Improving Spatiotopic Models of Vision using Retinotopic Input and Vector-based Saccadic Generation

A theoretical and methodological framework for the study of attentional control settings

Sofia Krasovskaya

Thesis for the degree of Philosophiae Doctor

June 2023

School of Health Sciences

FACULTY OF PSYCHOLOGY

Improving Spatiotopic Models of Vision using Retinotopic Input and Vector-based Saccadic Generation

A theoretical and methodological framework for the study of attentional control settings

Sofia Krasovskaya

Thesis for the degree of Philosophiae Doctor

Supervisors
Árni Kristjánsson, W. Joseph MacInnes

Doctoral committee
Árni Kristjansson, PhD, University of Iceland

W. Joseph MacInnes, PhD, Swansea University

Árni Gunnar Ásgeirsson, PhD, University of Akureyri

June 2023

School of Health Sciences
FACULTY OF PSYCHOLOGY
UNIVERSITY OF ICELAND

Abstract

Our surroundings are seldom stable and are filled with various information. Visual processing is one of the ways the brain deals with all this information; however, the capacity of our visual system is limited. To deal with this limitation, we have attentional mechanisms that help us extract relevant information from noisy surroundings. Within the visual modality, we can study these attentional mechanisms through the measurement of eye movements and computational approaches. This dissertation proposes a theoretical framework for the potential combination of eye movement data and computational approaches to construct a biologically inspired computational model of visual attention that would accommodate the attentional control settings under various context and task conditions.

Machine learning has gained popularity in the recent years. Many sophisticated algorithms and approaches fuel technological developments, such as face and object recognition for robotic applications like self-driving cars or security software. While these areas of research and technological application are undoubtedly important and interesting, there remain other domains that could benefit from such approaches, such as vision modelling. This dissertation aims to address existing gaps in the modelling of visual attention within the domain of cognitive neuroscience vision research. Specifically, the aims are: (1) to analyse the state-of-the-art in the field of visual attention modelling with respect to computational cognitive neuroscience; (2) to gain a deeper insight into some of the cognitive and biological mechanisms in visual attention; and (3) to introduce ways to improve existing computational models of vision using the results of the studies.

The aims are addressed in three studies. Study I focuses on computational modelling and consists of one major and two auxiliary manuscripts. The former provides an in-depth review of the field and proposes a novel theoretical guideline for computational cognitive neuroscience models of visual salience. The latter two papers investigate some computational approaches to create a biologically accurate model of visual attention, as well as their limitations. In Study II attentional control settings are observed within a gaze-contingent experimental setup, where the size of observers' functional field of view is directly manipulated. Performance (response times and accuracy rates) is assessed across several task conditions, revealing that the size of the functional field of view is dependent on more factors than initially thought. Study III proposes a novel experimental design to studying executive control and oculomotor suppression via microsaccade rates in the antisaccade paradigm. The results of this study demonstrate that microsaccade rates can be used as a measure of the degree of suppression of the oculomotor system.

The theoretical and experimental findings of the studies included in this dissertation are important for the field of cognitive computational modelling of visual attention and perception. The insights gained from them can improve our understanding of how visual attention operates within different contexts, while the use of computational modelling makes it possible to observe the interactions of the attentional mechanisms that guide our perception of the surroundings.

Keywords:

Visual attention, Attentional control settings, computational cognitive neuroscience, models of vision, executive control

Acknowledgements

This has been a very long and difficult journey, and I would never have accomplished it alone. The list of people to whom I am grateful is extremely long, and I apologise to those I do not mention here separately.

I want to express my deepest thanks to my supervisor, Dr Joe MacInnes, who has supported me through all my struggles – both professional and personal. His mentorship and friendship have been invaluable to me, while the VML lair he created at HSE University provided a source of income and endless aromatic coffee. To my supervisor, Dr Árni Kristjánsson, for his valuable insights and constructive feedback, as well as for providing safety and making the journey towards finishing my doctorate possible. To Dr Sabrina Hansmann-Roth, for providing professional advice and being a friend. To Dr Heiða María Sigurðardóttir, for sharing her academic and personal experience, and checking in on me. To Dr Hélène Devillez, for the proofreading, support, and the steed that has saved me hours of commuting and provided the excitement of riding with the wind.

I would like to acknowledge the support, assistance and smiles provided by the staff at the University of Iceland, the faculty of psychology and the IVL, for adopting and funding me and making me feel accepted. Their resources, facilities, as well as technical and mental support have made it possible to come to Iceland and finish my studies.

I would like to separately extend my appreciation to the participants who found the time to actively participate in my studies.

I want to mention my close friends. Henry-san, for his support and friendship over the decades, as well as our long promenades and comfortable silences. Alicia, for sharing the best and the worst, for the pies, for the coffee black as midnight on a moonless night, for Lynchian conversations and for Neuroscience. Úlla, for Íslenska, the care, the art, and the pain therapy. DD, for sharing her otherworldly beauty, inspiration, and gelatos with me. My coach, Greg, for the patience and the consistent training plans adapted to the often-shifting circumstances.

My endless gratitude, of course, goes to my family. To my mother Lana, for instilling curiosity in me from early childhood, teaching me, discussing books and the nature of things, and highlighting the importance of education. To my brother Ian, and Catherine, my sister-in-law, for believing in me and supporting me despite all odds. To my *altera particula*, Daniel, for sticking with me through all the plagues, wars, the software, hardware, and emotion-ware breakdowns, as well as distance and time

zone differences, among the other struggles. Your amaranthine patience, selfless care and COL-LA-BO-RAY-SHON have made this doctorate possible.

I would also like to mention Cla-Val Europe for helping me financially to quickly escape the political and military situation amidst the chaos and relocate.

Finally, I want to acknowledge several important sources of mental stability during my doctoral studies: the furry pawed creatures Dita von Tess, Baby Sandy Peso, Princess Millie and Druid Steris, Toulouse and Moxy, Móri and Kisa (the perfectly shaped one), Kisa (the little void of joy), and Harley, the goodest wolf-boy; the University gym, Vesturbæjarlaug, Laugardalslaug, Nauthólsvík and Ármann powerlifting club; and, of course, Sertraline, Venlafaxine, Quetiapine and Oxcarbazepine.

Contents

List of Abbreviations

2AFC – 2-alternative forced choice task

2D – two-dimensional

ANOVA – analysis of variance

CNN – convolutional neural network

dlPFC – dorsolateral prefrontal cortex

FEF – frontal eye fields

FFV – functional field of view

FIT – feature integration theory

GAN – generative adversarial network

GLME – generalised linear mixed effects

IOR – inhibition of return

IT – inferior temporal gyrus

LGN – lateral geniculate nucleus

LIF – leaky integrate-and-fire

LIP – lateral intraparietal area

LSTM – long short-term memory

PFC – prefrontal cortex

PPC – posterior parietal cortex

RNN – recurrent neural network

SC – superior colliculus

SEF – supplementary eye fields

SLCA – spatial leaky competing accumulator

vlPFC – ventrolateral prefrontal cortex

WTA – winner-take-all

List of Figures

List a short title of all figures in the thesis.

List of Tables

List of Original Papers

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I, Ia, Ib, II, III):

I. Krasovskaya, S., & MacInnes, W. J. (2019). Salience Models: A Computational Cognitive Neuroscience Review. *Vision, 3*(4), Article 4. https://doi.org/10.3390/vision3040056

Ia. Krasovskaya, S., Zhulikov, G., & MacInnes, J. (2021). Deep Learning Neural Networks as a Component of a Model of Saccadic Generation. *PsyArXiv.* https://doi.org/10.31234/osf.io/zr6ma

Ib. Merzon, L., Malevich, T., Zhulikov, G., Krasovskaya, S., & MacInnes, W. J. (2020). Temporal Limitations of the Standard Leaky Integrate and Fire Model. *Brain Sciences, 10*(1), Article 1. https://doi.org/10.3390/brainsci10010016

II. Krasovskaya, S., Kristjánsson, Á., & MacInnes, W. J. (manuscript submitted). Assessing the Optimal Size of the Functional Field of View by Task in a Gaze-Contingent Search Paradigm.

III. Krasovskaya, S., Kristjánsson, Á., & MacInnes, W. J. (2023). Microsaccade rate activity during the preparation of pro- and antisaccades. *Attention, Perception, & Psychophysics.* https://doi.org/10.3758/s13414-023-02731-3

All papers are reprinted by kind permission of the publishers.

Declaration of Contribution

I, Krasovskaya Sofia, declare that this dissertation represents my original work and contribution to the field of visual attention. I affirm that the following contributions have been made by me:

Research Design: I conceptualized and designed the research studies under the guidance of my supervisors and coauthors, Drs Joseph MacInnes and Árni Kristjánsson, including the formulation of research questions, the selection of appropriate methodologies, and the development of a comprehensive research framework.

Data Collection: I conducted data collection for this study, which involved experiments. I managed the logistics of data collection, recruited participants, and obtained informed consent.

Data Analysis: I performed detailed data analysis using appropriate qualitative or quantitative methods. I carefully examined the collected data, identified patterns, themes, or trends, and drew meaningful conclusions based on rigorous analysis. Supervision and control of the analyses were performed by my supervisors, Drs Joseph MacInnes and Árni Kristjánsson.

Literature Review: I conducted an extensive review of the relevant literature to establish the theoretical foundation of the dissertation. I critically analysed existing research, identified research gaps, and synthesized knowledge to situate my study within the broader academic discourse.

Manuscript Writing: I authored the dissertation manuscript, including the introduction, literature review, methodology, results, discussion, and conclusion chapters. I ensured clarity, coherence, and logical progression of ideas, while adhering to the specified formatting guidelines and academic standards.

Contribution to Knowledge: Through this dissertation, I have made a significant contribution to the field by advancing theoretical understanding, addressing research gaps, and proposing practical and methodological implications.

Ethical Considerations: I adhered to ethical principles and guidelines throughout the research process. I obtained necessary ethical approvals, ensured participant confidentiality, and conducted the study with integrity and professionalism.

I would like to express my sincere gratitude to my dissertation committee, advisors, colleagues, and participants for their support, guidance, and valuable input during the process.

Sofia Krasovskaya
July 2nd, 2023

Structure of the dissertation

The dissertation starts with a general introduction to the field of visual attention that sets the context of the research project, states its objectives, its significance, and the contribution to the field. After this we plunge deeper into the brief history of eye movement research, types of eye movements, their connection to attention, attentional control settings and higher-level cognitive functions, and an overview of eyetracking techniques. The neurophysiological mechanisms of eye movements are provided together with the ideas of a saliency and a priority map and their role in visual attention. With the understanding of these concepts, the reader is further submerged into Study I that focuses on the approaches to computational modelling of saliency and how cognitive neuroscience approaches differ from purely computational approaches. Two smaller (but crucial to Study I) computational modelling projects are briefly introduced. At the end of the chapter a specific issue of spatiotopic models is raised. Study II provides a brief introduction into retinotopy, gaze-contingent setups and the idea of a functional field of view (FFV). It then describes the methods and the results of three experiments followed by a discussion of the study. Study III continues the study of attentional settings within the antisaccade task. Once the reader is introduced to the background and the questions of the study, the chapter flows on into the description of the methods and the results of two eyetracking experiments. This is followed by a discussion of the results of the study and their implications. Finally, the three major studies converge in the General Discussion, where their implications and connection with the mechanisms of selective visual attention are considered. The dissertation ends with the original publications, included after the References in the *Original Publications* section and *Appendix A.*

Significance and Contributions

The dissertation proposes a methodological and theoretical framework for studying selective attention and attentional control settings using eye movements and computational modelling approaches.

Taken separately, each of the studies makes its own contribution on the verge of several disciplines. Study I takes a unique approach towards the assessment criteria of models from the standpoint of computational cognitive neuroscience. It proposes a classification system and analyses existing prominent models of visual saliency based on these criteria. The auxiliary projects add practical value to this study by implementing and testing newer deep learning computational approaches for spatial predictions, as well as optimising the temporal parameters of the Leaky integrate-and-fire algorithm for saccade reaction time simulations. These auxiliary studies helped identify the issues with the attempted methods and come up with the proposed guidelines. The separate projects are well cited, especially the review on saliency models (Krasovskaya & MacInnes, 2019).

Study II incorporates a sophisticated gaze-contingent setup to study visual attention in a visual search task. These types of tasks reveal how attention operates in visual perception. Visual search experiments examine how attention is allocated and directed towards visual stimuli, which can reveal the underlying mechanisms and processes involved in the visual perception of complex visual scenes. The gaze-contingent setup allows us to study the phenomenon in a retinotopic - native for the visual system - frame of reference. These data are highly valuable for updating the spatiotopic components of computational models with retinotopic information. Additionally, the identification of optimal FFV sizes can inform the design of visual displays and user interfaces, such as those used in aviation, automotive, and medical industries. By understanding how humans search for and identify visual targets, and how they can deal with certain amounts and layouts of visual information researchers can enable the design of more efficient displays.

The results of Study III imply a potential connection of microsaccade control with voluntary oculomotor suppression within an antisaccade setup (Hallett, 1978), suggesting a possible top-down control mechanism mediated by the frontal cortex. This observed trend may set a new direction of research of microsaccadic eye movements to further investigate their role(s) and perhaps provide a standardised framework for studying these micromovements. Visual search experiments have been used to study attentional disorders, such as attention deficit hyperactivity disorder (ADHD) and neglect. Experiments involving executive and top-down control,

such as the antisaccade task, can provide insights into the underlying cognitive processes and mechanisms involved in these disorders (Jóhannesson et al., 2013). Microsaccade rate data, combined with performance on the antisaccade task, can improve our understanding and diagnostics of attentional disorders.

Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth.

\- Jules Verne, *Journey to the Centre of the Earth*

Introduction

Eyes as a gateway into attention

The eyes: such a natural thing to see them bare in any culture and at any point in history. However, when you think of it, is it not an amazing and simultaneously terrifying realisation that an organ that evolved as an extension of the diencephalon should be so exposed? Due to this evolutionary feat, the eyes have attracted the attention not only of potential mates, but also of philosophers and researchers.

The history of oculomotor research goes way back to antiquity when Aristotle observed that human vision is binocular (as described in (N. J. Wade, 2010). Since then, we have learned a tremendous amount of fascinating things about the eyes and their relationship with the brain and the operation of the visual system in general. One of the most important discoveries, however, was that through the observation and estimation of oculomotor behaviour, we have access to a variety of cognitive processes without the need to use invasive neuroimaging techniques. In other words, eye movements are a rather valid proxy for attention (Kristjánsson, 2011).

To get to the part where sophisticated practical research methods appeared, it was necessary to go through an extended period of focussing on the theoretical aspects of vision. For example, Leonardo da Vinci studied the way we perceive the world around us and realised that the perception of the visual scene is not homogenous, but in fact we can only see a small part of it acutely – the one located directly within our 'line of sight' (Hunziker, 2006). Such observations led to more questions related to the way we can process the entire visual scene despite its large variations in acuity, which, in turn, led to the rise of experimental research of eye movements in the 18th century (for a more detailed history see (N. J. Wade & Tatler, 2011).

The beginning of the experimental period of eye movement research goes back as far as the late 18th century with Wells' study of optics (N. Wade et al., 2011). Almost a century later, researchers like Hering, Helmholtz and Dodge were looking into oculomotor behaviour during the execution of reading and visual perception tasks (N. J. Wade & Tatler, 2011). They noticed that during these tasks the eyes produced different patterns. In 1903, Dodge and his wife used photographic recording to document eye movements during the execution of reading tasks. This approach allowed them to distinguish the types and properties of eye movements.

