necessary to compensate for these delays.

In **Chapter 4**, we investigated how predictions are encoded in the context of apparent motion. When attention is directed towards an apparent motion stimulus, it may engage similar extrapolation mechanisms as those involved in smooth motion. However, unlike the smooth propagation of motion signals through feedforward and horizontal connections in neighboring neurons seen in smooth motion, apparent motion primarily relies on predictive mechanisms facilitated by feedback connections. This difference arises because neurons representing successive positions of the apparent motion are spaced further apart due to larger intervals between stimuli compared to the more continuous intervals in smooth motion (Blom et al., 2020, 2021). These feedback connections predict motion between the distant positions to create a cohesive and coherent perception of motion. Additionally, regularities in apparent motion (i.e., stimulus presentation every 100 ms at successive positions) facilitates the anticipatory allocation of attention to future positions along the motion trajectory, in line with the model's predictions.

Distracting information, such as motion-based regularities of the concurrently presented stream, can disrupt the model by introducing noise or bias to the processing of sensory data. This interference poses significant challenges as it may result in ambiguous representations of sensory information (Luck & Ford, 1998). To effectively manage these regularities, the model is likely to influence the neural coding, i.e., of position-specific information. This adaptability would be highly relevant in competitive and attention-demanding conditions, as presented in **Chapter 4**. In less competitive conditions involving single stimuli, prior studies such as Blom et al. (2020) observed pre-activation of P_{+1} as early as [70–90 ms] before information about P arrived. Such latency advantages served to reduce processing latency, essentially compensating for neural delays. However, unlike these findings, we did not observe pre-activations within the neural signal *before* the expected stimulus presentation at P. Instead, pre-activations were precisely aligned with the expected moment of arrival of sensory information for P_{+1}, consistent with other studies (Hsu et al., 2014; Kok et al., 2014). It would be insightful to replicate the latency advantages observed ahead of neural delays and investigate whether this interference scales with the complexity of the distractor stream. If the brain can effectively ignore an unpredictable distractor stream and still observe latency advantages, these dynamics would illuminate how our visual system navigates through complex stimuli to maintain a meaningful and coherent representation of motion across various dynamic environments.

One implication of this is the model's capability to encode changes in motion when the object velocity diverges from the expected trajectory. These deviations would typically result in prediction errors, but a number of studies have shown that the brain updates its internal model to respond to these changes in motion with a smaller delay. This has been demonstrated in the FLE. For instance, when the moving object reverses its direction at the time of the flash, the perceived position of the moving object aligns with the motion reversal (Whitney & Murakami, 1998). Depending on when the reversal occurs relative to the timing of the flash, the perceived position of the moving object may overshoot – except when the reversal occurs approximately 30 ms before the flash, their positions are perceived to be aligned, indicating no flash-lag (Eagleman & Sejnowski, 2000). These findings are congruent with observations in apparent motion, where the visual system anticipates the object's future positions along its previous motion trajectory before sensory information about the reversal is processed (Blom et al., 2020). These results suggest that the arrival of sensory data prompts the brain to rapidly recalibrate its predictions to accommodate the motion reversal. This adaptability allows the brain to perceive the position of the moving object following its new motion, ensuring that the model remains accurate and responsive to dynamic information that may be unpredictable (Duncan, 2001; Eagleman & Sejnowski, 2000; Khoei et al., 2017).

Thus far, I have discussed how our findings align with the interaction with top-down predictions and bottom-up input within the predictive coding framework. Although this thesis explores the relationship between attention and predictive motion processing, it is clear that mechanisms underlying the FLE extend beyond motion extrapolation alone. Temporal accounts, such as the involvement of horizontal connections and postdiction, provide additional perspectives (Eagleman & Sejnowski, 2000; Holcombe & Corbett, 2023; Murakami, 2001; Whitney et al., 2000) (see **Flash-lag effect**). While further research is needed to fully understand all complexities, and this thesis does not aim to adjudicate between theories, the three experiments presented can be understood within a comprehensive framework that includes motion extrapolation.

Reconciling effects of attention

Chapters 3 and 4 shed light on two complementary yet contrasting effects of attention in dynamic visual environments.

Chapter 3 examined how participants divide their attention across multiple moving objects in the FLE, demonstrating that when attention is spread thin, the brain focuses on maximizing efficiency by processing as much relevant information as possible within its constraints. This approach enables the brain to manage multiple trajectories in the environment, although it may compromise the precision of motion extrapolation (i.e., overcompensation) due to divided attentional resources.

In contrast, **Chapter 4** explored the effects of selective attention in apparent motion, revealing that focusing attention on a specific stimulus stream facilitated the prediction of future positions and allowed for more effective suppression of irrelevant information from the distractor stream. This study highlighted that, in scenarios requiring selective focus, accuracy in representation is prioritized over speed of processing.

The absence of motion extrapolation evidence when a stream is ignored underscores the critical role of attention in predictive motion processing. This finding suggests that motion extrapolation is not a fully automatic process; rather, it relies significantly on attention to enhance the brain's capacity to integrate motion information and enable extrapolation. Thus, when an apparent motion stimulus is ignored, the brain may perceive as a series of discrete events rather than a continuous trajectory, effectively preventing the extrapolation of future positions.

While the findings from the two chapters may seem contradictory – showing greater extrapolation with divided attention in **Chapter 3** and less with selective attention in **Chapter 4** – they actually reveal how attention guides extrapolation based on task demands and the observer's goals. This suggests that attention is flexibly employed to process and perceive dynamic information across various environments, with the brain using strategies to compensate for its limited capacity and processing delays.

Further research is needed to understand the neural mechanisms underlying these attentional strategies and their impact on perceptual processing (Nebel et al., 2005). Different brain networks are involved in divided and selective attention, although their relationship can vary with the type of stimuli and cognitive processes such as memory and decision-making (Corbetta et al., 1990). Some studies indicate that similar selective attention networks in the dorsal fronto-parietal areas, including the frontal eye fields and intraparietal sulcus, are activated differently depending on attentional demands (Fagioli & Macaluso, 2009; Hahn et al., 2008), suggesting multiple spotlights of attention (McMains & Somers, 2004). Investigating these neural basis within the paradigms presented in this thesis could enhance our understanding of how the brain adapts to optimize perceptual processing across diverse environments and reveal the temporal dynamics of attended versus ignored representations (Noah et al., 2023).

Altered attentional states

Altered states of attention, such as heightened alertness or increased sleepiness, are likely to influence our perception of motion. While variations in sleep quantity or quality can impact overall behavioral performance, I imagine that their effects on the FLE (**Chapters 2 and 3**) or classification performance (**Chapter 4**) would increase variability between individuals or are averaged out within general healthy populations. However, exploring this research question in clinical populations with attention deficits, such as those with visuospatial neglect, could provide valuable insights.

Visuospatial neglect is a common neurological condition following predominantly right-hemisphere damage to the ventral fronto-parietal cortex, characterized by attentional deficiencies in the contralesional visual space (Corbetta et al., 2005). Individuals with neglect frequently exhibit an inability to attend, perceive, or represent information in the contralesional hemifield. For example, they may overlook objects positioned on the contralesional side of their visual field or words located on the contralesional side of a page. In experimental settings, they typically demonstrate omissions of certain features on the contralesional side when reproducing or sketching simple figures, such as a clock.