Classification of eye movements

The Dodges distinguished five types of eye movements classified based on their functions into two general categories: movements that shift gaze and movements that stabilise gaze (Cullen & Van Horn, 2011). Sometimes the category of fixational eye movements is added to Dodge's traditional classification. Eye movements belonging to this category are considered to prevent neural adaptation that leads to the inability to perceive a visual image during a static eye fixation (Ditchburn & Ginsborg, 1952); (Martinez-Conde et al., 2006); (Macknik & Martinez-Conde, 2015); see also (Kowler, 2011), for an in-depth review). Thus, eye movements may be classified into:

1.   Gaze-shifting eye movements:
     a) *Saccadic eye movements (fig. 1a).* Rapid, ballistic shifts of gaze with velocities of up to 900°/sec.
     b) *Smooth pursuit (fig. 1c).* Slow and smooth eye movements used to track an external stimulus, such as a ball during a game of squash or a motorcycle moving on the racetrack. These movements are under voluntary control but cannot be executed without fixating on a moving target.
     c) *Vergence (fig. 1d).* These movements allow us to perceive objects at various distance without any blurring of the perceived image. They include converging and diverging movements of the eyes in order to place the object of focus in our line of sight depending on its proximity. In short, vergence movements make binocular vision possible.

2.   Gaze-stabilising eye movements:
     a) *Vestibulo-ocular movements (fig. 1f).* These movements serve to compensate for movements of the head, allowing us to maintain fixation on one point without any signal disruptions. They are not under our voluntary control and are reflexive.
     b) *Optokinetic movements (fig. 1e).* Another type of reflexive movement that compensates for the movement of objects relative to the head. Optokinetic movements may be divided into two distinct phases: a smooth pursuit movement followed by a quick shift of gaze in the opposite direction, often called nystagmus. One might experience optokinetic nystagmus when standing on a train platform in front of a passing train or when moving along a fence with a regular vertical pattern.

3.   Fixational eye movements (fig. 1b):
     a) *Drift.* Slow, barely noticeable eye movements that have a 'bending' pattern.
     b) *Tremor.* Tiny perturbations of the eye that coincide with drifts.
     c) *Microsaccades.* The most debatable micromovements in the field. They are very short (under 1°) saccades.

Figure 1. Types of Eye Movements
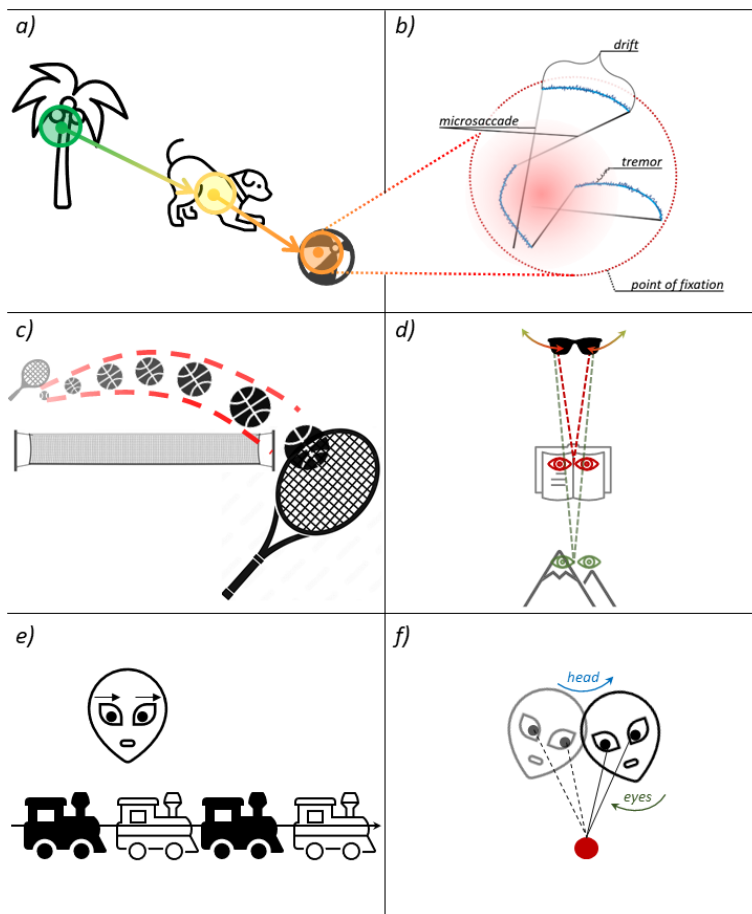
Note. A schematic representation of saccades and fixations (panel a); fixational eye movements (panel b); smooth pursuit movements (panel c); vergence (panel d); optokinetic movements (panel e); and vestibulo-ocular movements (panel f).

'Jerks' and their characteristics

Saccades are the most frequent eye movements humans make, with an average frequency of three saccades per second. Introduced by Javal (1879), observed in detail by Brown (1878) and finally adopted into the English language by Dodge (1916), saccades may be described as ballistic 'jerks' of the eye that shift one's gaze towards regions of interest in the visual field (Wade et al., 2003). With peak velocities as high as 900 degrees per second, saccades are initiated by the firing of neurons in the superior colliculus (SC) at rates of up to 450 spikes per second (Opstal & Kasap, 2019). Between each saccade the eyes briefly fixate (~250 ms) on some point in visual space. Such saccade and fixation combinations are referred to as saccade sequences.

Decades of research have led us to a more-or-less clear understanding of

saccadic eye movements and their underlying neurophysiological and neuroanatomical mechanisms. From the point of neurophysiological circuits, extensive research has been carried out, from single cell recordings in primates (Killian et al., 2015; Schlag-Rey & Schlag, 19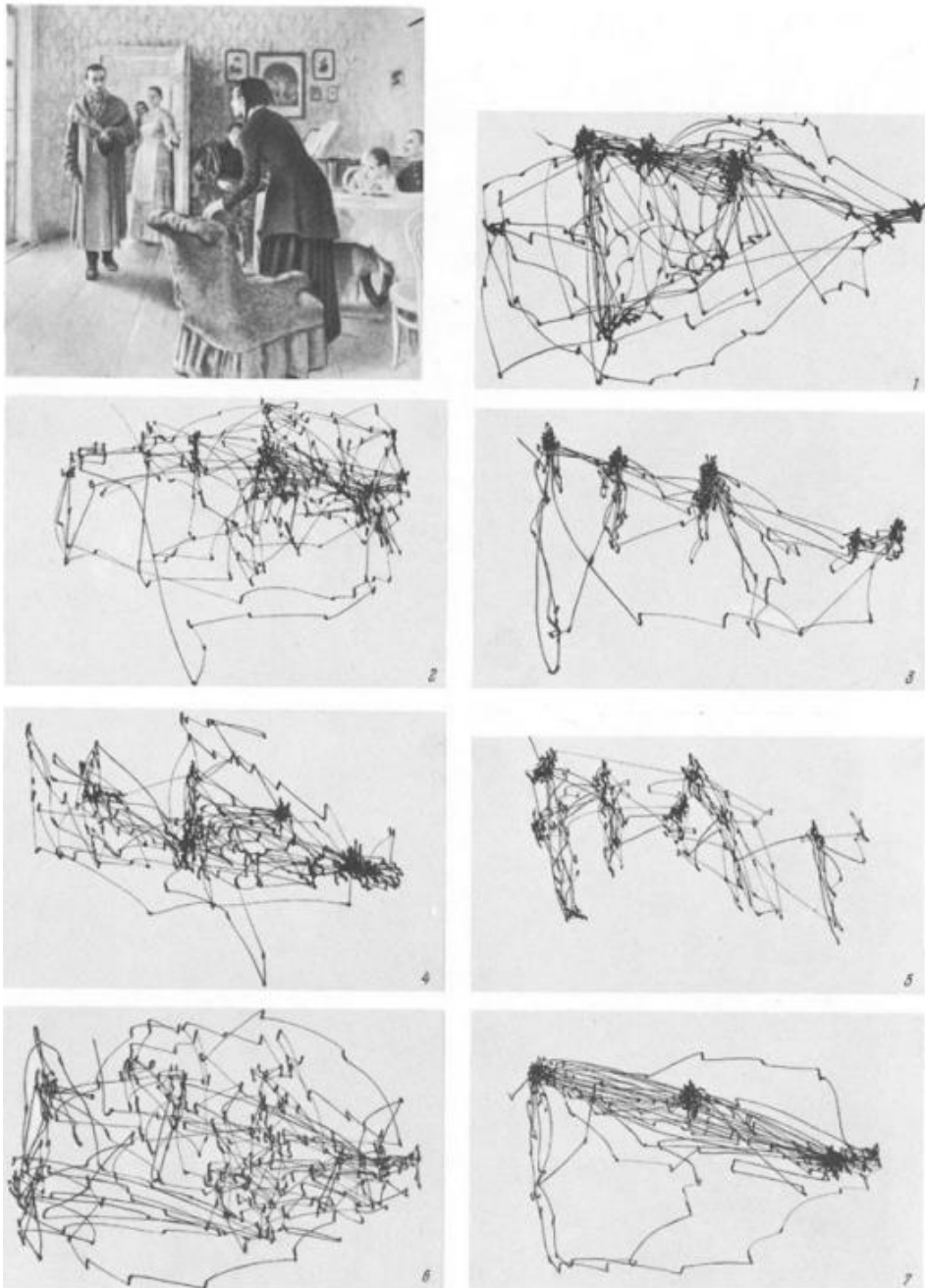84) to fMRI (Kimmig et al., 2001) to brain lesion studies (Pierrot-Deseilligny et al., 2004; also see Munoz & Everling, 2004, for a review). This corpus of knowledge makes it possible to study a variety of cognitive, neural, and psychiatric mechanisms and conditions based on oculomotor behaviour alone. One point of relevance is that saccades are not only reflexive, as when a person turns their gaze toward a source of a loud unexpected noise without even thinking about it, but also under volitional control. Such volitional control has been demonstrated during various tasks, one of which is the so-called antisaccade task (Hallett, 1978), where observers are asked to inhibit the reflexive saccade towards a stimulus and produce a saccade in the opposite direction. Thus, the process may be divided into two major steps: (1) the volitional selective suppression of the reflexive saccade and (2) the inversion of the saccadic vector.

Tracking

The emergence of eyetracking technology on the cusp of the 19th – 20th cc. marked a pivotal moment in the field of experimental research on attention. It introduced a new era of investigation, allowing researchers to delve into the intricate workings of the human visual system. From stethoscope-like devices connected to the eyelid (Lamare, 1892; Hering, 1879; Björk & Kugelberg, 1953; see historical review in Wade et al., 2003; Płużyczka, 2018) and suctioning eye cups with attached levers (Delabarre, 1898; Richardson & Spivey, 2008; also see van Gompel, 2007), eyetracking technology became more and more sophisticated. The most prominent idea that later became the foundation for modern P-CR eyetracking techniques (Merchant et al., 1974; Kimmel et al., 2012) was the measurement of light reflected from the eye (Vernon, 1928; Dodge & Cline, 1901)

With the possibility of making more precise measurements, researchers could move from observations during reading to more complex tasks. Buswell was the first researcher to observe how cognitive factors influence eye movements when he compared how patterns of visual behaviour changed during the viewing of the same images but when instructed with a specific task and when no task was given (Buswell, 1935). Yarbus further examined the role of cognitive factors in oculomotor behaviour by observing how patterns of eye movement changed within the same image but with different task instruction (Yarbus, 1967) (Fig.2). He also conducted a study related to face perception, where he noticed that ocular fixations happened primarily in a triangular pattern focusing on the eyes and mouth of a portrayed image of a face (Fig. 3). In the 1970's, Noton & Stark (1971) studied the role of eye movements in perception and attention.

Figure 2. Yarbus' Classical Study on Task-Related Eye Movements

Figure 3. Yarbus' Study of Eye Movements during Face Perception

These investigations of eye movements in 20th century paved the way for deeper research of the cognitive processes involved in visual perception.

Selective attention as a controlling mechanism for fixation locations

Perhaps one of the most important cognitive mechanisms in visual perception is selective attention. The choice to look at a specific spatial location is seldom random. Generally speaking, selective attention operates as a filter that determines which stimuli to process and which to ignore. To do this effectively, it employs so-called attentional control mechanisms that guide cognitive settings and biases to direct attention within a specific context, prioritising some kinds of information over others (Posner, 1988; Vecera et al., 2014; Kaye & Ruskin, 1990). This dynamic context-dependent framework allows us to flexibly distribute attentional resources based on a variety of factors. Perhaps the most basic factor that drives attention is saliency. Areas of the visual scene that attract our gaze more than others are usually called salient. The idea of salient locations is typically based on the notion of bottom-up attention which guides our selection of a fixation target based solely on the visual properties of a visual scene. This suggests that a saccade would be directed towards a salient location in space without consideration for any task being carried out at that time (Ludwig & Gilchrist, 2002), like in Yarbus' experiment during which participants tended to focus on the most salient and informative parts of the face – the eyes and

mouth. Bottom-up saliency typically includes such low-level visual properties as size, orientation, colour, motion, etc. (Itti & Koch, 2000; Theeuwes, 2004). Thus, bottom-up saliency is closely related to feature maps of the feature integration theory of selective attention (FIT) (Treisman & Gelade, 1980). According to FIT, we process object features separately in parallel feature maps before they are bound, or integrated, into a representation of an object under the guidance of a master attention map.

However, attending solely to physically salient stimuli without being able to ignore non-relevant locations would probably be very draining, as it is very demanding, if even possible, to constantly process everything in our visual field simultaneously. As Yarbus (1967) showed when he demonstrated *The Unexpected Visitor* seven times to one participant, in addition to bottom-up attentional processes there exist relevant top-down factors that guide our attention to specific locations (see figure 2). This has been verified many times ever since with various factors, from task requirements to personal preferences and even mental states (Johansson et al., 2001; Awh et al., 2012; Hershler & Hochstein, 2009; Corbetta & Shulman, 2002).