The neglect of the contralesional visual field has been examined in a motion illusion known as the flash-grab effect (Cavanagh & Anstis, 2013). In this illusion, the background of an object rotates and reverses direction, creating a percept of back-and-forth motion. When a static object is flashed on top of the moving pattern at the moment of reversal, the flash appears displaced in the direction of motion. Notably, the impact of motion *after* the flash on its perceived position highlights that the flash-grab effect arises due to a corrective mechanism linked to motion extrapolation (Blom et al., 2019). In a recent study, De Vito et al. (2015) compared the flash-grab effect across different locations and found that individuals with neglect do not experience the illusion in their neglected visual field. They perceived the flash close to or in its veridical position, unlike healthy controls who consistently misperceived the flash's position in the direction of motion reversal, irrespective of its location. This suggests that motion extrapolation mechanisms, which typically influence our perception of motion-induced position shifts, do not significantly impact the perception of the flash-grab effect in individuals with neglect. Despite these findings, several studies indicate that individuals with neglect still perceive motion without significant differences in sensitivity to direction, albeit with difficulty in judging the motion moving towards or away from the neglected visual field (Spinelli & Zoccolotti, 1992). Additionally, in visual search tasks, they can detect targets entering their neglected visual field more readily than static targets appearing in that visual field, suggesting some degree of motion extrapolation at play (Machner et al., 2012). Moreover, the effects of neglect seem to vary based on task demands (Bonato, 2012), and individuals with neglect also exhibit interference from competing information in visual processing. For example, Corben et al. (2001) observed that visually guided movements towards a contralesional target experience greater behavioral interference from distractors on the ipsilesional side of the target. The paradigms presented in this thesis offer promising avenues for exploring the mechanisms of predictive motion processing in individuals with neglect, especially those involving motion trajectories that traverse from one visual field to another.

Limitations

In **Chapter 1**, I provided an overview of various types of attention, including spatial attention, non-spatial feature-based attention, and object-based attention. However, in subsequent chapters, it was challenging to disentangle these attention types: **Chapter 3** examined spatial attention (cued quadrants) and object-based attention (moving bars) within attended locations, and **Chapter 4** manipulated feature-based attention (colors) while apparent motion facilitated allocation of spatial attention (future positions). In the real world, distinguishing between attention types is also challenging because they often interact, and motion presents a special case where changes in multiple features and locations are relevant. In principle, future research could employ paradigms designed to isolate individual attentional manipulations to better understand their respective effects on motion extrapolation. For instance, investigations similar to **Chapter 3** could be adapted to present multiple bars of different colors rotating atop each other along independent motion trajectories, comparable to the manipulation in **Chapter 4**. This would allow us to explore whether observers can predict the motion of overlapping objects simultaneously, or how the FLE might vary under conditions of divided or selective attention.

More specific to **Chapter 4**, one limitation is that the observed classification values were very small. While small effects have also been observed in similar decoding studies with single stimuli (Blom et al., 2020; Johnson et al., 2023), the additional noise introduced by the two stimulus streams in our study contributed to even weaker neural signals.

Additionally, our decoding analyses unveiled latent information concealed within the polarity changes of the TGMs. This study aimed to determine the presence or absence of predictive information of a stimulus' future position P_{+1} in the neural representation of its current position P, without knowing precisely when these representations first emerged due to latency advantages. Notably, we observed a qualitative reduction in the representation strength pertaining to the preceding position P_{-1} in the STOP condition compared to the ONGOING condition, and for the attended stimulus compared to the ignored stimulus. While this observation was not subjected to statistical testing, it nonetheless suggests a compelling pattern that warrants further investigation. One plausible explanation is interference from the representation of P_{+1}. The positive and negative biases could potentially nullify each other, obscuring the precise timing of when representations of P_{+1} are activated. It is possible that lingering information of the past sensory input P_{-1} might be overridden, or inhibited to optimize processing resources in anticipation of the future input P_{+1}. This interpretation was further supported by replication analyses, where the distance between trained positions P_{+1} and P_{-2} was wider than in the main analyses. However, in both cases, our methodology could not definitively pinpoint when representations were first pre-activated, indicating latency advantages, in addition to the weaker signals noted in the manuscript.

Finally, a general challenge within EEG refers to its poor spatial resolution. It would be beneficial to explore how signals from different brain regions are utilized at various time-points. Several studies have investigated apparent motion using fMRI (Maffei et al., 2010) and TMS (Donato et al., 2020; Vetter et al., 2015), and integrating these approaches with the current paradigm could further clarify the structure-function relationships between attention and predictive networks involved in motion perception.

Conclusion

In this thesis, I have explored the ways in which attention and prediction shape our perception of moving objects. Using a combination of behavioral and EEG experiments, we investigated how the brain strategically allocates attention during motion processing within dynamic visual environments. Our findings reveal that motion extrapolation is an obligatory consequence of motion processing (**Chapter 2**), playing a crucial role in compensating for neural delays and maintaining a coherent perception of the visual world. Moreover, we elucidate how attentional resources modulate extrapolation mechanisms, demonstrating the brain's capability to simultaneously monitor and predict the motion of multiple objects, especially when sensory information is delayed (**Chapter 3**). In environments where there is a continuous influx of information competing for processing, selective attention guides motion extrapolation to prioritize objects that are most relevant to behavior, facilitating perception and prediction of motion for task-relevant objects while suppressing distractors (**Chapter 4**). These findings are well accounted for by predictive coding theories of perception and open avenues for future exploration with potential experiments outlined in this chapter to advance our understanding of the brain's ability to make sense of dynamic information and navigate our visual world.

References

Adamian, N., & Andersen, S. K. (2022). Attentional Enhancement of Tracked Stimuli in Early Visual Cortex Has Limited Capacity. *The Journal of Neuroscience*, *42*(46), 8709–8715. https://doi.org/10.1523/JNEUROSCI.0605-22.2022

Agostino, C. S., Merkel, C., Ball, F., Vavra, P., Hinrichs, H., & Noesselt, T. (2023). Seeing and extrapolating motion trajectories share common informative activation patterns in primary visual cortex. *Human Brain Mapping*, *44*(4), 1389–1406. https://doi.org/10.1002/hbm.26123

Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219–227. https://doi.org/10.1016/j.conb.2017.08.010

Aitken, F., Menelaou, G., Warrington, O., Koolschijn, R. S., Corbin, N., Callaghan, M. F., & Kok, P. (2020). Prior expectations evoke stimulus-specific activity in the deep layers of the primary visual cortex. *PLOS Biology*, *18*(12), e3001023. https://doi.org/10.1371/journal.pbio.3001023

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology: CB*, *23*(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, *52*(6), 1106–1130. https://doi.org/10.1152/jn.1984.52.6.1106

Anderson, C. H., & Van Essen, D. C. (1987). Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Sciences*, *84*(17), 6297–6301. https://doi.org/10.1073/pnas.84.17.6297

Apthorp, D., Schwarzkopf, D. S., Kaul, C., Bahrami, B., Alais, D., & Rees, G. (2013). Direct evidence for encoding of motion streaks in human visual cortex. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1752), 20122339. https://doi.org/10.1098/rspb.2012.2339

Arazi, A., Yeshurun, Y., & Dinstein, I. (2019). Neural Variability Is Quenched by Attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(30), 5975–5985. https://doi.org/10.1523/JNEUROSCI.0355-19.2019

Arnold, D. H., Durant, S., & Johnston, A. (2003). Latency differences and the flash-lag effect. *Vision Research*, *43*(17), 1829–1835. https://doi.org/10.1016/S0042-6989(03)00281-5

Arnold, D. H., Ong, Y., & Roseboom, W. (2009). Simple differential latencies modulate, but do not cause the flash-lag effect. *Journal of Vision*, *9*(5), 4–4. https://doi.org/10.1167/9.5.4

Baldo, M. V. C., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an Attentional Component of the Perceptual Misalignment between Moving and Flashing Stimuli. *Perception*, *31*(1), 17–30. https://doi.org/10.1068/p3302

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature*, *378*(6557), Article 6557. https://doi.org/10.1038/378565a0

Baldo, M. V. C., & Klein, S. A. (2010). Paying attention to the flash-lag effect. In R. Nijhawan & B. Khurana (Eds.), *Space and Time in Perception and Action* (1st ed., pp. 396–407). Cambridge University Press. https://doi.org/10.1017/CBO9780511750540.023

Battistoni, E., Stein, T., & Peelen, M. V. (2017). Preparatory attention in visual cortex: Preparatory attention in visual cortex. *Annals of the New York Academy of Sciences*, *1396*(1), 92–107. https://doi.org/10.1111/nyas.13320

Benvenuti, G., Chemla, S., Boonman, A., Perrinet, L., Masson, G. S., & Chavane, F. (2020). *Anticipatory responses along motion trajectories in awake monkey area V1* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.03.26.010017

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature*, *398*(6725), 334–338. https://doi.org/10.1038/18678

Blakemore, M. R., & Snowden, R. J. (1999). The Effect of Contrast upon Perceived Speed: A General Phenomenon? *Perception*, *28*(1), 33–48. https://doi.org/10.1068/p2722

Blom, T., Bode, S., & Hogendoorn, H. (2021). The time-course of prediction formation and revision in human visual motion processing. *Cortex*, *138*, 191–202. https://doi.org/10.1016/j.cortex.2021.02.008

Blom, T., Feuerriegel, D., Johnson, P., Bode, S., & Hogendoorn, H. (2020). Predictions drive neural representations of visual events ahead of incoming sensory information.

Proceedings of the National Academy of Sciences, *117*(13), 7510–7515. https://doi.org/10.1073/pnas.1917777117

Blom, T., Liang, Q., & Hogendoorn, H. (2019). When predictions fail: Correction for extrapolation in the flash-grab effect. *Journal of Vision*, *19*(2), 3. https://doi.org/10.1167/19.2.3

Bonato, M. (2012). Neglect and Extinction Depend Greatly on Task Demands: A Review. *Frontiers in Human Neuroscience*, *6*, 195. https://doi.org/10.3389/fnhum.2012.00195

Braddick, O. (2001). Occipital Lobe (Visual Cortex): Functional Aspects. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 10826–10828). Pergamon. https://doi.org/10.1016/B0-08-043076-7/03470-7

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash–lag effect. *Vision Research*, *40*(13), 1645–1648. https://doi.org/10.1016/S0042-6989(00)00067-5

Broadbent, D. E. (1958). *Perception and Communication*. Pergamon Press.

Büchel, C. (1997). Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex*, *7*(8), 768–778. https://doi.org/10.1093/cercor/7.8.768

Büchel, C., Josephs, O., Rees, G., Turner, R., Frith, C. D., & Friston, K. J. (1998). The functional anatomy of attention to visual motion. A functional MRI study. *Brain*, *121*(7), 1281–1294. https://doi.org/10.1093/brain/121.7.1281

Burkitt, A. N., & Hogendoorn, H. (2021). Predictive Visual Motion Extrapolation Emerges Spontaneously and without Supervision at Each Layer of a Hierarchical Neural Network with Spike-Timing-Dependent Plasticity. *Journal of Neuroscience*, *41*(20), 4428–4438. https://doi.org/10.1523/JNEUROSCI.2017-20.2021

Burr, D. C. (1979). Acuity for apparent vernier offset. *Vision Research*, *19*(7), 835–837. https://doi.org/10.1016/0042-6989(79)90162-7

Carlson, T. A., Grootswagers, T., & Robinson, A. K. (2020). *An introduction to time-resolved decoding analysis for M/EEG* (pp. 679–690). https://researchdirect.westernsydney.edu.au/islandora/object/uws%3A59844/

Carrasco, M. (2018). How visual spatial attention alters perception. *Cognitive Processing*, *19*(Suppl 1), 77–88. https://doi.org/10.1007/s10339-018-0883-4

Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., & Goebel, R. (2002). Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(21), 13914–13919. https://doi.org/10.1073/pnas.202049999

Cavanagh, P., & Anstis, S. (2013). The flash grab effect. *Vision Research*, *91*, 8–20. https://doi.org/10.1016/j.visres.2013.07.007

Cavanagh, P., Caplovitz, G. P., Lytchenko, T. K., Maechler, M. R., Tse, P. U., & Sheinberg, D. L. (2023). The Architecture of Object-Based Attention. *Psychonomic Bulletin & Review*, *30*(5), 1643–1667. https://doi.org/10.3758/s13423-023-02281-7

Changizi, M. A., Hsieh, A., Nijhawan, R., Kanai, R., & Shimojo, S. (2008). Perceiving the Present and a Systematization of Illusions. *Cognitive Science*, *32*(3), 459–503. https://doi.org/10.1080/03640210802035191

Chappell, M., Hine, T. J., Acworth, C., & Hardwick, D. R. (2006). Attention ‘capture’ by the flash-lag flash. *Vision Research*, *46*(19), 3205–3213. https://doi.org/10.1016/j.visres.2006.04.017

Cheong, D., Zubieta, J.-K., & Liu, J. (2012). Neural Correlates of Visual Motion Prediction. *PLoS ONE*, *7*(6), e39854. https://doi.org/10.1371/journal.pone.0039854

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975–979. https://doi.org/10.1121/1.1907229

Chong, E., Familiar, A. M., & Shim, W. M. (2016). Reconstructing representations of dynamic visual objects in early visual cortex. *Proceedings of the National Academy of Sciences*, *113*(5), 1453–1458. https://doi.org/10.1073/pnas.1512144113

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Cohen, M. S., Kosslyn, S. M., Breiter, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., Brookheimer, S. Y., Rosen, B. R., & Belliveau, J. W. (1996). Changes in cortical activity during mental rotation. A mapping study using functional MRI. *Brain: A Journal of Neurology*, *119 (Pt 1)*, 89–100. https://doi.org/10.1093/brain/119.1.89

Compte, A., & Wang, X.-J. (2006). Tuning curve shift by attention modulation in cortical neurons: A computational study of its mechanisms. *Cerebral Cortex (New York, N.Y.: 1991)*, *16*(6), 761–778. https://doi.org/10.1093/cercor/bhj021

Corben, L. A., Mattingley, J. B., & Bradshaw, J. L. (2001). A kinematic analysis of distractor interference effects during visually guided action in spatial neglect. *Journal of the International Neuropsychological Society*, *7*(3), 334–343. https://doi.org/10.1017/S1355617701733073

Corbetta, M., Kincade, M. J., Lewis, C., Snyder, A. Z., & Sapir, A. (2005). Neural basis and recovery of spatial attention deficits in spatial neglect. *Nature Neuroscience*, *8*(11), Article 11. https://doi.org/10.1038/nn1574

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional Modulation of Neural Processing of Shape, Color, and Velocity in Humans. *Science*, *248*(4962), 1556–1559. https://doi.org/10.1126/science.2360050

Culham, J. C., Brandt, S. A., Cavanagh, P., Kanwisher, N. G., Dale, A. M., & Tootell, R. B. H. (1998). Cortical fMRI Activation Produced by Attentive Tracking of Moving Targets. *Journal of Neurophysiology*, *80*(5), 2657–2670. https://doi.org/10.1152/jn.1998.80.5.2657