From the point of view of neurophysiology, bottom-up visual information processes may be generally described as follows: visual information from the retina is first sent via the lateral geniculate nucleus (LGN) towards V1. From here it reaches extrastriate areas through two major pathways, the dorsal and the ventral streams (Desimone & Duncan, 1995; Saalmann et al., 2007). Visual information is able to reach the dorsolateral prefrontal cortex (dlPFC) via areas that are in the dorsal pathway: areas V1, V2, V3, MT, MST, and the lateral intraparietal area (LIP) of the posterior parietal cortex (Kravitz et al., 2011; Rauschecker, 2018). This pathway is mainly responsible for the processing of spatial location information and motion detection. The ventral stream includes areas V1, V2, V4, the inferior temporal cortex (IT) and the ventrolateral prefrontal cortex (vlPFC) and is associated with object recognition (Hamker, 2002; Kravitz et al., 2013; Mishkin et al., 1983; Pietrini et al., 2004)

While bottom-up visual processing seems relatively straightforward, the mechanisms of top-down attention are more complicated. Neuronal activation happens in the dorsal and ventral streams too; however, top-down processes rely on reciprocal connections between areas in the PPC and the PFC (Buschman & Miller, 2007; Katsuki & Constantinidis, 2014; Zanto et al., 2011). Their participation in this process is not surprising, as top-down attention suggests the use of working memory and decision-making, both of which require the participation of higher order cortical areas. The control of attention necessary for determining the relevance or irrelevance of an object to the task is associated with neuronal activation and inhibition to this object in fronto-parietal areas (Buschman & Miller, 2010; Corbetta et al., 1993, 1998; T. Donner et al., 2000; Suzuki & Gottlieb, 2013). On top of that, there is enough

evidence suggesting the participation of the fronto-parietal network during the execution of visual search tasks (Corbetta et al., 1995; Donner et al., 2000; Donner et al., 2002; Fischer et al., 2021; Ischebeck et al., 2021; Nobre et al., 2002) and its role in decision-making related to fixation locations (T. Liu et al., 2003; Schafer & Moore, 2011), with the frontal eye fields (FEF) being one of the most active regions involved in various tasks requiring the use of top-down attention (Bourgeois et al., 2022; Miller & D'Esposito, 2005; Moore & Fallah, 2004; Veniero et al., 2021; Wardak et al., 2006; Zhou & Desimone, 2011)

Certainly, this theoretical differentiation of selective attention as driven by bottom-up or top-down factors is quite simplistic. While attention is driven by the saliency of low-level stimuli features as well as our goals and intentions, it is not limited by them. Many other factors are involved in the allocation of attention. The amount of cognitive load or mental effort required to successfully perform a task can impact attention (Amadieu et al., 2011; Y.-C. Lee et al., 2007, 2009). When a task is demanding, attention may be focused on specific aspects likely to help with task performance, while other irrelevant information is ignored, or even suppressed. Our attention is often drawn to stimuli that elicit emotional responses (Bertels et al., 2010; Failing & Theeuwes, 2018; Finucane, 2011; Hartikainen, 2021; Jefferies et al., 2008; Krebs et al., 2011; Baddeley, 1972). Emotionally valent stimuli (for e.g., threatening or rewarding) tend to capture attention more easily. The surrounding context also influences attention (Janzen & van Turennout, 2004; Peelen & Kastner, 2014). It can guide attention toward relevant objects or locations based on contextual cues or scene familiarity, like safely navigating through a busy city or concentrating on the slides and presentation in a classroom setting. Attention is influenced by social factors, including the presence of others and social cues such as eye gaze or body language (Böckler et al., 2014; Chan et al., 2020; Nummenmaa & Calder, 2009). We tend to prioritize such social stimuli and cues in our environment to facilitate social interaction and choose appropriate behaviours in different settings. Attention can be shaped by learning and past experience (Awh et al., 2012; Chetverikov et al., 2020; Gunseli et al., 2016; Nakayama et al., 2004). Through repeated exposure and reinforcement, we learn to selectively attend to stimuli that are associated with important outcomes or consequences. Individual factors, such as personality traits, expertise, and prior knowledge can also influence our attentional control settings (Abernethy & Russell, 1987; Feldmann-Wüstefeld et al., 2015; Hershler & Hochstein, 2009; McIntyre & Graziano, 2016; Tipper & Baylis, 1987). For example, individuals with expertise in a certain domain may have enhanced attentional abilities for stimuli relevant to that domain, like a painter or photographer would be drawn to specific details that affect scene composition when assessing a scene (S. Park et al., 2022; Tallon et al., 2021). All of these factors interact and contribute to attentional control settings to determine the next target of fixation (Irons et al., 2012; Irons & Leber, 2016; Taylor et al., 2016; Yantis, 1998), however, how exactly this happens remains

unclear.

Study I: Modelling saliency, priority, and other creatures of selective attention

A map is not the territory it represents, but, if correct, it has a similar structure to the territory, which accounts for its usefulness.

— Alfred Korzybski, *Science and Sanity,* p. 58

One approach to studying the possible processes underlying target selection and their interactions is through computational means, and the most successful interpretations so far have been based on the idea of a saliency map. A saliency map may be defined as a 2D accumulation of visuo-spatial information that includes, but is not limited to, bottom-up information from input image features. The saliency map is thought to reflect neural activity in the superior colliculus (SC) where the field size increases with depth (Sparks & Hartwich-Young, 1989). The SC is a layered structure in the dorsal part of the mammalian midbrain that orients our attention toward specific locations. Activity in superficial layers of SC has been used to predict fixation locations in free scene viewing and visual search paradigms (Adeli et al., 2017)

Another proposal that has been suggested relatively recently based on neurophysiological studies of saccade generation in monkeys is a priority map (Bisley & Mirpour, 2019). Studies in non-human primates have shown that besides the SC, the neural map driving eye movements lies in the lateral intraparietal area (LIP) (Haji-Abolhassani & Clark, 2014; Henderson et al., 2013). This simple map may then work with the frontal eye fields (FEF) to tag previous locations (Mirpour et al., 2019) and SC as the result of the winner-take-all (WTA) process (Bisley & Mirpour, 2019).

To simplify, the difference between a saliency and a priority map is in that the priority map is a more accurate reflection of selective attention mechanisms. It accounts for attentional control settings by integrating multiple factors that manipulate the allocation of attention. Nevertheless, it is not uncommon to see a priority map referred to as a saliency map, especially in older models. For many models, the saliency map has been the primary component driving attentional shifts over time. Findlay and Walker (1999) proposed a model of saccade generation which avoided using the notion of attention. Instead of attentive processes, they used parallel processing and competitive inhibition as factors explaining the choice of fixation targets. Their model had two pathways – a spatial and a temporal pathway. Similarly, Logan (1996) avoided references to attention in his CODE theory of visual attention. Both approaches were based on the idea of a salience map created via spatially distributed coding, a mechanism determining new fixation locations. Perhaps the most influential model of visual saliency was the one proposed by Itti & Koch (2001). It provided a plausible explanation of the mechanisms behind bottom-up saliency, and it pushed the field into producing new ideas and provided suggestions for improving the original proposal. However, models of visual saliency have become so

diverse over the decades and there has been such an influx of them that it has become almost impossible to determine which ones are relevant to the specific goal of cognitive neuroscience: to explain how the brain selects where to look. To address this issue, we performed an in-depth analysis of the current literature and of existing models and created a guide consisting of five major requirements to help determine whether a saliency model is useful for computational cognitive neuroscience.

Manuscript I: *Salience Models: A Computational Cognitive Neuroscience Review*

Problem

Patterns of saccades and fixations from experimental studies have been used in deep learning and robotics to train networks to perform various tasks, from algorithms for basic object recognition to more complex tasks. However, the goal of understanding the underlying mechanisms of visual attention remains. Most of these models excel at functional tasks but lack the critical aspects of visual salience modelling from the perspective of computational cognitive neuroscience.

Significance of the paper

The paper provides an in-depth analysis of the state of the field of vision modelling. It allowed us to identify existing weaknesses and gaps in current models of visual salience from the point of view of computational cognitive neuroscience. This review of existing models and approaches in the field, along with their advantages and shortcomings not only helped us identify the focus of our research but has proved to be a useful guide to the field in general.

Goal

We are interested in using deep learning to create a model that accurately portrays the neurobiology of visual salience with explicit computational tools, generates shifts of attention, and considers spatial and temporal aspects of visual attention. To address the issue of the overwhelming number of existing saliency models, we devised a classification approach based on these five parameters.

Summary

Over the past decade we have been observing an ongoing trend with deep-learning neural network models used at every turn. The reason is simple: these computational tools offer incredible versatility with regards to field of employment and, when used correctly, can outperform humans at some tasks. In cognitive psychology and neuroscience, computational modelling has made it possible to implement precise independent theories (Broadbent, 2013; Guest & Martin, 2021; Schlesinger & McMurray, 2012) and combine them with real behavioural data to observe their interactions. This makes it possible to come up with such research questions and

hypotheses that would not have been possible with individual experiments alone. Together, experimental data from oculomotor measurements combined with a computational approach create a long train of research questions which, upon investigation, create even more research questions; but this simultaneously deepens our understanding of what works, how it works and –if we are lucky and persistent – why it works…or does not work.

To help sort existing models of visual attention, we propose five major aspects that we believe the perfect computational cognitive neuroscience model of visual saliency should strive to include. First, such a model should be biologically plausible. This implies that it should aim to mimic some physiological aspect of visual processing. An example would be a component of the model that would simulate centre-surround receptive fields, or the way specific neurons in area V1 respond to orientation. Second, the model should be computational. This means that it should use computational algorithms as a tool for modelling visual processing. These algorithms may be based on various mathematical and computational approaches, like using a Bayesian framework for a model, or integrating a specific technique for finding a global maximum within an array of values representing the strength of independent neuronal responses. Third, it should be generative. This distinguishes such models from classifiers based on the output they produce. Classifiers learn to categorically discern between the types of input ('either 1 or 0'), literally sorting the input into separate classes it has the knowledge of. Generative models, on the other hand, learn to produce more complex output. Usually, the aim of such models is to produce new output based on the given training input. In our case, a generative model should be able to catch the patterns and distributions of human oculomotor behaviour and learn to generate similar behavioural patterns. Fourth, such patterns of the model's resulting behaviour should incorporate spatial locations of the most interesting and/or relevant locations in the visual scene. The fifth, and final point, is that the output should also include the temporal characteristics accompanying the spatial shifts of attention. Together, the spatial and temporal aspects would mimic saccades and fixations and their temporal dynamics.

In accordance with our classification and due to the growing popularity of deep learning approaches to modelling, we distinguish two categories of saliency models: the category of computational saliency models that are further subdivided into six more distinct groups; and a separate broad category of deep-learning classifier models. A summary of the classification is provided in Table 1.

Table 1. A Classification of Saliency Models of Vision from a Computational Cognitive Neuroscience Approach

Category	Subgroup	Description	Examples
Computational saliency models	Direct variations of the original saliency model	Models that are similar in basic structure to the original Itti and Koch (2001) model but alter it in some crucial aspects	Attentional Selection for Object Recognition (Walther et al., 2002); SAFE (Draper & Lionelle, 2005); Guided Search (Wolfe, 1994); GBVS (Harel et al., 2006);
	Biologically driven models	Models with architecture similar to the original Itti and Koch (2001) model but a focus on the detailed biological and neurophysiological aspect of visual saliency	Bottom-up saliency map in V1 (Koene & Zhaoping, 2007); Bottom-up saliency model with LGN (Park et al., 2002); cortical magnification model (Aboudib et al., 2016); feature-based saliency model with top-down inference and FEF (Hamker, 2006)
	Top-down models	Models that improved the top-down component of the original Itti and Koch (2001) model	Guided search (Wolfe, 1994, 2021); The Feature Gate model (Cave, 1999); probabilistic model of attention (Rao & Ballard, 2005); predictive coding model (Rao & Ballard, 1999)
	The 'what' and 'where' models	Models that improve on the original saliency model taking the theory of the ventral and dorsal streams into account	Attention-guided recognition (Rybak et al., 2005); the Neurodynamical model (Deco & Rolls, 2004)

	Object models	Models that use object instead of feature salience combined with the bottom-up approach of the original Itti and Koch model (2001)	NAVIS (Backer et al., 2001); object-based visual attention inspired by the ICT theory (Duncan et al., 1997; Sun & Fisher, 2003)
	Computationally driven models	Non-deep learning models that use complex computational approaches to enhance the original model	The Selective Tuning model (Tsotsos et al., 1995); Model of attention and recognition by information maximisation (Schill, 2005); attention guidance model using a Bayesian architecture and global scene context (Torralba et al., 2006); SUN (Kanan et al., 2009); Boolean map saliency detection (Zhang & Sclaroff, 2013); SVM classifier fixation prediction model (Judd et al., 2009)
Deep learning classifiers	CNN	Saliency models based on convolutional neural network architecture	Deep Gaze 2 (Kummerer et al., 2017); EML-NET (Jia & Bruce, 2020); DeepFix (Kruthiventi et al., 2017)
	GAN	Saliency models based on generative adversarial networks	SalGAN (Pan et al., 2018); task-specific visual saliency with memory augmented GAN model (Fernando et al., 2018)
	RNN	Saliency models based on recurrent neural networks	DSRCNN model (Tang et al., 2016); Recurrent attentional saliency network (Kuen et al.,

			2016); DSCLRCN (Liu & Han, 2018)

Metrics. In addition to the analysis and classification we describe available evaluation and metrics approaches that can be used to assess the performance of saliency. The metrics included in the manuscript provide a brief description of the evaluation techniques used in the Tübingen/MIT saliency benchmark (Bylinskii et al., 2019; Judd et al., 2012; Kümmerer et al., n.d.) that we believe should serve as the minimal assessment criteria for saliency models. We believe that together with our proposed classification approach this can help standardise the way models of visual saliency are evaluated, at least in the field of computational cognitive neuroscience.

Conclusion

We conclude that in the recent years the focus in saliency modelling has been shifted from visual attention towards more computationally heavy AI and deep learning fuelled spatial fixation classification. Most widely used evaluative tools also tend to focus more on model performance related to spatial location predictions while ignoring other aspects of visual attention. Though these approaches are undoubtedly important in the field of AI and robotic vision, they tend to be overused in the field of cognitive neuroscience the goal of which is to improve our understanding of the human brain. Therefore, we believe that there needs to be a shift of attention in the computational cognitive neuroscience community, where '*computational* would imply using powerful mathematical and programming approaches as a tool for enhancing our understanding of the brain, not a distractor that leads to us questioning the field's value. Namely, we propose that our models should be tested against human data not only in terms of spatial prediction performance but include temporal distributions of typically human shifts of visual attention and neurophysiological patterns of oculomotor behaviour, as well as visual processing errors that people often make.

Manuscript Ia: *Deep Learning Neural Networks as a Component of a Model of Saccadic Generation*

To address the issue of the spatially focused classifier saliency models, we attempted to create our own model inspired by the classic Itti and Koch saliency model (2001). The aim was to improve the bottom-up component of the original model by replacing the parallel feature maps with an autoencoder (Bengio, 2009; Japkowicz et al., 2000; Kramer, 1991) deep learning neural network and adding a leaky-integrate-and-fire (LIF) (Gerstner et al., 2014; Usher & McClelland, 2001) component to imitate the temporal distribution of the shifts of attention typical for human observers.

Our approach proved the model to be efficient with regards to spatial accuracy, producing fixation map predictions that matched human data (see Fig. 4). The

temporal LIF approach did fall into the range of normal human saccade reaction time distributions, but it was far from optimal. The major issues were the deterministic nature of the predictions and a bimodality present in all testing phases (see Fig. 5).

Figure 4. Model Performance: Saliency Map Spatial Predictions
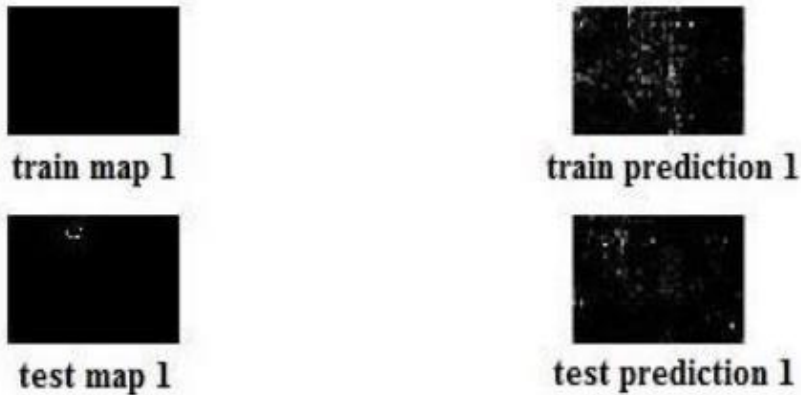
Note. The figure demonstrates an example of the saliency map predictions. White patches represent fixation locations.