Culham, J. C., He, S., Dukelow, S., & Verstraten, F. A. J. (2001). Visual motion and the human brain: What has neuroimaging told us? *Acta Psychologica*, *107*(1–3), 69–94. https://doi.org/10.1016/S0001-6918(01)00022-1

De Vito, S., Lunven, M., Bourlon, C., Duret, C., Cavanagh, P., & Bartolomeo, P. (2015). When brain damage “improves” perception: Neglect patients can localize motion-shifted probes better than controls. *Journal of Neurophysiology*, *114*(6), 3351–3358. https://doi.org/10.1152/jn.00757.2015

Demarchi, G., Sanchez, G., & Weisz, N. (2019). Automatic and feature-specific prediction-related neural activity in the human auditory system. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-11440-1

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, *11*(5), 219–226. https://doi.org/10.1016/0166-2236(88)90130-0

Donato, R., Pavan, A., Nucci, M., & Campana, G. (2020). The neural mechanisms underlying directional and apparent circular motion assessed with repetitive transcranial magnetic stimulation (rTMS). *Neuropsychologia*, *149*, 107656. https://doi.org/10.1016/j.neuropsychologia.2020.107656

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*(4), 501–517. https://doi.org/10.1037/0096-3445.113.4.501

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, *2*(11), Article 11. https://doi.org/10.1038/35097575

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion Integration and Postdiction in Visual Awareness. *Science*, *287*(5460), 2036–2038. https://doi.org/10.1126/science.287.5460.2036

Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias localization judgments: A unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. *Journal of Vision*, *7*(4), 3–3. https://doi.org/10.1167/7.4.3

Ekman, M., Kok, P., & De Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*(1), 15276. https://doi.org/10.1038/ncomms15276

Ekman, M., Roelfsema, P. R., & Lange, F. P. de. (2020). Object Selection by Automatic Spreading of Top-Down Attentional Signals in V1. *Journal of Neuroscience*, *40*(48), 9250–9259. https://doi.org/10.1523/JNEUROSCI.0438-20.2020

Fagioli, S., & Macaluso, E. (2009). Attending to Multiple Visual Streams: Interactions between Location-based and Category-based Attentional Selection. *Journal of Cognitive Neuroscience*, *21*(8), 1628–1641. https://doi.org/10.1162/jocn.2009.21116

Feldman, H., & Friston, K. (2010). Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, *4*. https://www.frontiersin.org/articles/10.3389/fnhum.2010.00215

Fox, K. J., Birman, D., & Gardner, J. L. (2023). Gain, not concomitant changes in spatial receptive field properties, improves task performance in a neural network attention model. *eLife*, *12*, e78392. https://doi.org/10.7554/eLife.78392

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126–132. https://doi.org/10.1037/0278-7393.10.1.126

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325–1352. https://doi.org/10.1016/j.neunet.2003.06.005

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2008). Hierarchical Models in the Brain. *PLOS Computational Biology*, *4*(11), e1000211. https://doi.org/10.1371/journal.pcbi.1000211

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Fröhlich, F. W. (1923). Uber die messung der empfindungszeit [Measuring the time of sensation]. *Zeitschrift Für Sinnesphysiologie*, *54*, 58–78.

Galashan, F. O., Saßen, H. C., Kreiter, A. K., & Wegener, D. (2013). Monkey area MT latencies to speed changes depend on attention and correlate with behavioral reaction times. *Neuron*, *78*(4), 740–750. https://doi.org/10.1016/j.neuron.2013.03.014

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision*, *7*(8), 9. https://doi.org/10.1167/7.8.9

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998). The constructive nature of vision: Direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery: fMRI studies of apparent motion and motion imagery. *European Journal of Neuroscience*, *10*(5), 1563–1573. https://doi.org/10.1046/j.1460-9568.1998.00181.x

Grill-Spector, K., & Malach, R. (2004). THE HUMAN VISUAL CORTEX. *Annual Review of Neuroscience*, *27*(1), 649–677. https://doi.org/10.1146/annurev.neuro.27.070203.144220

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience*, *29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Hahn, B., Wolkenberg, F. A., Ross, T. J., Myers, C. S., Heishman, S. J., Stein, D. J., Kurup, P. K., & Stein, E. A. (2008). Divided versus selective attention: Evidence for common processing mechanisms. *Brain Research*, *1215*, 137–146. https://doi.org/10.1016/j.brainres.2008.03.058

Harrison, L. M., Stephan, K. E., Rees, G., & Friston, K. J. (2007). Extra-classical receptive field effects measured in striate cortex with fMRI. *NeuroImage*, *34*(3), 1199–1208. https://doi.org/10.1016/j.neuroimage.2006.10.017

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), Article 7238. https://doi.org/10.1038/nature07832

Harvey, B. M., & Dumoulin, S. O. (2016). Visual motion transforms visual space representations similarly throughout the human visual hierarchy. *NeuroImage*, *127*, 173–185. https://doi.org/10.1016/j.neuroimage.2015.11.070

Hogendoorn, H. (2020). Motion Extrapolation in Visual Processing: Lessons from 25 Years of Flash-Lag Debate. *The Journal of Neuroscience*, *40*(30), 5698–5705. https://doi.org/10.1523/JNEUROSCI.0275-20.2020

Hogendoorn, H., & Burkitt, A. N. (2018). Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. *NeuroImage*, *171*, 55–61. https://doi.org/10.1016/j.neuroimage.2017.12.063

Hogendoorn, H., & Burkitt, A. N. (2019). Predictive Coding with Neural Transmission Delays: A Real-Time Temporal Alignment Hypothesis. *Eneuro*, *6*(2), ENEURO.0412-18.2019. https://doi.org/10.1523/ENEURO.0412-18.2019

Hohwy, J. (2012). Attention and Conscious Perception in the Hypothesis Testing Brain. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00096

Holcombe, A. O., & Corbett, J. (2023). *Temporal errors: Researchers should stop studying the flash-lag effect*. https://doi.org/10.31234/osf.io/swzr7

Howe, P. D. L., & Holcombe, A. O. (2012). Motion information is sometimes used as an aid to the visual tracking of objects. *Journal of Vision*, *12*(13), 10–10. https://doi.org/10.1167/12.13.10

Hsu, Y.-F., Hamalainen, J. A., & Waszak, F. (2014). Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00152

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews. Cognitive Science*, *2*(5), 580–593. https://doi.org/10.1002/wcs.142

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, *12*(5), 822–851. https://doi.org/10.3758/BF03196775

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

James, W. (1950). *The Principles of Psychology* (Vol. 2). Dover Publications Inc.