Figure 5. Model Performance: Saccade Reaction Time
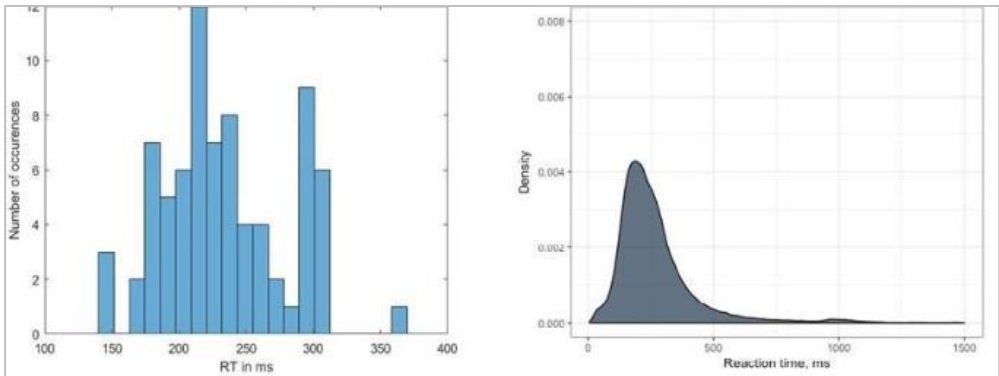
Note. The figure demonstrates an example of the temporal distribution of saccade reaction times. The left panel shows the model predictions. The right panel represents a distribution of average human saccade reaction times. The model matches the mean of human saccade reaction times, but there is a bimodality – two peaks ca. 220ms and 300ms – that is not typical in human observers.

Manuscript Ib: *Temporal Limitations of the Standard Leaky Integrate and Fire Model*

The assessment of the temporal LIF approach was described in a separate publication (Merzon et al., 2020). We used the default parameters of the algorithm to test its efficiency in our specific task – accurately modelling the saccade reaction times based on input scenes from the *LabelMe* dataset (Russell et al., 2008). We further optimised the parameters to produce a better temporal prediction output. The

general structure of the tested model included the use of a deep learning approach similar to the one proposed in Manuscript Ia for the visual component, and a LIF approach (Usher & McClelland, 2001) for the temporal component. Specifically, the spatial component used both supervised and unsupervised learning by combining stacked autoencoders (Bengio, 2009; Japkowicz et al., 2000; Kramer, 1991) and a softmax approach (Goodfellow et al., 2016) to produce maps of spatial salience. These maps were forwarded to the temporal LIF component that incorporated a winner-take-all algorithm (Walther & Koch, 2006) to produce respective temporal distributions for the spatial (saccadic and fixational) predictions. This general structure is visually implemented in Fig. 6.

Figure 6. The Structure of the Model Tested in the Study
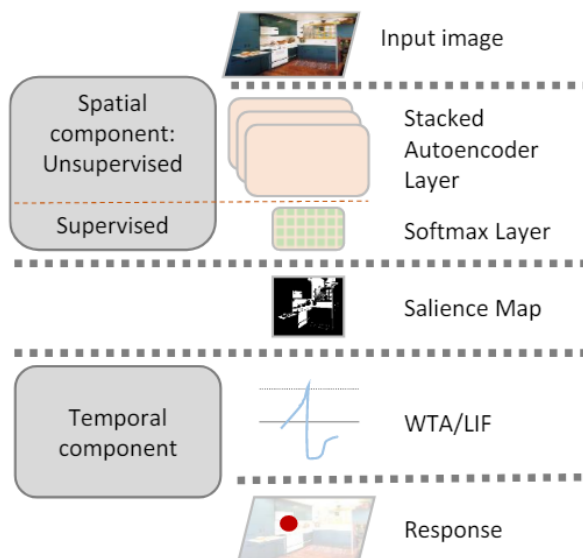
Note. The figure visually demonstrates the structure of the tested model. The output of the spatial component was a saliency map of the input image which was then used to generate a saccadic latency estimate. The final model response was a predicted visual fixation with its spatial parameter – a spatial location, and its temporal parameter – the 'saccade' latency.

The results of our tests showed that the algorithm set to default parameters was not the best predictor of visual temporal characteristics. Our parameter optimisation allowed us to find a fit that was able to match real human data with respect to the mean and z-test, and saccade latency distributions, though less so in the latter. We believe that the inability to produce biologically accurate saccade latency predictions as well the deterministic nature of the LIF model prediction is connected to the lack of a randomness component necessary for outputting a distribution.

Discussion

The set of the three abovementioned studies have shown that, in general, it is possible to use newer computational approaches to improve classical theories. We believe that each level of a computational model of saliency should make different contributions to the model, like a real visual system would process visual information in stages. Therefore, we introduced a theoretical framework based on five major aspects for computational cognitive neuroscience modelling and driven by that framework, built and tested several models of visual saliency. Our attempts have been rooted in cognitive theory, used several computational approaches, and took the physiological and temporal aspects of saccade generation by the SC into account. These attempts were more of a 'proof of concept' rather than a full model of biological vision, but they allowed us to define the next important stages in our research. Namely, the physiological accuracy of the model implied shifting the focus towards the improvement of its spatial predictions by considering the brain's native frame of reference, which is retinotopic, while taking into account the theoretical underpinnings of the cognitive mechanisms that govern selective attention.

Study II: Retinal relativity and gaze contingency

A process cannot be understood by stopping it. Understanding must move with the flow of the process, must join it and flow with it.

— Frank Herbert, *Dune*

Diving deeper into the issues of existing models of vision we identify another major trait most of them share - the spatial predictions they generate are based on the spatiotopic coordinate system. One way of improving spatial predictions in such models is through the implementation of a retinotopy-based component. What do we mean by '*spatiotopic*' and '*retinotopic*'? In a nutshell, the term *retinotopic* is used to refer to the location of the objects surrounding us in relation to our retinae. This is different from the *spatiotopic* reference system where the location of objects remains fixed even if we move (given that the objects are static). We can use retinotopic dimensions as opposed to spatiotopic dimensions to better understand how the visual system operates (see section *Beyond the End: On the Way to the Next Fixation Location* on p. 45 for a more detailed explanation of the two coordinate systems and why they are important for visual perception).

A non-invasive approach to retinotopic space is through the implementation of so-called gaze-contingent displays. In such setups, the point of gaze fixation is constantly monitored and updated at high frequencies by the eyetracker. One important difference from passive gaze position recording is that the location information is used for active interaction, such as a visual mask or patch displayed in the visual scene on the monitor that is 'tethered' to the observer's gaze. Such setups make it possible to simulate scotoma and other visual abnormalities in healthy subjects and study their performance. Such setups are usually built on the constant feedback-feedforward communication between high-frequency eyetracking equipment and a computer powerful enough to maintain the speed of the incoming updates, the perpetual generation of stimuli and data recording. They make it possible to control the visual input presented to participants in real-time, by these constant online updates of their gaze location. Thus, we can manipulate the way participants see visual information located in the fovea and in the periphery to study the effects of foveal eccentricity on retinal visual processing. This artificial manipulation of visual input is similar to an internal cognitive mechanism that is believed to be adaptable to the task being performed. In our study we refer to this as the Functional Field of View (FFV). The FFV may be broadly defined as the region in space where attention is allocated, and visual information is processed efficiently. Attentional control settings play a crucial role in determining the size and allocation

of the FFV. By investigating the relationship between the FFV and task performance, we can gain a deeper understanding of how attentional control settings impact cognitive processes, such as the processing of visual information, decision-making, and task performance, as well as their interactions. Therefore, we attempted a study to collect more behavioural data related to retinotopic processing using a gaze-contingent paradigm to see if the amount of effectively processed visual input changed based on the cognitive demands of the task.

Manuscript II: *Assessing the Optimal Size of the Functional Field of View by Task in a Gaze-Contingent Search Paradigm*

Problem

While bottom-up saliency models perform well with regards to computer vision and novel computational approaches, they fail to capture the more natural aspects of human oculomotor behaviour, like adapting to task demands, making mistakes, or translating spatiocentric to retinocentric coordinates. By introducing new data with different parameters of search to vision models we can try to model the way attention is allocated based on the various context-dependent control settings. Consequently, this would allow us to reverse engineer the mechanisms that drive certain types of visual behaviour. A model that operates in brain-native visual space coordinates while also taking task and context demands into account can potentially improve our understanding of how much visual information is taken in at a given time at a certain location in the visual scene.

Significance of the paper

In general, the study adds valuable information for our understanding of selective visual attention and cognition. Specifically, our gaze-contingent setup helps us understand how we process visual input based on the size of the aperture limiting the area available for extracting useful information from the visual scene. The stimuli used in this study are more semantically rich than stimuli consisting of low-level features like simple lines and line combinations. The effects we observed suggest that the changes in our visual field are driven by more factors than we considered previously.

Goal

The major questions in this study were: *does the amount of information we can process remain fixed in any situation? If not, what exactly does it depend on and what are the underlying cognitive processes associated with the changes in the 'volume' of visual processing ? Does access to more information in the visual field improve performance during visual search? Does the mental difficulty of the given task affect this performance?* These questions served the main goal of learning how changes in the FFV affect attentional set and consequently the performance in a cognitively

demanding search task. The collected data may be used to improve the spatial predictions of visual saliency models.

Summary

An aperture for attention?

There has been an ongoing debate related to the way we perform visual search. So far, the tendency has been leaning towards easier object-based, or item-based visual search paradigms, such as those used in Guided Search (Wolfe, 1994, 2007, 2021), Search via Recursive Rejection (SERR) (Humphreys & Muller, 1993) based on the Attentional Engagement Theory (Duncan & Humphreys, 1992) or Competitive Guided Search (Moran et al., 2013). However, there have been alternative proposals stating that out of the laboratory we tend to perform visual search in a fixation-based manner, suggesting that eye movements play an important role in visual search and are not just a confound to control for. One of the most stimulating debates over visual search in the recent years has erupted in response to a proposal to shift from item-based towards fixation-based paradigms by Hulleman and Olivers (Hulleman & Olivers, 2017b). According to a fixation-based approach, instead of making visual fixations per individual objects, we make a fixation to process all the information available within the limits of that one fixation. The spatial area around the current gaze fixation point is limited, and different terminology has been used to describe it by different researchers over the years. Among the most common are *attentional window* (Belopolsky & Theeuwes, 2010) functional viewing field, FVF (Sanders, 1970), *useful field of view*, UFOV (Ball et al., 1988), *area of visual conspicuity* (Engel, 1971) and *visual span* (Jacobs, 1986; O'Regan et al., 1983). In this study we use the term *functional field of view* – FFV. Though related, the concept of the FFV should not be confused with the spotlight of attention (Posner et al., 1980) or the zoom lens model (C. W. Eriksen & St. James, 1986). While the spotlight of attention and the attentional zoom lens are attentional mechanisms that are not necessarily connected with oculomotor activity, the FFV is tied to the locus of fixation at any timepoint (Young & Hulleman, 2013). One major characteristic of the FFV is its dynamic adaptability: its size may vary depending on the cognitive load associated with the task performed at a given time, increasing in response to lower load and narrowing with increasing load (Gaspar et al., 2016; Young & Hulleman, 2013).

Limited by Physiology

The amount of visual information that the brain can take in within one fixation is determined on the one hand by physiological factors which can be seen as 'hardware', and by cognitive factors on the other. Physiological factors imply that input from the visual field is limited due to physiological constraints. Like the way electric conductivity depends on the area of the conductor and the materials it is made of, so is visual conductivity related to human 'hardware' parameters. This leads to the

following questions: is it possible to quantitatively assess the amount of information entering our brain? Is this a fixed amount or does it fluctuate in some way? Koch and colleagues (Koch et al., 2006) addressed these specific questions by recording information transmission rates in guinea pig retinal ganglion cells. This study allowed them to extrapolate the results to humans and conclude that the human retina has a general physical data transmission rate comparable to the rate of an ethernet connection (10 bits · s−1). This transmission rate is dependent on the degree of cell activation caused by the stimuli, allowing to take in more information from the visual field at a time. This implies that the rate is by no means fixed, but rather depends on the kind of visual stimuli.

Another biological factor important for the way we perceive the visual scene is the difference in visual acuity between the fovea and the periphery. It has been determined that the area with the highest resolution is the fovea, which is only as large as a thumbnail on an extended hand, or approximately 1° visual angle. The fovea is responsible for high acuity vision and does this via a tight array of photoreceptors that has been compared to the density further in the periphery (Curcio & Allen, 1990). However, existing terminology can make the division of the eye into visual areas slightly confusing. Functionally, the area responsible for foveal vision is actually called the foveola that is 1° in diameter (Hendrickson, 2005; Kroell & Rolfs, 2022; Wandell, 1995; Yellott et al., 1980). According to Wandell (1995), the fovea is a region with a diameter of 5.2°, which, together with the parafovea surrounding it (size, 5°-9°) and perifovea (9°-17°), comprises the macula. The macula is itself a part of the area generally referred to as the central visual field with a diameter of 60°. The region extending beyond this point to almost 180° horizontally maintains our peripheral vision. Thus, we can conclude that our visual resolution deteriorates as a function of eccentricity from the foveola with no rigid boundaries between the foveal and extrafoveal regions, allowing for a relatively smooth transition from one into the other.

Our visual perception of the surrounding world is also influenced by the cellular structure of the retina (Strasburger et al., 2011). The distribution of rods and cones across the retina is not uniform, therefore our perception of the environment is not uniform across the visual field (Anderson et al., 1991; Curcio et al., 1987; Curcio & Allen, 1990; Jóhannesson et al., 2018; Perry et al., 1984). The foveal part of the retina consists mainly of cones while the peripheral part of the retina is more densely packed with rods than cones (Curcio et al., 1990; Grünert & Martin, 2020; S. C. S. Lee et al., 2019; Wells-Gray et al., 2016). This allows us to see in dimly lit conditions and notice changes in environment, however, as these cells have large receptive fields, this comes at the price of reduced colour perception and lesser sensitivity to detail.

All these physiological factors set the physical limitations of the capacity of visual processing. However, the performance of this processing system is guided by cognitive mechanisms that depend on various external and internal factors.

Shaped by Cognition

Although the FFV is under physiological constraints via the limited capabilities of our 'hardware', it is an attentional mechanism that is dependent on the kind of visual input available at a given moment. Thus, a common approach to studying the FFV is through the manipulation of the performed task. The hypothesis that the FFV is dynamic based on the task has been shown multiple times. Perhaps one of the most interesting studies of the FFV was carried out by Motter and Simoni (Motter & Simoni, 2008). They observed the changes in the size of the FFV in visual search with and without overt eye movements. They found that in the condition when participants had to suppress eye movements while searching for a target among distractors, the size of the FFV gradually increased to fit as much information on the search set as possible. They observed that despite this ability to attend to stimuli in the periphery, participants preferred to perform search while keeping the target of attention in their fovea by making eye movements – the natural way humans explore their environment. Direct support for the idea that FFV size varies by task demands comes from Young and Hulleman (2013). They argued that visual search is based on visual fixations (rather than items), where the size of the FFV - and the number of items within it - depends on several factors, such as the complexity of the visual scene, the difficulty of the search task and the characteristics of the target. They claim that the FFV is dynamic and depends on how discriminable the target is: low discriminability results in a smaller FFV, while higher discriminability results in a larger FFV. These claims have led to much debate in the visual attention and perception community (Hulleman & Olivers, 2017a). To gain insight into what determines the amount of visual information that passes through FFV, we conducted a visual search experiment. The main goal was to manipulate FFV directly, via external manipulations. We were interested in assessing the relationship between the load associated with the task being performed and the size of the FFV. We also wanted to gain insight into what determines the amount of visual information passing through it. In order to try answer these questions, we conducted three experiments.

Experiment 1: Non-Gaze-Contingent Test Setup

Methods. To test if our condition manipulations were successful, we decided to run an initial experiment without a gaze-contingent setup. We designed a visual search task where participants had to find a target among a set of distractors in a grid of 45 items per trial (Fig. 7). The presented stimuli were borrowed from a set created by Chetverikov and colleagues (Chetverikov et al., 2018) for an earlier study. We manipulated the size of the stimuli in two task conditions – trials with more discriminable search targets had larger stimuli (1.3°) while trials with less

discriminable targets had smaller stimuli (0.7°). The number of items presented on the display remained the same across all trials and conditions.