Jancke, D., & Erlhagen, W. (2010). Bridging the gap: A model of common neural mechanisms underlying the Fröhlich effect, the flash-lag effect, and the representational momentum effect. In R. Nijhawan & B. Khurana (Eds.), *Space and Time in Perception and Action* (1st ed., pp. 422–440). Cambridge University Press. https://doi.org/10.1017/CBO9780511750540.025

Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex: Population coding of position in visual cortex. *The Journal of Physiology*, *556*(3), 971–982. https://doi.org/10.1113/jphysiol.2003.058941

Jensen, A. R., & Rohwer, W. D. (1966). The stroop color-word test: A review. *Acta Psychologica*, *25*, 36–93. https://doi.org/10.1016/0001-6918(66)90004-7

Jiang, L. P., & Rao, R. P. N. (2023). *Dynamic Predictive Coding: A Model of Hierarchical Sequence Learning and Prediction in the Neocortex* (p. 2022.06.23.497415). bioRxiv. https://doi.org/10.1101/2022.06.23.497415

Johnson, P. A., Blom, T., Van Gaal, S., Feuerriegel, D., Bode, S., & Hogendoorn, H. (2023). Position representations of moving objects align with real-time position in the early visual response. *eLife*, *12*, e82424. https://doi.org/10.7554/eLife.82424

Johnston, J., & Lagnado, L. (2015). General features of the retinal connectome determine the computation of motion anticipation. *eLife*, *4*, e06250. https://doi.org/10.7554/eLife.06250

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: Spatial uncertainty is the key to perceptual mislocalization. *Vision Research*, *44*(22), 2605–2619. https://doi.org/10.1016/j.visres.2003.10.028

Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The Flash-Lag Effect as a Motion-Based Predictive Shift. *PLOS Computational Biology*, *13*(1), e1005068. https://doi.org/10.1371/journal.pcbi.1005068

Khurana, B., & Nijhawan, R. (1995). Extrapolation or attention shift? *Nature*, *378*(6557), Article 6557. https://doi.org/10.1038/378566a0

Khurana, B., Watanabe, K., & Nijhawan, R. (2000). The Role of Attention in Motion Extrapolation: Are Moving Objects 'Corrected' or Flashed Objects Attentionally Delayed? *Perception*, *29*(6), 675–692. https://doi.org/10.1068/p3066

Klein, B. P., Harvey, B. M., & Dumoulin, S. O. (2014). Attraction of position preference by spatial attention throughout human visual cortex. *Neuron*, *84*(1), 227–237. https://doi.org/10.1016/j.neuron.2014.08.047

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719. https://doi.org/10.1016/j.tins.2004.10.007

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. *Current Biology*, *26*(3), 371–376. https://doi.org/10.1016/j.cub.2015.12.038

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience*, *26*(7), 1546–1554. https://doi.org/10.1162/jocn_a_00562

Kok, P., Mostert, P., & de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences*, *114*(39), 10473–10478. https://doi.org/10.1073/pnas.1705652114

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & De Lange, F. P. (2012). Attention Reverses the Effect of Prediction in Silencing Sensory Signals. *Cerebral Cortex*, *22*(9), 2197–2206. https://doi.org/10.1093/cercor/bhr310

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research*, *39*(16), 2669–2679. https://doi.org/10.1016/S0042-6989(98)00287-9

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Research*, *40*(2), 201–215. https://doi.org/10.1016/S0042-6989(99)00168-6

Kutschireiter, A., Surace, S. C., Sprekeler, H., & Pfister, J.-P. (2017). Nonlinear Bayesian filtering and learning: A neuronal dynamics for perception. *Scientific Reports*, *7*, 8722. https://doi.org/10.1038/s41598-017-06519-y

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–579. https://doi.org/10.1016/S0166-2236(00)01657-X

Lamme, V. A., Supèr, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8*(4), 529–535. https://doi.org/10.1016/s0959-4388(98)80042-1

Laycock, R., Crewther, D. P., Fitzgerald, P. B., & Crewther, S. G. (2007). Evidence for Fast Signals and Later Processing in Human V1/V2 and V5/MT+: A TMS Study of Motion Perception. *Journal of Neurophysiology*, *98*(3), 1253–1262. https://doi.org/10.1152/jn.00416.2007

Liu, B., Hong, A., Rieke, F., & Manookin, M. B. (2021). Predictive encoding of motion begins in the primate retina. *Nature Neuroscience*, *24*(9), Article 9. https://doi.org/10.1038/s41593-021-00899-1

Liu, T., Slotnick, S. D., & Yantis, S. (2004). Human MT+ mediates perceptual filling-in during apparent motion. *NeuroImage*, *21*(4), 1772–1780. https://doi.org/10.1016/j.neuroimage.2003.12.025

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, *77*(1), 24–42. https://doi.org/10.1152/jn.1997.77.1.24

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 825–830.

Luu, T., & Howe, P. D. L. (2015). Extrapolation occurs in multiple object tracking when eye movements are controlled. *Attention, Perception, & Psychophysics*, *77*(6), 1919–1929. https://doi.org/10.3758/s13414-015-0891-8

Machner, B., Dorr, M., Sprenger, A., von der Gablentz, J., Heide, W., Barth, E., & Helmchen, C. (2012). Impact of dynamic bottom-up features and top-down control on the visual exploration of moving real-world scenes in hemispatial neglect. *Neuropsychologia*, *50*(10), 2415–2425. https://doi.org/10.1016/j.neuropsychologia.2012.06.012

Mackay, D. M. (1958). Perceptual Stability of a Stroboscopically Lit Visual Field containing Self-Luminous Objects. *Nature*, *181*(4607), 507–508. https://doi.org/10.1038/181507a0

Maffei, V., Macaluso, E., Indovina, I., Orban, G., & Lacquaniti, F. (2010). Processing of Targets in Smooth or Apparent Motion Along the Vertical in the Human Brain: An fMRI Study. *Journal of Neurophysiology*, *103*(1), 360–370. https://doi.org/10.1152/jn.00892.2009

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Current Biology*, *14*(9), 744–751. https://doi.org/10.1016/j.cub.2004.04.028

Marzecová, A., Widmann, A., SanMiguel, I., Kotz, S. A., & Schröger, E. (2017). Interrelation of attention and prediction in visual processing: Effects of task-relevance and stimulus probability. *Biological Psychology*, *125*, 76–90. https://doi.org/10.1016/j.biopsycho.2017.02.009

Masri, R. A., Grünert, U., & Martin, P. R. (2020). Analysis of Parvocellular and Magnocellular Visual Pathways in Human Retina. *Journal of Neuroscience*, *40*(42), 8132–8148. https://doi.org/10.1523/JNEUROSCI.1671-20.2020

Maunsell, J. H., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology*, *68*(4), 1332–1344. https://doi.org/10.1152/jn.1992.68.4.1332

Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. https://doi.org/10.1016/j.tins.2006.04.001

Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, *49*(5), 1127–1147. https://doi.org/10.1152/jn.1983.49.5.1127

Maus, G. W., Fischer, J., & Whitney, D. (2013). Motion-dependent representation of space in area MT+. *Neuron*, *78*(3), 554–562. https://doi.org/10.1016/j.neuron.2013.03.010

Maus, G. W., Ward, J., Nijhawan, R., & Whitney, D. (2013). The Perceived Position of Moving Objects: Transcranial Magnetic Stimulation of Area MT+ Reduces the Flash-Lag Effect. *Cerebral Cortex*, *23*(1), 241–247. https://doi.org/10.1093/cercor/bhs021

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. *Journal of Neuroscience*, *19*(1), 431–441. https://doi.org/10.1523/JNEUROSCI.19-01-00431.1999

McKeefry, D. J., Watson, J. D., Frackowiak, R. S., Fong, K., & Zeki, S. (1997). The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. *NeuroImage*, *5*(1), 1–12. https://doi.org/10.1006/nimg.1996.0246

McMains, S. A., & Somers, D. C. (2004). Multiple Spotlights of Attentional Selection in Human Visual Cortex. *Neuron*, *42*(4), 677–686. https://doi.org/10.1016/S0896-6273(04)00263-6

Millidge, B., Tang, M., Osanlouy, M., & Bogacz, R. (2023). *Predictive Coding Networks for Temporal Prediction* (p. 2023.05.15.540906). bioRxiv. https://doi.org/10.1101/2023.05.15.540906

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, *6*(1), 57–77. https://doi.org/10.1016/0166-4328(82)90081-x