Figure 7. An Example of the Stimuli Grids in the two Conditions used in the FFV Experiment

| Large | Small |
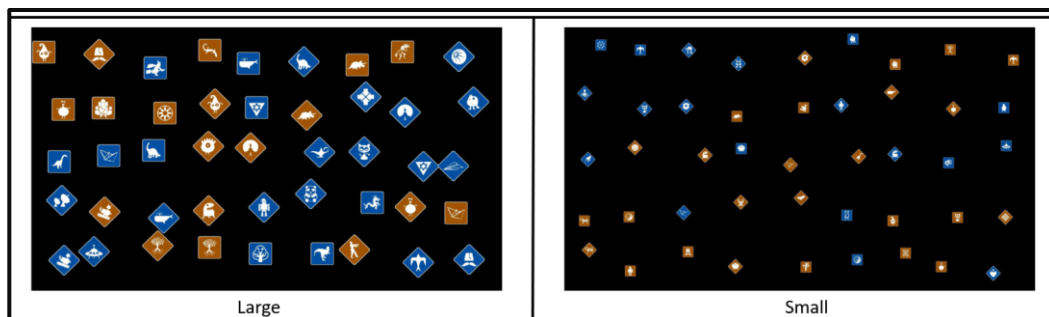
Note. The figure depicts some of the stimuli we used in our FFV visual search task randomly placed for one separate trial. The search conditions were manipulated through the size of the items in the set. Large stimuli (1.3°) were used in the less demanding condition (left panel), small stimuli (0.7°) were used in the more demanding condition (right panel).

Another parameter we manipulated was the absence or presence of the target. Participants had 5000 ms to respond if a target shown at the start of the trial was present among the search set: in case it was present they had to press the 'f' key on the keyboard, if it was absent, they were asked to respond by pressing the 'j' key on the keyboard. If no response was given before the end of the trial, then it was registered as an incorrect trial. A trial consisted of the following events: at the start of the trial, a search target was briefly (500 ms) shown at the centre of the screen. Immediately after this the search set appeared for 5000 ms where the participants had to provide a manual response (Fig. 8).

Nine participants took part in the study. Each one contributed to six blocks of trials with 36 trials in each block. All the blocks contained an even number of absent/present and large/small trials. Participants performed a practice block before the main experiment to get accustomed to the task.

Figure 8. General Design of the Experiment
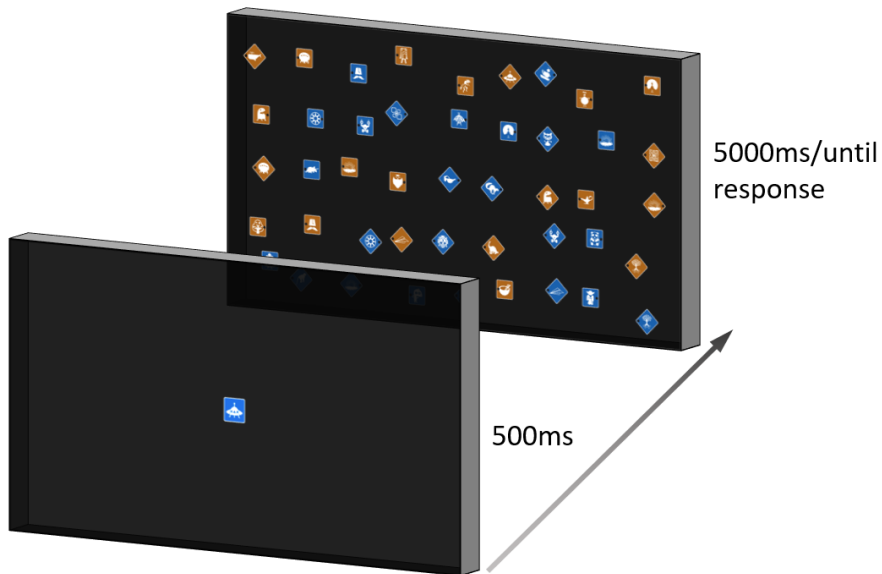
Note. The figure depicts the general design of the experiment. The timing applies to one trial in the experiment. First, a target was shown for 500 ms. Then, a search set containing 45 icons appeared for 5 seconds during which a response was expected. The trial ended as soon as there was a response or when it reached the search time limit.

Results. We performed a 2x2 ANOVA analysis (item size *(large vs small)* x target presence *(present vs absent)*) to assess response times and accuracy. We found that item size and target presence affected the accuracy of participant responses (Fig.9, left panel). Participants provided more correct responses on large item set trials as compared to small item set trials. They were also more accurate on target present trials than on target absent trials.

The analysis of response times supported the results of the accuracy analysis, showing longer search times on sets with small items than on sets with large items (Fig.9, right panel). Response times on trials with no targets were also longer than on target-present trials.

Figure 9. Participants' Performance in the Visual Search Task
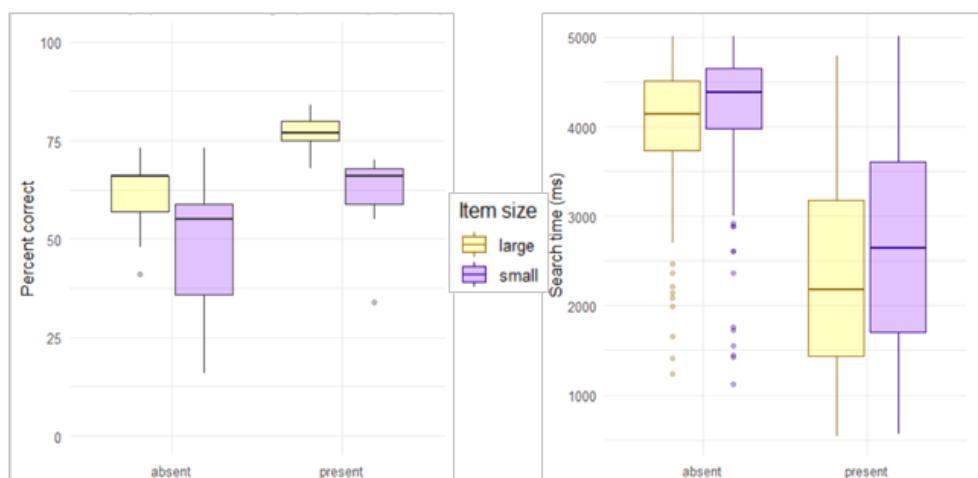
Note. Participants' performance in the visual search task given item size and whether the target was present or absent. Accuracy (left panel) is shown as the percentage of correct responses (%). Search time (right panel) is based on manual response times (ms).

Discussion. We manipulated stimuli size and target presence in a visual search task to see if accuracy and response times would differ in the manipulated conditions. The results demonstrated that the manipulation of item size and target presence affected participant performance. On trials with small item size and absent targets they demonstrated longer search times and lower accuracies, whereas trials with larger item size and present targets resulted in the best performance.

However, we ran into a problem when adding the gaze-contingent component to the current design. The problem was in the way we manipulated our search parameters. By adding manipulations of the aperture size to the manipulations of item size, we were introducing a potential confound. The number of items available in the aperture at any given time would differ for the two item size conditions, resulting in a crowding effect and making comparisons between the two item size conditions difficult. We therefore decided to change our approach within a new experiment with a gaze-contingent setup.

Experiment 2: Gaze-Contingent Test Setup

Methods. We set up a version of the study that included the aperture. In this setup we changed the way we manipulated the load associated with the search task: the manipulation of item size was abandoned, and instead of that we used item rotation to control the level of abstraction of the stimuli. We employed a gaze-contingent display with three aperture sizes, Large (16°), Medium (8°) and Small (4°), while participants searched for a target among distractors on either upright or inverted trials (Fig. 10, left panel). As in the non-gaze-contingent study, at the start of the trial a stimulus icon was presented for 500 ms. Immediately after this the search set

appeared for 5000 ms. These icons were placed among the search set on target-present trials, while on target-absent trials they were not present in the search set. If it was present, participants had to press the 'f' key on the keyboard; if it was absent, they were asked to respond by pressing the 'j' key on the keyboard. If no response was given before the end of the trial, it was registered as a no-response trial.

Eight participants took part in the gaze-contingent pilot study. Each one contributed to six blocks of trials with 100 trials in each block. All the blocks contained an even number of absent/present and upright/inverted trials. A practice block with a very large (20°) aperture was used to introduce the participants to the task.

Figure 10. The Conditions used in the Gaze-Contingent Test Experiment
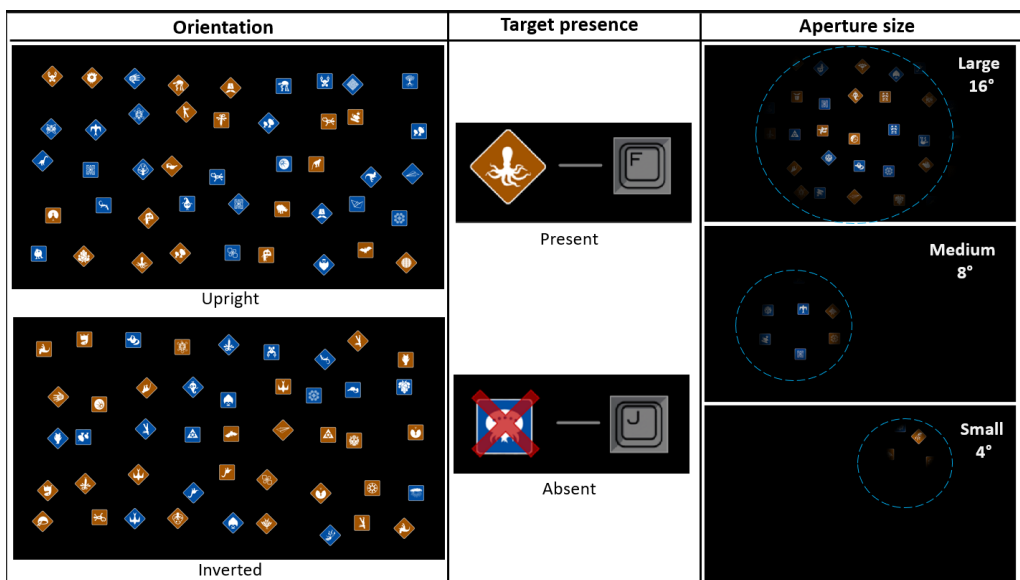
Note. The parameters of the search task were manipulated through the orientation of the set stimuli. Upright stimuli were used in the high-discriminability condition (left panel, top), inverted stimuli were used in the low-discriminability condition (left panel, bottom). The target was either present (central panel, top) or absent (central panel, bottom), adding an additional source of cognitive load to the search. Aperture size differences are depicted in the right panel.

Results. We performed a 2x2x3 ANOVA analysis (orientation *(upright vs inverted)* x target presence *(present vs absent)* x aperture size *(large vs medium vs small)*) for accuracy and response times. For accuracies, we observed a large amount (1815; >50%) of trials where no response was given as compared to correct (2092) and incorrect (895) responses (Fig. 11).

Figure 11. Accuracy by Response Type
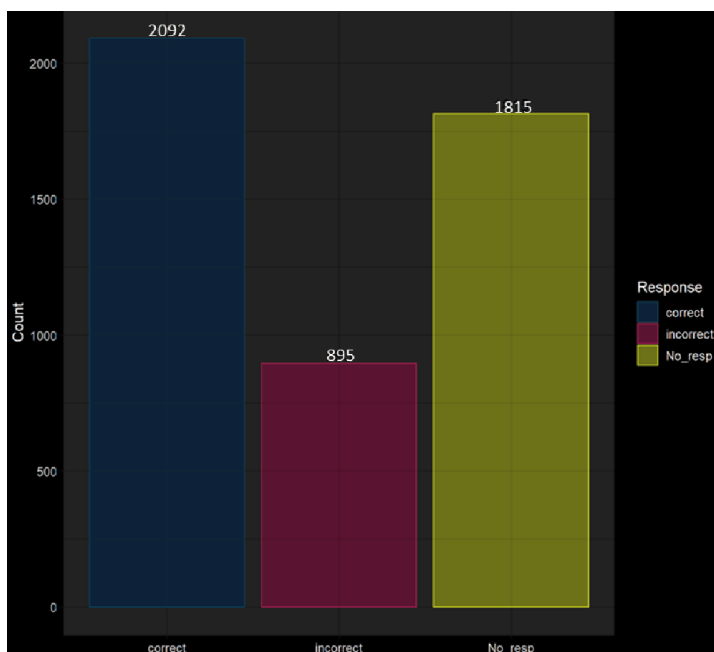
Note. The figure depicts the number of correct (blue), incorrect (red) and absent (yellow) responses. The corresponding count values are displayed at the top of the bars.

Taking aperture size into account, trials with the small aperture had the lowest number of correct responses and largest number of no response trials. The analysis of mean response times did not reveal any significant effects.

Discussion. The response time analysis resulted in a lack of sufficient data to draw any conclusions. The analysis of accuracy revealed a large number of trials with absent responses. Participants reported having extreme difficulty performing the task. Their biggest challenge was associated with scanning the display with the smallest aperture within 5000 ms: they were not able to build a complete visual image of the entire set, decide if they saw a target and produce a manual response before the trial ended. This resulted in them admitting to guess-responding on most trials with the small aperture. Once again, we were forced to reconsider our experiment design.

Experiment 3: Main Gaze-Contingent Setup

Given the participant reports and poor performance in the test experiment, we decided to change the design of the task. As in experiment 2, we employed a gaze-contingent display while participants searched for a target among distractors in two conditions: upright and inverted. However, we increased the search time from 5000 ms to 7000 ms. We also changed the size of all the three apertures from 4°, 8° and 16° to 6°, 12° and 18° for the small, medium, and large apertures, respectively. To

substitute the dependent variable, we used the same stimuli but replaced target presence with a 2-alternative-forced-choice task (2AFC) (Fig. 12). Participants (N=20) manually responded which side the dot was on using a keyboard – 'a' key for the left side and 'd' key for the right. The eye-tracking procedure was the same as in the previous gaze-contingent experiment.

Figure 12. Stimuli Modified for the 2AFC Gaze-Contingent Task
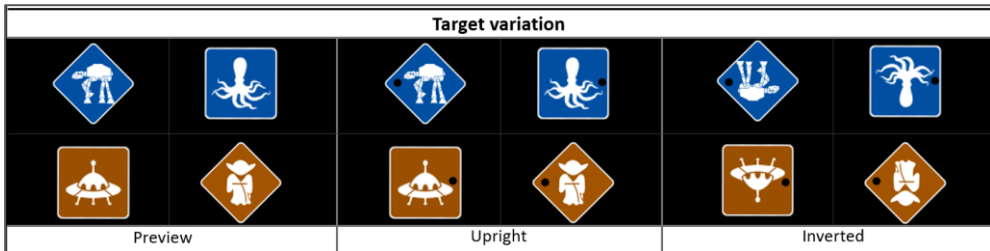
Note. The figure depicts some of the stimuli we used in our FFV visual search task. They were placed randomly for each separate trial. The original stimuli were adapted for the 2AFC task - black dots were randomly placed either on the left or the right side of the stimuli. The target flash at the beginning of each trial ('Preview') appeared without the modification (examples in left panel). In the search set we used the modified target (the dot was placed randomly on one side), so it did not differ from the distractors.