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782–785.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary Visual Cortex Activity along the Apparent-Motion Trace Reflects Illusory Perception. *PLoS Biology*, *3*(8), e265. https://doi.org/10.1371/journal.pbio.0030265

Muckli, L., Kriegeskorte, N., Lanfermann, H., Zanella, F. E., Singer, W., & Goebel, R. (2002). Apparent Motion: Event-Related Functional Magnetic Resonance Imaging of Perceptual Switches and States. *The Journal of Neuroscience*, *22*(9), RC219–RC219. https://doi.org/10.1523/JNEUROSCI.22-09-j0003.2002

Murakami, I. (2001). A flash-lag effect in random motion. *Vision Research*, *41*(24), 3101–3119. https://doi.org/10.1016/S0042-6989(01)00193-6

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalizations with moving stimuli: The Fröhlich effect, the flash-lag, and representational momentum. *Visual Cognition*, *9*(1–2), 120–138. https://doi.org/10.1080/13506280143000359

Namba, J., & Baldo, M. V. C. (2004). The Modulation of the Flash-Lag Effect by Voluntary Attention. *Perception*, *33*(5), 621–631. https://doi.org/10.1068/p5212

Nebel, K., Wiese, H., Stude, P., de Greiff, A., Diener, H.-C., & Keidel, M. (2005). On the neural basis of focused and divided attention. *Brain Research. Cognitive Brain Research*, *25*(3), 760–776. https://doi.org/10.1016/j.cogbrainres.2005.09.011

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*(6487), 256–256. https://doi.org/10.1038/370256a0

Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences*, *31*(2), 179–198. https://doi.org/10.1017/S0140525X08003804

Nijhawan, R., & Wu, S. (2009). Compensating time delays with neural predictions: Are predictions sensory or motor? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *367*(1891), 1063–1078. https://doi.org/10.1098/rsta.2008.0270

Noah, S., Meyyappan, S., Ding, M., & Mangun, G. R. (2023). Time Courses of Attended and Ignored Object Representations. *Journal of Cognitive Neuroscience*, *35*(4), 645–658. https://doi.org/10.1162/jocn_a_01972

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*(1), 44–64. https://doi.org/10.1016/0010-0285(75)90004-3

Nowak, L. G., & Bullier, J. (1997). The Timing of Information Transfer in the Visual System. In K. S. Rockland, J. H. Kaas, & A. Peters (Eds.), *Extrastriate Cortex in Primates* (Vol. 12, pp. 205–241). Springer US. https://doi.org/10.1007/978-1-4757-9625-4_5

O’Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, *401*(6753), Article 6753. https://doi.org/10.1038/44134

Ororbia, A., Mali, A., Giles, C. L., & Kifer, D. (2020). Continual Learning of Recurrent Neural Networks by Locally Aligning Distributed Representations. *IEEE Transactions on Neural Networks and Learning Systems*, *31*(10), 4267–4278. https://doi.org/10.1109/TNNLS.2019.2953622

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature*, *396*(6710), 424–424. https://doi.org/10.1038/24766

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197. https://doi.org/10.1163/156856888x00122

Rahnev, D. (2019). The Bayesian brain: What is it and do humans have it? *The Behavioral and Brain Sciences*, *42*, e238. https://doi.org/10.1017/S0140525X19001377

Raiguel, S. E., Xiao, D. K., Marcar, V. L., & Orban, G. A. (1999). Response latency of macaque area MT/V5 neurons and its relationship to stimulus parameters. *Journal of Neurophysiology*, *82*(4), 1944–1956. https://doi.org/10.1152/jn.1999.82.4.1944

Ramachandran, V. S., & Anstis, S. M. (1986). The Perception of Apparent Motion. *Scientific American*, *254*(6), 102–109. https://doi.org/10.1038/scientificamerican0686-102

Rao, R. P. N., & Ballard, D. H. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation*, *9*(4), 721–763. https://doi.org/10.1162/neco.1997.9.4.721

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Rees, G., Friston, K., & Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience*, *3*(7), 716–723. https://doi.org/10.1038/76673

Reynolds, J. H., & Chelazzi, L. (2004). ATTENTIONAL MODULATION OF VISUAL PROCESSING. *Annual Review of Neuroscience*, *27*(1), 611–647. https://doi.org/10.1146/annurev.neuro.26.041002.131039

Reynolds, J. H., & Heeger, D. J. (2009). The Normalization Model of Attention. *Neuron*, *61*(2), 168–185. https://doi.org/10.1016/j.neuron.2009.01.002

Robinson, A. K., Grootswagers, T., Shatek, S. M., Gerboni, J., Holcombe, A., & Carlson, T. A. (2020). *Overlapping neural representations for the position of visible and imagined objects* (arXiv:2010.09932). arXiv. http://arxiv.org/abs/2010.09932

Sarich, D., Chappell, M., & Burgess, C. (2007). Dividing attention in the flash-lag illusion. *Vision Research*, *47*(4), 544–547. https://doi.org/10.1016/j.visres.2006.09.029

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, *79*(6), 3272–3278. https://doi.org/10.1152/jn.1998.79.6.3272

Schneider, K. A. (2018). The Flash-Lag, Fröhlich and Related Motion Illusions Are Natural Consequences of Discrete Sampling in the Visual System. *Frontiers in Psychology*, *9*, 1227. https://doi.org/10.3389/fpsyg.2018.01227

Schneider, M., Marquardt, I., Sengupta, S., Martino, F. D., & Goebel, R. (2019). *Motion Displaces Population Receptive Fields in the Direction Opposite to Motion* (p. 759183). bioRxiv. https://doi.org/10.1101/759183

Schultz, D. W., & Eriksen, C. W. (1978). Stimulus size and acuity in information processing. *Bulletin of the Psychonomic Society*, *12*(6), 397–399. https://doi.org/10.3758/BF03329719

Seidemann, E., & Newsome, W. T. (1999). Effect of Spatial Attention on the Responses of Area MT Neurons. *Journal of Neurophysiology*, *81*(4), 1783–1794. https://doi.org/10.1152/jn.1999.81.4.1783

Shioiri, S., Yamamoto, K., Oshida, H., Matsubara, K., & Yaguchi, H. (2010). Measuring attention using flash-lag effect. *Journal of Vision*, *10*(10), 10–10. https://doi.org/10.1167/10.10.10

Silvanto, J., Lavie, N., & Walsh, V. (2005). Double dissociation of V1 and V5/MT activity in visual awareness. *Cerebral Cortex (New York, N.Y.: 1991)*, *15*(11), 1736–1741. https://doi.org/10.1093/cercor/bhi050

Simanova, I., Francken, J. C., de Lange, F. P., & Bekkering, H. (2016). Linguistic priors shape categorical perception. *Language, Cognition and Neuroscience*, *31*(1), 159–165. https://doi.org/10.1080/23273798.2015.1072638

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743–761. https://doi.org/10.1016/S0042-6989(97)00183-1

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074. https://doi.org/10.1068/p281059

Smeets, J. B. J., & Brenner, E. (1995). Prediction of a moving target's position in fast goal-directed action. *Biological Cybernetics*, *73*(6), 519–528. https://doi.org/10.1007/BF00199544

Smout, C. A., Tang, M. F., Garrido, M. I., & Mattingley, J. B. (2019). Attention promotes the neural encoding of prediction errors. *PLOS Biology*, *17*(2), e2006812. https://doi.org/10.1371/journal.pbio.2006812

Souihel, S., & Cessac, B. (2021). On the potential role of lateral connectivity in retinal anticipation. *The Journal of Mathematical Neuroscience*, *11*(1), 3. https://doi.org/10.1186/s13408-020-00101-z