Results. We measured how participants' response times and search accuracies changed in response to the manipulation of the aperture and rotation of the stimuli. We found that neither the small nor the medium apertures had negative effects on search efficiency in the inverted search condition. There was no strong evidence suggesting any benefit of the small aperture over the large one either. We did make some interesting observations related to the inverted versus upright search conditions: there was a trend signifying a performance improvement on inverted trials with the medium-sized aperture as compared to upright trials. There was no such pattern of significant improvement observed in the small aperture condition, but neither was there any associated deterioration in performance. With the large aperture, however, the task performance was worse on inverted trials than it was on upright trials.

Potential improvements

The study allowed us to identify several potential approaches to improving the design. First, we could improve the clarity of our observations by introducing a control condition. For example, we could have search sets where stimuli features are fully congruent and fully incongruent with the target shape and colour, i.e., where shape and colour are fixed within a block. This would potentially eliminate the manual errors driven by shape/colour mismatch that resulted in quicker response times on incorrect inverted trials (compared to incorrect upright trials). For example, searching for a cat icon inside a blue diamond but responding to a cat icon inside a blue square would be considered a shape mismatch. This could also provide a baseline that mixed

blocks (i.e., containing trials where stimuli with different colours and shapes are intermixed) could be compared to.

Second, making the aperture look more natural could be a potential improvement. The opaque black mask used in the current version of the study does not allow participants to accurately predict the position of the peripheral stimuli nor perform any summary statistics on them. This makes the search with the small aperture ineffective as more eye movements must be made to random locations to see if there are any stimuli there - a process similar to groping the wall in an unfamiliar room in search for a light switch in complete darkness. A more biologically accurate aperture would be a non-opaque representation of a function of distance from the foveola. A potential solution is to use a blurring mask that will not only increase the blur towards the farthest locations, but also gradually strip the colour properties of stimuli farther away from the foveola.

Finally, we analysed participants' patterns of saccades and fixations. Such analyses allow us to see if participants tended to fixate more on matching shapes and colours even if their manual responses were incorrect, compare saccade amplitudes and velocities between the different aperture size conditions and even look at pupil sizes, as higher task demands are associated with a decrease in pupil size. This could provide an insight into attentional mechanisms like IOR and into various search strategies such as overt and covert search, and parallel feature search versus sequential item-by-item search.

Discussion

An important conclusion we can make based on our findings is that for this specific task there is no benefit of using the large aperture. It seems that the medium sized aperture was closest to participants' natural FFV, making it the most effective of the three for the task used in our study. During visual search, attention should be optimally allocated to efficiently detect a target among a set of distractors. At the same time, executive control maintains a level of inhibition to suppress the processing of items with irrelevant features. The stimuli used in this study combined several feature categories, such as shape, colour, and semantics, increasing the cognitive demands of the task. In line with the theories supporting a decrease in the size of the FFV, our task seemed to have had a baseline level of the cognitive demands imposed by the task, where the large aperture did not provide any advantage over the medium-sized aperture. The additional load associated with stimuli inversion even impaired performance on trials with the large aperture as compared to upright trials, leading to longer response times. With the medium aperture participants were probably able to take in visual information optimally. It allowed them to maintain an area large enough to process the several stimuli in the scene in parallel, scanning and picking the next location of interest while efficiently filtering out the distractors that had to be additionally suppressed in the large-aperture condition. At the same time, the

difference in the degree of abstraction between the upright and inverted conditions was not large enough to observe significant improvements on trials using the small aperture. The lack of any deterioration in performance on small-aperture inverted trials does imply that its size did not negatively impact the adaptability of selective attention mechanisms to the task.

Study III: Towards Visual Cognition, Step by Microstep

It has long been an axiom of mine that the little things are infinitely the most important.

- A. Conan Doyle, A Study in Scarlet

Being task-dependent, attentional control settings need to be studied within a variety of contexts. The FFV adjusts to the task being performed and it is tied to the centre of gaze where foveal vision occurs, and we move our eyes to focus the centre of gaze on things that attract our attention (Kowler, 2011). Fixations on these salient or meaningful locations in space usually last around 250 ms. Although they are typically called fixations, these pauses between eye movements are actually not entirely fixed, as the eyes continue to produce so-called micromovements such as drift and microsaccades (Collewijn & Kowler, 2008; Engbert & Kliegl, 2003b; Krauzlis et al., 2017; Rolfs, 2009). Of all the micromovements of the eye, microsaccades are the most debated ones, as neither is their functional purpose certain, nor can researchers agree on a single definition of a microsaccade. We do know that microsaccades, specifically their frequency rates, can be used to study various aspects of cognition and attention (Chen et al., 2015; Engbert & Kliegl, 2004), as well as to gain insight into the mechanisms of oculomotor behaviour (Hafed et al., 2015; Zuber & Stark, 1966). One such insight states that microsaccades and typical saccades share a single inhibitory mechanism (Johnston & Everling, 2006; Paneri & Gregoriou, 2017; Rolfs et al., 2008). Inhibition is a mechanism of executive control – one of the functions of attention (Dodds et al., 2011; Heidlmayr et al., 2020). The oculomotor system is well-suited to studying executive control because of solid theoretical foundations for gaze control mechanisms (Kristjánsson, 2011; Munoz & Everling, 2004). Suppressing an automatic response requires the use of gaze control, and a relatively straightforward approach to observe this suppression is through the antisaccade paradigm (Hallett, 1978).

Executive control and oculomotor behaviour. A major part of our understanding of the role of eye movements in executive control and the neurophysiology associated with it has become available through the study of antisaccades. Antisaccades are voluntary eye movements made to the opposite location to where a stimulus is presented (initially described by Hallett, 1978; see Kristjánsson, 2011 for a review). In a typical prosaccade task, participants fixate on a central fixation point, and should make a prosaccade towards a peripheral stimulus onset. For antisaccades they should suppress reflexive saccades towards the stimulus onset and instead make a

saccade of equal amplitude in the opposite direction. Antisaccades therefore require the participation of the frontal cortex and the basal ganglia to successfully voluntarily suppress stimulus-driven reflexive prosaccades driven by the parietal and occipital areas before the saccadic stimulus appears (Coe & Munoz, 2017), followed by deliberate programming of the opposite saccade vector (Kristjánsson et al., 2001; Munoz & Everling, 2004; Stuyven et al., 2000). If this top-down suppression builds-up throughout the trial, then it might be found in prestimulus microsaccade rates.

Another experimental manipulation requiring executive control is task-switching (Monsell, 2003). Performance costs associated with switching between two or more tasks are seen in a large variety of paradigms and may interact with saccades. For example, task-switching involving prosaccades with a short suppression of response may introduce interference and delay the planning of subsequent saccades (Tari et al., 2019; Tari & Heath, 2019). Pierce and colleagues (2014) have shown that when prosaccades are less probable within a trial block, prosaccade errors increase. In their study, when pro- and antisaccades were mixed within blocks, the percentage of correct responses was lower than in fixed blocks. Moreover, a preceding antisaccade trial increased saccade latency on subsequent trials, irrespective of type. They suggested that the repeated antisaccades inhibited saccade-generating neurons in the frontal eye fields and superior colliculus, increasing their thresholds.

Cognitive load and attentional control. It has been shown that higher levels of cognitive effort may affect attentional control settings (Irons & Leber, 2016). Although attentional control is flexible and serves as a mechanism of performance optimization, the difficulty of the task being performed can impair this performance. According to Irons and Leber (2016) one of the reasons is resource conservation which results in the tendency to avoid adaptation to external conditions via updates of attentional control settings to a new task (task switching). Conflict monitoring associated with the assessment of one's approach to task performance in real time is a taxing process. Several notable studies have studied the effects of the higher cognitive demands imposed by antisaccades by looking at microsaccade rates. Hermens et al. (2010) found lower microsaccade rates for antisaccades than prosaccades, but only on trials with peripheral targets where participants had to delay their responses by 1500 ms. Watanabe and colleagues (2013) observed a reduced frequency of fixational microsaccades with longer fixations preceding peripheral stimulus onset. Microsaccade frequency was also lower on correct antisaccade trials than on prosaccade trials. This differs from the results in Hermens et al. (2010), where there were no observable differences in microsaccade rates preceding pro- and antisaccades in the non-delayed condition. Finally, Dalmaso, Castelli and Galfano (2020) compared the two previous studies by combining the methods they used and additionally analysing pupil size dynamics during the antisaccade task. They found fewer microsaccades and larger pupil sizes before antisaccades, but only within blocks where pro- and antisaccade trials were randomly mixed. Dalmaso et al. (2020)

argued that this reflected differences in cognitive load causing decreased microsaccade rates during more difficult tasks. In an earlier study, Dalmaso, Castelli, and Galfano (2019) also investigated the effects of cognitive load on microsaccade rates in a flanker task (B. A. Eriksen & Eriksen, 1974) finding reduced microsaccade rates before trials associated with so-called 'cognitive conflict', where the tasks were incongruent with the preceding cues.

As such, we were interested if microsaccades could be used to measure the degree of suppression in oculomotor activity driven by top-down attentional control settings.

Manuscript III: *Microsaccade Rate Activity During the Preparation of Pro- and Antisaccades*

Problem

Microsaccades belong to the category of fixational micromovements and may be crucial for image stability on the retina. Eye movement paradigms typically require fixational control, but this does not eliminate all oculomotor activity. Of all the types of micromovements, the specific functions of microsaccades still remain a mystery. We build on previous studies of microsaccades in the antisaccade task that allows the separation of planning and execution.

Significance of the paper

We propose a new research paradigm to study attentional control settings. We build on previous studies of microsaccades in the antisaccade task using a combination of fixed and mixed pro- and antisaccade blocks. This may provide clues about the way we allocate visual attention and how cognitive load affects fixational eye movements. In addition, the study adds important conclusions to the literature on microsaccades. We also provide a short history of research into microsaccades. While there are certainly high-quality reviews available on the topic, they tend to be quite long, making it difficult to grasp the current state of the field for readers not specialising in microsaccades. Therefore, the paper is valuable as it provides useful context for such readers but at the same time not overwhelming them with detailed information.

Goal

The main question of the study was whether microsaccades can be used as a measure of oculomotor inhibition related to attentional control settings.

Summary

Dynamic Fixation

In most experiments we tend to control participants' gaze using fixation points.

However, these fixational protocols do not eliminate all oculomotor activity. Tiny movements still persist, and one type of such movements are microsaccades. But what exactly are microsaccades? For decades, these tiny movements of the eyes have been a point of debate among vision researchers. First, there have been various opinions regarding a single distinct definition of a microsaccade, starting with basic parameters such as amplitude and velocity (Ditchburn, 1973; Engbert & Kliegl, 2004; Rolfs, 2009; Zuber et al., 1965; and see Collewijn & Kowler, 2008 for a historical overview), and moving on to other characteristics including the typical direction of microsaccades (Rolfs, 2009), whether they may happen in one eye or must be synchronised between the two eyes (Cornsweet, 1956; Engbert & Kliegl, 2003b, 2003a; Krauskopf et al., 1960; Møller et al., 2002; St-Cyr & Fender, 1969; see Collewijn & Kowler, 2008, for a review), the ability to voluntarily control them, and the underlying neural mechanisms (Willeke et al., 2019).

To make things even more complicated, after over 60 years of research we are still not certain what functional purpose these micromovements serve. The variety of opinions here is even more diverse: returning the eyes to the point of fixation after a drift (Cornsweet, 1956), prevention of Troxler fading (Martinez-Conde et al., 2009), fine spatial vision (Kagan & Hafed, 2013; Poletti et al., 2013) and error correction (Ditchburn, 1980). The last point was a large point of contention that became a source of many interesting consequent research studies (Bridgeman & Palca, 1980; Engbert & Kliegl, 2004; Kowler & Steinman, 1980). One important more recent finding that hints at a potential role of microsaccades is that there may be more than one type of microsaccades —fixational and exploratory (Ko et al., 2010; Kowler & Collewijn, 2010), with possible voluntary suppression of the former.

Studies have revealed that microsaccades and saccades share one neurophysiological mechanism that is responsible for their generation (see Hafed et al., 2015, for a review) and inhibition (Rolfs et al., 2008). This is a crucial finding that has allowed us to use microsaccades as one of the ways to study visual cognition through the observation of the dynamics of their properties and rates (Chen et al., 2015; Engbert & Kliegl, 2004; Zuber & Stark, 1966). Since inhibition is one of the characteristics of executive control (Johnston & Everling, 2006; Paneri & Gregoriou, 2017), one approach to studying microsaccades could be within the context of using executive control to suppress the activity of the oculomotor system. The antisaccade paradigm was chosen for studying specifically executive control and the inhibition of automatic responses (Hallett, 1978; Kristjánsson, 2011).

Research Question

We expected oculomotor inhibition to be higher on antisaccade trials than on prosaccade trials due to reflexive saccade suppression (Kristjánsson et al., 2001, 2004). Consequently, this suppression would manifest more strongly before an antisaccade and therefore lead to a lower rate of microsaccades before an

antisaccade than before a prosaccade. We also expected this suppression to be strongest closer to the signal indicating the type of saccade to be made. Our third assumption was that we would observe higher rates in the randomly mixed block of pro- and antisaccade trials due to task-switching. Additionally, we measured microsaccade rates preceding pro- and antisaccades in two modes of eyetracking – monocular and binocular.

Approach to Data Analysis

The pipeline for data analysis was similar in both experiments. We used the microsaccade detection toolbox in R (Engbert et al., 2015). The main analyses for both experiments included calculations of saccade latency, the estimation of microsaccade rates (the average rate per second using generalized linear mixed effects (GLME) models and a running rate analysis using a binned sliding window algorithm) and a cluster permutation analysis of the data. Experiment 2 included additional detailed analyses such as microsaccade rates with relation to following saccade errors and the congruency of microsaccade direction with the direction of the following saccadic eye movement based on saccade error.

Experiment 1: Monocular

Methods. Participants (N=22, 10 male, 1 excluded) contributed to three blocks of trials: 50 fixed prosaccade trials, 50 fixed antisaccade trials and 200 randomly mixed block of pro- and antisaccade trials (100 trials of each saccade type). On all trials fixation lasted for 800 ms, after which a cue appeared for 100 ms, signalling the direction of the saccade to be made (Fig. 13). Up to 3000 ms were allowed for an eye movement before the trial ended. Experiment 1 was carried out monocularly using the SR-Research Eyelink 1000 Plus eyetracker (SR Research Ltd, Osgoode, Ontario, Canada) set to a sampling rate of 1000 Hz to track eye position. Stimuli were presented using the Psychophysics toolbox (Brainard, 1997) for MatLab ® (The Mathworks Inc., Natick, MA).

Figure 13. The General Design Used in Both Microsaccade Experiments
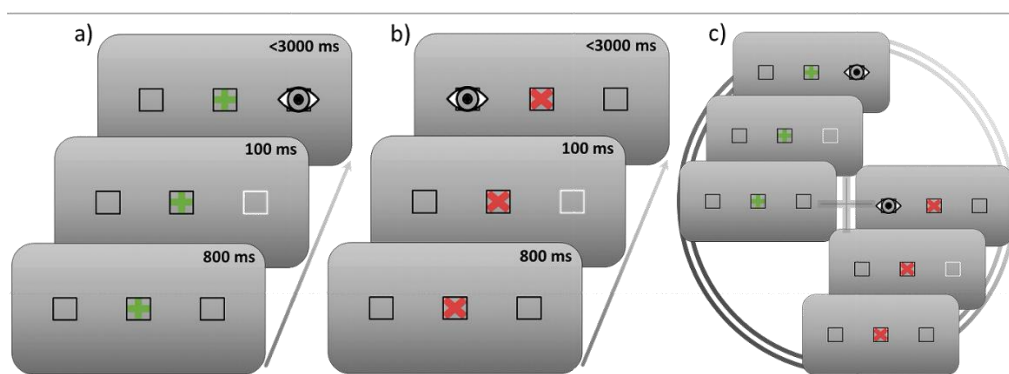
Note. a) Fixed prosaccade trials; b) fixed antisaccade trials; c) randomly mixed block of pro- and antisaccade trials.