Spinelli, D., & Zoccolotti, P. (1992). Perception of moving and stationary gratings in brain damaged patients with unilateral spatial neglect. *Neuropsychologia*, *30*(4), 393–401. https://doi.org/10.1016/0028-3932(92)90112-Y

Spratling, M. W. (2008). Predictive coding as a model of biased competition in visual attention. *Vision Research*, *48*(12), 1391–1408. https://doi.org/10.1016/j.visres.2008.03.009

Sterzer, P., Haynes, J.-D., & Rees, G. (2006). Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT+/V5. *NeuroImage*, *32*(3), 1308–1316. https://doi.org/10.1016/j.neuroimage.2006.05.029

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, *32*(8), 1535–1549. https://doi.org/10.1016/0042-6989(92)90209-2

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. https://doi.org/10.1016/j.tics.2009.06.003

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139. https://doi.org/10.1146/annurev.ne.19.030196.000545

Teichmann, L., Edwards, G., & Baker, C. I. (2021). Resolving visual motion through perceptual gaps. *Trends in Cognitive Sciences*, *25*(11), 978–991. https://doi.org/10.1016/j.tics.2021.07.017

Teichmann, L., Moerel, D., Rich, A. N., & Baker, C. I. (2022). The nature of neural object representations during dynamic occlusion. *Cortex*, *153*, 66–86. https://doi.org/10.1016/j.cortex.2022.04.009

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, *22*(3), 377–380. https://doi.org/10.1016/0042-6989(82)90153-5

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), Article 6582. https://doi.org/10.1038/381520a0

Töllner, T., Zehetleitner, M., Gramann, K., & Müller, H. J. (2011). Stimulus Saliency Modulates Pre-Attentive Processing Speed in Human Visual Cortex. *PLOS ONE*, *6*(1), e16276. https://doi.org/10.1371/journal.pone.0016276

Treisman, A. M. (1964). SELECTIVE ATTENTION IN MAN. *British Medical Bulletin*, *20*, 12–16. https://doi.org/10.1093/oxfordjournals.bmb.a070274

Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences*, *24*(5), 295–300. https://doi.org/10.1016/S0166-2236(00)01814-2

Treue, S., & Maunsell, J. H. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, *382*(6591), 539–541. https://doi.org/10.1038/382539a0

Treue, S., & Maunsell, J. H. (1999). Effects of Attention on the Processing of Motion in Macaque Middle Temporal and Medial Superior Temporal Visual Cortical Areas. *The Journal of Neuroscience*, *19*(17), 7591–7602. https://doi.org/10.1523/JNEUROSCI.19-17-07591.1999

Turatto, M., Vescovi, M., & Valsecchi, M. (2007). Attention makes moving objects be perceived to move faster. *Vision Research*, *47*(2), 166–178. https://doi.org/10.1016/j.visres.2006.10.002

Ungerleider, L. G., & Haxby, J. V. (1994). ‘What’ and ‘where’ in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157–165. https://doi.org/10.1016/0959-4388(94)90066-3

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In *Analysis of visual behavior* (Ingle, D. J., Goodale, M. A. and Mansfield, R. J. W., eds, pp. 549–586). MIT Press. https://cir.nii.ac.jp/crid/1571698600088970624

Van Heusden, E., Harris, A. M., Garrido, M. I., & Hogendoorn, H. (2019). Predictive coding of visual motion in both monocular and binocular human visual processing. *Journal of Vision*, *19*(1), 3. https://doi.org/10.1167/19.1.3

Vetter, P., Grosbras, M.-H., & Muckli, L. (2015). TMS Over V5 Disrupts Motion Prediction. *Cerebral Cortex*, *25*(4), 1052–1059. https://doi.org/10.1093/cercor/bht297

Vreven, D., & Verghese, P. (2005). Predictability and the Dynamics of Position Processing in the Flash-Lag Effect. *Perception*, *34*(1), 31–44. https://doi.org/10.1068/p5371

Walsh, K. S., McGovern, D. P., Clark, A., & O’Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. https://doi.org/10.1111/nyas.14321

Wang, W., Lei, X., Gong, W., Liang, K., & Chen, L. (2022). Facilitation and inhibition effects of anodal and cathodal tDCS over areas MT+ on the flash-lag effect. *Journal of Neurophysiology*, *128*(1), 239–248. https://doi.org/10.1152/jn.00091.2022

Wässle, H. (2004). Parallel processing in the mammalian retina. *Nature Reviews Neuroscience*, *5*(10), Article 10. https://doi.org/10.1038/nrn1497

Wertheimer, M. (1912). *Experimentelle Studien über das Sehen von Bewegung*. J.A. Barth.

Wexler, M., Glennerster, A., Cavanagh, P., Ito, H., & Seno, T. (2013). Default perception of high-speed motion. *Proceedings of the National Academy of Sciences*, *110*(17), 7080–7085. https://doi.org/10.1073/pnas.1213997110

White, P. A. (2018). Is the perceived present a predictive model of the objective present? *Visual Cognition*, *26*(8), 624–654. https://doi.org/10.1080/13506285.2018.1530322

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience*, *1*(8), 656–657. https://doi.org/10.1038/3659

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research*, *40*(2), 137–149. https://doi.org/10.1016/S0042-6989(99)00166-2

Wojtach, W. T., Sung, K., Truong, S., & Purves, D. (2008). An empirical explanation of the flash-lag effect. *Proceedings of the National Academy of Sciences*, *105*(42), 16338–16343. https://doi.org/10.1073/pnas.0808916105

Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nature Neuroscience*, *9*(9), 1156–1160. https://doi.org/10.1038/nn1748

Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive Field Shift and Shrinkage in Macaque Middle Temporal Area through Attentional Gain Modulation. *The Journal of Neuroscience*, *28*(36), 8934–8944. https://doi.org/10.1523/JNEUROSCI.4030-07.2008

Wyatte, D., Jilk, D. J., & O'Reilly, R. C. (2014). Early recurrent feedback facilitates visual object recognition under challenging conditions. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00674

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience*, *1*(6), 508–512. https://doi.org/10.1038/2226

Yoo, S.-A., Martinez-Trujillo, J. C., Treue, S., Tsotsos, J. K., & Fallah, M. (2022). Attention to visual motion suppresses neuronal and behavioral sensitivity in nearby feature space. *BMC Biology*, *20*(1), 220. https://doi.org/10.1186/s12915-022-01428-7

Zaksas, D., & Pasternak, T. (2006). Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(45), 11726–11742. https://doi.org/10.1523/JNEUROSCI.3420-06.2006

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *The Journal of Physiology*, *236*(3), 549–573. https://doi.org/10.1113/jphysiol.1974.sp010452

Zeki, S. M., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *11*(3), 641–649. https://doi.org/10.1523/JNEUROSCI.11-03-00641.1991

Author Contributions

Chapter 2

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2022). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. *Vision Research.*

Yook, J.	Conceptualization, Methodology, Investigation, Formal analysis, Writing (original draft), Writing (review & editing), Visualization
Lee, L.	Conceptualization, Methodology, Investigation, Writing (review & editing), Visualization
Vossel, S.	Supervision, Writing (review & editing)
Weidner, R.	Supervision, Writing (review & editing)
Hogendoorn, H.	Conceptualization, Methodology, Supervision, Writing (review & editing), Funding acquisition

Chapter 3

Yook, J., Hogendoorn, H., Fink, G.R., Vossel, S., & Weidner, R. (in press). When visual attention is divided in the flash-lag effect. *Journal of Vision.*

Yook, J.	Conceptualization, Methodology, Investigation, Formal analysis, Writing (original draft), Writing (review & editing), Visualization
Hogendoorn, H.	Conceptualization, Supervision, Writing (review & editing)
Fink, G.R.	Writing (review & editing), Funding acquisition
Vossel, S.*	Conceptualization, Methodology, Supervision, Writing (review & editing)
Weidner, R.*	Conceptualization, Methodology, Supervision, Writing (review & editing)

*These authors contributed equally

Chapter 4

Yook, J., Viswanathan S., Turner, W., Johnson, P., Weidner, R., Vossel, S., & Hogendoorn, H. (in prep). Meaningful objects: Selective visual attention guides predictive motion extrapolation.