Results. We analysed the data and looked at the following variables: the mean latencies of saccadic responses, the correlation between microsaccade velocity and amplitude (the 'main sequence') (Bahill et al., 1975), microsaccade amplitudes, microsaccade direction, average microsaccade rate preceding a saccadic response, microsaccade running rate preceding a saccadic response and meaningful clusters of microsaccade activity preceding a saccadic response.

The mean latency analysis showed a significant effect of saccade type but not block type, and a significant interaction between saccade and block types, with longer latencies on antisaccade trials than on saccade trials. The main sequence correlation analysis of microsaccade velocity and amplitude demonstrated a significant relationship between the two parameters, which confirmed that the microsaccades were detected correctly. For microsaccade amplitudes we found significant effects of saccade type and saccade correctness, with larger microsaccade amplitudes on prosaccade trials than on antisaccade trials and an even stronger effect on fixed incorrect trials. In the mixed condition, however, microsaccade amplitudes were larger before incorrect antisaccades than prosaccades. The analysis of microsaccade direction revealed a horizontal bias for all conditions, which is in line with existing literature (Rolfs, 2009). The analysis of the average rate of microsaccades did not show any significant effects of saccade or block types. In summary, mean microsaccade rates were higher before antisaccades than prosaccades. The running rate analysis did not reveal any clear differences between block types or saccade types. The running rate data were then subjected to a cluster permutation test (Barr, 2020) to identify if there were any distinct clusters of microsaccade activity present. This procedure showed that there were four small significant clusters of microsaccade rate differences between saccade types in the fixed condition. The mixed condition did not contain any such significant clusters of activity.

These results could reflect higher control levels necessary for suppressing reflexive saccades in the antisaccade condition (Coe & Munoz, 2017), but they could also be a consequence of the monocular recording. Most existing studies agree that microsaccades occur in both eyes simultaneously, so monocular data may contain artifacts. So, while the results of experiment 1 were suggestive, we further addressed these issues in experiment 2, this time using binocular eye recording.

Experiment 2: Binocular

Methods. Participants (N=23, 11 male, 2 excluded) contributed to three blocks of trials: 50 fixed prosaccade trials, 50 fixed antisaccade trials and 100 randomly mixed trials (50 of each saccade type) (Fig 13). The methods were identical to the methods of experiment 1 with one crucial exception: eye movement recordings were conducted binocularly at a rate of 1000 Hz.

Results. The same analyses over the same variables as in experiment 1 were conducted on data acquired from the binocular recordings in experiment 2.

The mean latency analysis uncovered significant effects of saccade and block types, and an interaction between them, with longer responses on antisaccade than on prosaccade trials and on mixed than on fixed trials. The interaction was present due to the larger latency difference between saccade types in the fixed condition than in the mixed condition. The main sequence analysis was in line with existing literature on the typical relationship between microsaccade velocity and amplitude. The analysis of microsaccade amplitudes showed a significant effect of saccade correctness, with higher amplitudes and variance on incorrect than on correct trials. Saccade type and correctness interacted, with larger microsaccade amplitudes on incorrect prosaccade trials than on incorrect antisaccade trials. The microsaccade direction analysis revealed a horizontal bias supported by the literature (Rolfs, 2009). We additionally looked at the congruency of the direction of microsaccades and the direction of the following saccade. We observed a significant effect of saccade correctness, an interaction between saccade correctness and congruency and an interaction between congruency, saccade correctness, saccade type and block type. On correct fixed antisaccade trials incongruent microsaccades were more frequent than congruent ones, while on correct fixed prosaccade trials congruent microsaccades were more frequent than incongruent directions. On mixed trials, the same pattern for correct antisaccades was observed, while the frequency of incongruent microsaccades was higher than the frequency of congruent microsaccades on correct prosaccade trials. The average rate of microsaccades was significantly affected by saccade type, where we found less microsaccades preceding antisaccades than prosaccades. This was mostly due to the difference between the rates in the mixed condition. We additionally looked at the rates on correct and incorrect trials and found an interaction between saccade type and saccade correctness with higher rates preceding incorrect antisaccades. The running rate and cluster permutation analyses disclosed four significant clusters of rate differences between saccade types in the fixed condition and one such significant cluster in the mixed condition.

Discussion

Overall, we observed significant differences in experiment 1 in both latency and microsaccade rate, though we suspect the results may not be entirely reliable due to monocular recording. In experiment 2 microsaccade rates prior to antisaccades were reduced, consistent with the hypothesis that microsaccade rates are influenced by which saccade type is being prepared. Specifically, suppression of the reflexive oculomotor system to prevent incorrect express saccades to the peripheral stimulus (Coe & Munoz, 2017) may have reduced overall microsaccade rates in the pre-stimulus period.

Interestingly, our results suggest that participants may have better control over microsaccadic eye movements than previously suggested (Buonocore et al., 2017). Our design may have encouraged participants to suppress most microsaccades prior to saccade generation since we enforced fixation with an error signal and a reminder to fixate and started each trial with drift correction. Both manipulations may have resulted in stronger ocular inhibition than otherwise. Nevertheless, in both experiments the suppression during fixed antisaccade trials was released more strongly right before the cue, while this release was not as strong in the mixed condition. This probably reflects the load associated with switching between randomly appearing pro- and antisaccade cues. No such release effect occurred for prosaccades in either the fixed or mixed conditions, probably because less suppression is required.

The cluster permutation test additionally showed larger differences in microsaccade activity on fixed trials in both experiments. On mixed trials, however, there were either no significant clusters of rate differences or they were very small. Since the saccade was executed immediately following the cue and the pre-cue interval was 800 ms, there simply may not have been enough time to observe larger clusters. Participants may have been anticipating the cue flash and combined with saccade preparation, they may have suppressed all oculomotor activity (including microsaccades, Hermens et al., 2010; see also evidence from monkey studies: Hafed et al., 2011)

Conclusion

Our results suggest that microsaccades can be used to measure oculomotor suppression in an antisaccade task. This may be important for current theories that suppose that suppression of the reflexive oculomotor system may gate other attentional mechanisms. For example, Redden and colleagues (2021) have suggested a model of IOR that changes form (input or output) based on whether the reflexive oculomotor system is in a state of inhibition. Further, our cluster permutation analyses may set a new direction for investigating the role of these micromovements and may provide a framework for their study.

General discussion

We performed three major studies and two supplementary studies that in general aim to advance our understanding of attentional control processes and their influence on cognitive performance. They contribute to the theoretical foundations of visual attention research by investigating different aspects of attentional control settings and their implications. Collectively, they propose a theoretical framework to investigate the role of attentional control settings in visual performance on two levels: executive functioning and visual perception. Taken separately, each of these studies provide theoretical and methodological contributions to the existing literature in the field of computational cognitive neuroscience of visual attention and perception.

Namely, Study I was comprised of a large review of existing models (Krasovskaya & MacInnes, 2019) classified in accordance with five proposed major aspects that a computational cognitive neuroscience model should strive to incorporate: neurobiological and cognitive plausibility, computational implementations, generative capabilities, the spatial and temporal characteristics of visual attention. Two supplementary studies (Manuscript Ia & Manuscript Ib) proposed a practical implementation of the theoretical framework and existing gaps suggested in Manuscript I (Krasovskaya & MacInnes, 2019). The focus of one of these studies (Krasovskaya et al., 2021) was a conceptual deep-neural network implementation of the classical Itti and Koch (2001) model that attempted to incorporate all the vision modelling aspects of computational cognitive neuroscience outlined in Manuscript I. This model was modular and provided a solid foundation for future implementations that could work flexibly with other deep-learning architectures. A second model was built for the study described in Manuscript Ib (Merzon et al., 2020). Using this model, we identified the limitations of the leaky integrate-and-fire approach with default parameters for modelling the temporal component of visual attention. The results of these projects, especially the elements did not work well in terms of temporal performance, were taken into consideration in the proposal of a new sophisticated computational architecture for temporal predictions - the Spatial Leaky Competing Accumulator (SLCA) model (Zemliak & MacInnes, 2022).

The theoretical framework proposed in Study I brought into light the realisation that many models of vision tend to build their predictions based on bottom-up saliency, ignoring the more intricate workings of other attentional control mechanisms. Models tend to learn based on the input they receive. To potentially improve spatial model predictions, we gathered human data in two large eyetracking studies (Studies II and III) that could be used separately as well as model input. In these studies, oculomotor behaviour was observed during the performance of

complex visual tasks that affected the attentional control settings in participants. Study II focussed on the control of observers' attentional set by enforcing a constant fixed area of visual input during visual search. The gaze-contingent display provided a retinotopic setup where participants had to make overt eye movements to find the target before the trial ended. The important findings of this study are that (1) the FFV is adaptable to task instructions; and (2) the optimal size of the FFV may vary for different objectives. This implies that it would be worthwhile to observe how different task and cognitive load variations, combined with a larger set of aperture sizes, would affect performance. These data reflect the top-down and context-driven influences of attentional control settings on oculomotor behaviours. Besides the valuable experimental methodological approaches and behavioural findings, crucial eyetracking data were acquired during this study. These data contain participants' gaze locations at very specific time points, making them an essential source of input for the retinotopic aspect of spatial model predictions.

In Study III, the provided instructions forced participants to adapt to the task and either release or suppress their reflexive responses towards a peripheral cue by utilising their executive control. The difference in cognitive load and task-switching demands influenced the degree of suppression of the oculomotor system, as seen in microsaccade rate dynamics. Combined with the results of previous studies using a similar paradigm, our results may help uncover the role(s) of microsaccadic eye movements in various tasks. Using these data for modelling purposes will help train models that more precisely simulate human behaviour and potentially help us spot patterns that are not yet visible given the small number of studies in this area. These approaches could be a useful source of insight into the modulation of visual attention via changes in task difficulty and attentional set.

One of the attentional control mechanisms that is involved in the process of selective attention is executive control. Executive control is a major player in the presented studies. During the antisaccade task used in Study III, executive control is crucial for the suppression of the reflexive saccade. It is also an important factor of attentional control when searching for a target stimulus among similar distractors (Study II). It influences the amount of visual information entering the system for consequent processing by suppressing the surrounding irrelevant signals as determined by attentional control mechanisms. As such, it is important to take these factors into account when building computational models if our goal is to study human visual attention. We can suppress lower-level bottom-up signals when overwritten by other attentional control factors, thus such models should focus more on the interplay of these factors under various conditions and their levels of priority.

Cumulatively, the obtained results and data are beneficial for the field of visual attention and perception. They provide a behavioural framework that, taken either alone or in combination with neurophysiological data from primate and human

studies, makes it possible to observe how visual decision-making is affected by attentional control settings - namely, how exactly areas in the visual field are prioritized based on the context of the task and the intentions of the observer. Computational modelling makes it possible to combine the behavioural and neurophysiological data into a single system and observe the way different module settings interact within this system, as well as how their interactions influence the entire system. The important part is that it excludes fatal consequences that are often a possibility in invasive studies, and it does not require volunteers for such procedures.

To summarise, the impact of the current project is manifold. From the theoretical perspective, the impact of the project lies in its potential to deepen our understanding of attentional control processes, refine existing theories, and pave the way for future advancements in the field of visual attention. Methodologically, it proposes several sophisticated experimental frameworks for studying selective attention and its underlying cognitive mechanisms via various oculomotor proxy-signals. From the computational modelling side, it adds value to the field by suggesting a structure for future architectures whose objective is to study visual cognition. Finally, it paves the way for the next exciting stage of the journey into visual cognition.

Beyond the End: On the Way to the Next Fixation Location

> *It's not what's on this island, but what's beneath it, that interests me.*
>
> — Jules Verne, *Journey to the Centre of the Earth*

Selective attention drives our perception, but can the study of attentional control settings via eye movements cover existing gaps in the cognitive neuroscience modelling of vision? This dissertation is but the beginning of a larger journey towards the territory of the precise cognitive mechanisms that regulate and sustain visual cognition.

Ahead of this lies one of the next visible destinations of the larger project - a model of attentional priority that uses eyetracking and neuroimaging data to help us learn how humans build a spatiotopic understanding of the world based on the interplay of attentional settings and retinotopic visual input.

Consider the following questions: *How do you judge the locations of objects in the surrounding environment? Would you provide a latitude and longitude for the location of a cup of aromatic coffee or soothing tea sitting on your desk while you read this manuscript? Perhaps it is located on the left or right side of the computer screen? Or would you rather see it as being on your left or right side when you face the computer screen? What if you turn away from the screen to talk to somebody behind you? Would it still be on the same side in relation to your body? Would the latitude and longitude location of the cup change in this case?*

When we take the longitude and latitude of the cup in the example above, we are talking about the spatiotopic coordinates of the cup which are fixed in relation to the external environment. This means that irrespective of our location in the world, the position of the cup will stay the same. However, when we use our eyes or head as a reference point, the position of the cup will change based on our position in space. If the cup is on our left side when we sit at our desk and somebody comes up behind us, forcing us to turn around to face them, the position of the cup would change relative to our new position – it will probably be to our right and slightly posterior. At the same time, its spatiotopic location would not change.

Both frames of reference are important for visual processing. The idea of a spatiotopic reference frame facilitates our understanding of how we process visual data while changes in gaze direction happen. Specifically, how we maintain a stable image of our surroundings when our eyes are never entirely still. Retinotopic

coordinates are crucial for visual processing as the brain is retinotopically organised (Golomb et al., 2008; Golomb & Kanwisher, 2012), while the representation of the visual image rapidly changes with each saccade. Receptive fields of neurons are responsive to particular locations within the visual field (Hirsch & Martinez, 2009). This retinotopic organisation is maintained throughout the entire stream of visual signal processing, from the retina through the optic nerve into the associated cortical areas. This retinotopic organisation and sensitivity to gaze position has been demonstrated multiple times (Machner et al., 2020; Paraskevoudi & Pezaris, 2021; Wandell et al., 2007). But how exactly and at which level of visual processing does the translation or integration of the different coordinate systems happen? Given all these complexities, how does the brain know where to look and how does this happen so quickly? One of the brain regions responsible for saccade generation, as mentioned before, is the SC. Its superficial layers receive input directly from the retina and from the visual cortex via the retinotectal and the retino-geniculo-cortical pathways, respectively (Munoz & Everling, 2004). The deeper intermediate layers of the SC are involved in a much more complex network. They receive excitatory and inhibitory inputs from a variety of other brain regions such as the LIP, dlPFC, FEF, SEF and basal ganglia, as well as regions responsible for synapses connected with other sensory modalities. All these signals are integrated and coordinated to decide whether an eye movement should be made and in which direction based on the retinotopic map of 2D space in the SC. In the case of a positive decision, neurons on this map are activated and excitatory signals are sent to the cerebellum and/or the reticular formation. Consequently, a saccadic eye movement is generated in the direction matching the retinal vector defined by the location of the activated neuron on the retinotopic map. Despite having the separate components and their functions mapped out, we still have a long way to go in terms of understanding the details of how all these mechanisms interact. Perhaps computational modelling may be the Nautilus that carries us deeper into this unknown area?

Conclusions

To summarise the work accomplished within the doctoral project, we have (1) analysed the state-of-the-art in the field of visual attention modelling with respect to computational cognitive neuroscience (Study I, Manuscript I); (2) gained a deeper insight into some of the cognitive and attentional mechanisms in visual attention (Studies II and III); and (3) introduced ways to improve existing computational models of vision using the results of the studies (Study I, Manuscripts I, Ia and Ib; General discussion).