Yook, J.	Conceptualization, Methodology, Investigation, Formal analysis, Writing (original draft), Writing (review & editing), Visualization
Viswanathan, S.	Formal analysis, Writing (original draft), Writing (review & editing)
Turner, W.	Investigation, Formal analysis, Writing (review & editing)
Johnson, P.	Writing (review & editing)
Weidner, R.	Conceptualization, Supervision, Writing (review & editing)
Vossel, S.	Conceptualization, Supervision, Writing (review & editing)
Hogendoorn, H.	Conceptualization, Methodology, Supervision, Writing (review & editing), Funding acquisition

Curriculum Vitae

Scientific Education

Doctor of Philosophy, Psychology

University of Melbourne, University of Cologne 12/2019 – present

- Joint-PhD degree, Jülich-University of Melbourne Postgraduate Academy (JUMPA) program
- Melbourne Neuroscience PhD program at The Florey Institute of Neuroscience and Mental Health

Honours, Psychology (1st Class)

Australian National University

Bachelor of Arts, Psychology

University of Melbourne

US High School Diploma

Dalat International School

Relevant Experience

Doctoral Researcher

Forschungszentrum Jülich Jülich (GER) 2021 – 2024

- Managed experimental studies at the Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM-3)

Research Assistant

University of Melbourne Melbourne (AUS) 2019

- Coordinated the planning and initiation of electrophysiological studies at the Melbourne School of Psychological Sciences

Science Communication Intern

Biotext Canberra (AUS) 2018

- Supported the development of project timelines and deliverables at a City science communications firm

Psychology Intern

Women's Centre for Change Penang (MYS) 2017

- Provided support to the counseling and legal team, assisting with crisis intervention documentation and case management for individuals affected by child sexual abuse and domestic violence at a non-profit human rights organization

Skills

Transferrable skills

Ethical Awareness, Organization, Regulatory Compliance, Data Visualization

Technical skills

MATLAB, Psychtoolbox, EEGLAB, Python, Psychopy, R, Microsoft Office, Technical Writing, Project Management

Management skills

Leadership, Communication, Adaptability, Empathy

Languages

English, Korean, Mandarin, Bahasa Melayu, German

Additional Experience

Doctoral Researchers Association (DocTeam) Spokesperson

Forschungszentrum Jülich 2021 – 2022

- Facilitated communication between doctoral researchers and leaders of the Forschungszentrum including Board of Directors, administration, and the Helmholtz Association to enhance working conditions of the research center
- Spearheaded a "Wellness Series" campaign for Mental Health Month in May 2022

All-Staff Choir

IgniteLAB Melbourne Conservatorium of Music, University of Melbourne 2019

Complex Human Data Summer School

Melbourne School of Psychological Sciences, University of Melbourne 2018

ARC Centre of Excellence in Cognition and its Disorders

Australian Research Council 2018

Journal Publications

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2022). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. *Vision Research.*

Feuerriegel, D., **Yook, J**., Quek, G. L., Hogendoorn, H., & Bode, S. (2020). Visual mismatch responses index surprise signalling but not expectation suppression. *Cortex*.

Published Abstracts

Yook, J., Turner, W., Johnson, P., Weidner, R., Vossel, S., & Hogendoorn, H. (2023). Prediction in the visual perception of moving objects: the role of attention. Invited symposium talk given at the 65th *Tagung experimentell arbeitender Psycholog:innen; Conference of Experimental Psychologists (TeaP)*, Trier, Germany.

Yook, J., Turner, W., Johnson, P., Weidner, R., Vossel, S., & Hogendoorn, H. (2022). Predictive activation of neural position representations for moving objects with and without visual attention. Poster presented at *European Conference of Visual Perception (ECVP),* Nijmegen, The Netherlands.

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2021). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Poster presented at the *INM-3 Retreat,* Jülich, Germany.

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2020). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Talk given at *Neuromatch Conference 3.0* (online).

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2020). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Talk given at the *MDHS Graduate Research Conference,* Melbourne, Australia (online).

Yook, J., Lee, L., Vossel, S., Weidner, R., & Hogendoorn, H. (2020). Motion extrapolation in the flash-lag effect depends on perceived, rather than physical speed. Poster presented at the *UNSW workshop on Expectation, Perception, and Cognition*, Sydney, Australia (online).

Feuerriegel, D., **Yook, J.**, Quek, G. L., Hogendoorn, H., & Bode, S. (2019). Visual mismatch responses are a product of surprise, but not fulfilled expectations. Talk given at *Australian Cognitive Neuroscience Society Conference (ACNS)*, Launceston, Australia.

Dawel, A., Rippon, K., **Yook, J.**, Crookes, K., Foo, Y. Z., & Rhodes, G. (2019). Differences in EEG responses to human vs computer-generated (FaceGen) faces revealed by Fast Periodic Visual Stimulation (FPVS). Poster presented at *BIOPAC T4 Human Physiology Conference*, Santa Barbara, USA.

Dawel, A., Dietrich, E. J., Easton, B., Rippon, K., **Yook, J.**, Crookes, K., Foo, Y. Z., & Rhodes, G. (2018). When computer-generated (CG) faces behave in unreal ways we get unreal answers: The value of seeking convergent evidence from naturalistic facial behaviour. Talk given at *Emerging Research in Cognition and Emotion Forum, Australasian Experimental Psychology Conference*, Wellington, New Zealand.

23.02.2024, Jeong In (Jane) Yook

Lebenslauf

Jeong In (Jane) Yook was born on the 3rd of January 1996 in Ulsan, South Korea. After graduating from high school at Dalat International School located at the tropical island of Penang, Malaysia in 2013, she went on to study at the renowned University of Melbourne, Australia. Initially drawn to a career in Forensic Psychology, she started out in Psychology and minoring in Criminology. She decided to switch her minor to Media and Communication, which turned out to be a switch for the better. In 2016, she obtained her Bachelor's degree in Psychology, with a newfound fascination for cognitive neuroscience, namely, investigating the link between brain and behavior, sparked by courses such as The Integrated Brain and The Unconscious Mind. In 2018, she obtained her Honours' degree in Psychology at the Australian National University. Particularly, the Honours thesis studies conducted with Amy Dawel evolved the aforementioned spark into a fervent interest. In 2019, Jane moved back to Melbourne to contribute to research endeavors with Stefan Bode and Daniel Feuerriegel at the Decision Neuroscience Lab. The culmination of these experiences led her to start a PhD within the JUMPA program in December of the same year, under the supervision of Hinze Hogendoorn in the Timing Lab. In 2021, she relocated to Cologne, Germany to join her supervisors Ralph Weidner and Simone Vossel at the Spatial Cognition and Prediction and Attention groups within the Institute of Neuroscience and Medicine, Forschungszentrum Jülich, where she finished her PhD thesis.