The logical progression following this project would involve integrating the theoretical and experimental findings into an actual practical computational model of visual attention and perception. By incorporating the insights gained from the research, we can develop a computational model to simulate and understand how attention operates in visual processing. This model will aim to replicate the mechanisms and behaviour of human visual attention, providing a valuable tool for studying and predicting attentional processes within various contexts and tasks while incorporating the retinotopic processes the visual system uses to navigate the external world.

References

Abernethy, B., & Russell, D. G. (1987). Expert-Novice Differences in an Applied Selective Attention Task. *Journal of Sport and Exercise Psychology*, *9*(4), 326–345. https://doi.org/10.1123/jsp.9.4.326

Aboudib, A., Gripon, V., & Coppin, G. (2016). A Biologically Inspired Framework for Visual Information Processing and an Application on Modeling Bottom-Up Visual Attention. *Cognitive Computation*, *8*(6), 1007–1026. https://doi.org/10.1007/s12559-016-9430-8

Adeli, H., Vitu, F., & Zelinsky, G. J. (2017). A Model of the Superior Colliculus Predicts Fixation Locations during Scene Viewing and Visual Search. *The Journal of Neuroscience*, *37*(6), 1453–1467. https://doi.org/10.1523/JNEUROSCI.0825-16.2016

Amadieu, F., Mariné, C., & Laimay, C. (2011). The attention-guiding effect and cognitive load in the comprehension of animations. *Computers in Human Behavior*, *27*(1), 36–40. https://doi.org/10.1016/j.chb.2010.05.009

Anderson, S. J., Mullen, K. T., & Hess, R. F. (1991). Human peripheral spatial resolution for achromatic and chromatic stimuli: Limits imposed by optical and retinal factors. *The Journal of Physiology*, *442*, 47–64.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Backer, G., Mertsching, B., & Bollmann, M. (2001). Data- and model-driven gaze control for an active-vision system. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(12), 1415–1429. https://doi.org/10.1109/34.977565

Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, *24*(3–4), 191–204. https://doi.org/10.1016/0025-5564(75)90075-9

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of The Optical Society of America A-Optics Image Science and Vision*. https://doi.org/10.1364/JOSAA.5.002210

Barr, D. (2020). *Cluster-Permutation Tests for Time-Series Data from Factorial Experiments* (0.0.0.9200) [R]. https://github.com/dalejbarr/clusterperm.git

Belopolsky, A. V., & Theeuwes, J. (2010). No capture outside the attentional window. *Vision Research*, *50*(23), 2543–2550. https://doi.org/10.1016/j.visres.2010.08.023

Bengio, Y. (2009). Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning*, *2*(1), 1–127. https://doi.org/10.1561/2200000006

Bertels, J., Kolinsky, R., & Morais, J. (2010). Emotional valence of spoken words influences the spatial orienting of attention. *Acta Psychologica*, *134*(3), 264–278. https://doi.org/10.1016/j.actpsy.2010.02.008

Bisley, J. W., & Mirpour, K. (2019). The neural instantiation of a priority map. *Current Opinion in Psychology*, *29*, 108–112. https://doi.org/10.1016/j.copsyc.2019.01.002

Björk, A., & Kugelberg, E. (1953). The electrical activity of the muscles of the eye and eyelids in various positions and during movement. *Electroencephalography and Clinical Neurophysiology*, *5*(4), 595–602. https://doi.org/10.1016/0013-4694(53)90037-6

Böckler, A., van der Wel, R. P. R. D., & Welsh, T. N. (2014). Catching Eyes: Effects of Social and Nonsocial Cues on Attention Capture. *Psychological Science*, *25*(3), 720–727. https://doi.org/10.1177/0956797613516147

Bourgeois, A., Sterpenich, V., Iannotti, G. R., & Vuilleumier, P. (2022). Reward-driven modulation of spatial attention in the human frontal eye-field. *NeuroImage*, *247*, 118846. https://doi.org/10.1016/j.neuroimage.2021.118846

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Bridgeman, B., & Palca, J. (1980). The role of microsaccades in high acuity observational tasks. *Vision Research*, *20*(9), 813–817. https://doi.org/10.1016/0042-6989(80)90013-9

Broadbent, D. E. (2013). *Perception and Communication*. Elsevier.

Buonocore, A., Chen, C.-Y., Tian, X., Idrees, S., Münch, T. A., & Hafed, Z. M. (2017). Alteration of the microsaccadic velocity-amplitude main sequence relationship after visual transients: Implications for models of saccade control. *Journal of Neurophysiology*, *117*(5), 1894–1910. https://doi.org/10.1152/jn.00811.2016

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science (New York, N.Y.)*, *315*(5820), 1860–1862. https://doi.org/10.1126/science.1138071

Buschman, T. J., & Miller, E. K. (2010). Shifting the spotlight of attention: Evidence for discrete computations in cognition. *Frontiers in Human Neuroscience*, *4*, 194. https://doi.org/10.3389/fnhum.2010.00194

Buswell, G. T. (1935). *How people look at pictures: A study of the psychology and perception in art* (p. 198). Univ. Chicago Press.

Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., & Durand, F. (2019). What Do Different Evaluation Metrics Tell Us About Saliency Models? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(3), 740–757. https://doi.org/10.1109/TPAMI.2018.2815601

Cave, K. R. (1999). The FeatureGate model of visual selection. *Psychological Research*, *62*(2), 182–194. https://doi.org/10.1007/s004260050050

Chan, H. F., Bodiuzzman, S. M., & Torgler, B. (2020). The power of social cues in the battle for attention: Evidence from an online platform for scholarly commentary. *Journal of Informetrics*, *14*(4), 101077. https://doi.org/10.1016/j.joi.2020.101077

Chen, C.-Y., Ignashchenkova, A., Thier, P., & Hafed, Z. M. (2015). Neuronal Response Gain Enhancement prior to Microsaccades. *Current Biology*, *25*(16), 2065–2074. https://doi.org/10.1016/j.cub.2015.06.022

Chetverikov, A., Hansmann-Roth, S., Tanrıkulu, Ö. D., & Kristjánsson, Á. (2020). Feature Distribution Learning (FDL): A New Method for Studying Visual Ensembles Perception with Priming of Attention Shifts. In S. Pollmann (Ed.), *Spatial Learning and Attention Guidance* (pp. 37–57). Springer US. https://doi.org/10.1007/7657_2019_20

Chetverikov, A., Kuvaldina, M., MacInnes, W. J., Jóhannesson, Ó. I., & Kristjánsson, Á. (2018). Implicit processing during change blindness revealed with mouse-contingent and gaze-contingent displays. *Attention, Perception, & Psychophysics*, *80*(4), 844–859. https://doi.org/10.3758/s13414-017-1468-5

Coe, B. C., & Munoz, D. P. (2017). Mechanisms of saccade suppression revealed in the anti-saccade task. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160192. https://doi.org/10.1098/rstb.2016.0192

Collewijn, H., & Kowler, E. (2008). The significance of microsaccades for vision and oculomotor control. *Journal of Vision*, *8*(14), 20. https://doi.org/10.1167/8.14.20

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., Van Essen, D. C., & Shulman, G. L. (1998). A Common Network of Functional Areas for Attention and Eye Movements. *Neuron*, *21*(4), 761–773. https://doi.org/10.1016/S0896-6273(00)80593-0

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience*, *13*(3), 1202–1226. https://doi.org/10.1523/JNEUROSCI.13-03-01202.1993

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), Article 3. https://doi.org/10.1038/nrn755

Corbetta, M., Shulman, G. L., Miezin, F. M., & Petersen, S. E. (1995). Superior Parietal Cortex Activation During Spatial Attention Shifts and Visual Feature

Conjunction. *Science*, *270*(5237), 802–805. https://doi.org/10.1126/science.270.5237.802

Cornsweet, T. N. (1956). Determination of the Stimuli for Involuntary Drifts and Saccadic Eye Movements*. *JOSA*, *46*(11), 987–993. https://doi.org/10.1364/JOSA.46.000987

Cullen, K. E., & Van Horn, M. R. (2011). The neural control of fast vs. Slow vergence eye movements. *European Journal of Neuroscience*, *33*(11), 2147–2154. https://doi.org/10.1111/j.1460-9568.2011.07692.x

Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *The Journal of Comparative Neurology*, *300*(1), 5–25. https://doi.org/10.1002/cne.903000103

Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *The Journal of Comparative Neurology*, *292*(4), 497–523. https://doi.org/10.1002/cne.902920402

Curcio, C. A., Sloan, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of cones in human and monkey retina: Individual variability and radial asymmetry. *Science (New York, N.Y.)*, *236*(4801), 579–582. https://doi.org/10.1126/science.3576186

Dalmaso, M., Castelli, L., & Galfano, G. (2019). Anticipation of cognitive conflict is reflected in microsaccades: Evidence from a cued-flanker task. *Journal of Eye Movement Research*, *12*(6), Article 6. https://doi.org/10.16910/jemr.12.6.3

Dalmaso, M., Castelli, L., & Galfano, G. (2020). Microsaccadic rate and pupil size dynamics in pro-/anti-saccade preparation: The impact of intermixed vs.

blocked trial administration. *Psychological Research*, *84*(5), 1320–1332. https://doi.org/10.1007/s00426-018-01141-7

Deco, G., & Rolls, E. T. (2004). A Neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research*, *44*(6), 621–642. https://doi.org/10.1016/j.visres.2003.09.037

Delabarre, E. B. (1898). A Method of Recording Eye-Movements. *The American Journal of Psychology*, *9*(4), 572–574. https://doi.org/10.2307/1412191

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

Ditchburn, R. W. (1973). *Eye-movements and visual perception* (pp. xv, 421). Clarendon.

Ditchburn, R. W. (1980). The function of small saccades. *Vision Research*, *20*(3), 271–272. https://doi.org/10.1016/0042-6989(80)90112-1

Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a Stabilized Retinal Image. *Nature*, *170*(4314), Article 4314. https://doi.org/10.1038/170036a0

Dodds, C. M., Morein-Zamir, S., & Robbins, T. W. (2011). Dissociating Inhibition, Attention, and Response Control in the Frontoparietal Network Using Functional Magnetic Resonance Imaging. *Cerebral Cortex*, *21*(5), 1155–1165. https://doi.org/10.1093/cercor/bhq187

Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review*, *8*(2), 145–157. https://doi.org/10.1037/h0076100

Donner, T. H., Kettermann, A., Diesch, E., Ostendorf, F., Villringer, A., & Brandt, S. A. (2002). Visual Feature and Conjunction Searches of Equal Difficulty

Engage Only Partially Overlapping Frontoparietal Networks. *NeuroImage*, *15*(1), 16–25. https://doi.org/10.1006/nimg.2001.0951

Donner, T., Kettermann, A., Diesch, E., Ostendorf, F., Villringer, A., & Brandt, S. A. (2000). Involvement of the human frontal eye field and multiple parietal areas in covert visual selection during conjunction search. *European Journal of Neuroscience*, *12*(9), 3407–3414. https://doi.org/10.1046/j.1460-9568.2000.00223.x

Draper, B. A., & Lionelle, A. (2005). Evaluation of selective attention under similarity transformations. *Computer Vision and Image Understanding*, *100*(1), 152–171. https://doi.org/10.1016/j.cviu.2004.08.006

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 578–588. https://doi.org/10.1037/0096-1523.18.2.578

Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, *7*(2), 255–261. https://doi.org/10.1016/S0959-4388(97)80014-1

Engbert, R., & Kliegl, R. (2003a). Binocular Coordination in Microsaccades. In *The Mind's Eye* (pp. 103–117). Elsevier. https://doi.org/10.1016/B978-044451020-4/50007-4

Engbert, R., & Kliegl, R. (2003b). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*(9), 1035–1045. https://doi.org/10.1016/S0042-6989(03)00084-1

Engbert, R., & Kliegl, R. (2004). Microsaccades Keep the Eyes' Balance During Fixation. *Psychological Science*, *15*(6), 431–431. https://doi.org/10.1111/j.0956-7976.2004.00697.x

Engbert, R., Sinn, P., Mergenthaler, K., & Trukenbrod, H. (2015). *Microsaccade Toolbox for R*. http://read.psych.uni-potsdam.de/index.php?option=com_content&view=article&id=140:engbert-et-al-2015-microsaccade-toolbox-for-r&catid=26:publications&Itemid=34

Engel, F. L. (1971). Visual conspicuity, directed attention and retinal locus. *Vision Research*, *11*(6), 563–575. https://doi.org/10.1016/0042-6989(71)90077-0

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*(4), 225–240. https://doi.org/10.3758/BF03211502

Failing, M., & Theeuwes, J. (2018). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin & Review*, *25*(2), 514–538. https://doi.org/10.3758/s13423-017-1380-y

Feldmann-Wüstefeld, T., Uengoer, M., & Schubö, A. (2015). You see what you have learned. Evidence for an interrelation of associative learning and visual selective attention. *Psychophysiology*, *52*(11), 1483–1497. https://doi.org/10.1111/psyp.12514

Fernando, T., Denman, S., Sridharan, S., & Fookes, C. (2018). Task Specific Visual Saliency Prediction with Memory Augmented Conditional Generative

Adversarial Networks. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1539–1548. https://doi.org/10.1109/WACV.2018.00172

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, *22*(4), 661–721. https://doi.org/10.1017/S0140525X99002150

Finucane, A. M. (2011). The effect of fear and anger on selective attention. *Emotion*, *11*(4), 970–974. https://doi.org/10.1037/a0022574

Fischer, M., Moscovitch, M., & Alain, C. (2021). A systematic review and meta-analysis of memory-guided attention: Frontal and parietal activation suggests involvement of fronto-parietal networks. *WIREs Cognitive Science*, *12*(1), e1546. https://doi.org/10.1002/wcs.1546

Gaspar, J. G., Ward, N., Neider, M. B., Crowell, J. A., Carbonari, R., Kaczmarski, H., Ringer, R. V., Johnson, A. P., Kramer, A. F., & Loschky, L. C. (2016). Measuring the Useful Field of View During Simulated Driving With Gaze-Contingent Displays: *Human Factors*. https://doi.org/10.1177/0018720816642092

Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press. https://books.google.is/books/about/Neuronal_Dynamics.html?id=D4j2AwA AQBAJ&redir_esc=y

Golomb, J. D., Chun, M. M., & Mazer, J. A. (2008). The Native Coordinate System of Spatial Attention Is Retinotopic. *Journal of Neuroscience*, *28*(42), 10654–10662. https://doi.org/10.1523/JNEUROSCI.2525-08.2008

Golomb, J. D., & Kanwisher, N. (2012). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex (New York, N.Y.: 1991)*, *22*(12), 2794–2810. https://doi.org/10.1093/cercor/bhr357

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Grünert, U., & Martin, P. (2020). Cell types and cell circuits in primate retina. *Progress in Retinal and Eye Research*, *78*, 100844. https://doi.org/10.1016/j.preteyeres.2020.100844

Guest, O., & Martin, A. E. (2021). How Computational Modeling Can Force Theory Building in Psychological Science. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Gunseli, E., Olivers, C. N. L., & Meeter, M. (2016). Task-irrelevant memories rapidly gain attentional control with learning. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 354–362. https://doi.org/10.1037/xhp0000134

Hafed, Z. M., Chen, C.-Y., & Tian, X. (2015). Vision, Perception, and Attention through the Lens of Microsaccades: Mechanisms and Implications. *Frontiers in Systems Neuroscience*, *9*, 167. https://doi.org/10.3389/fnsys.2015.00167

Hafed, Z. M., Lovejoy, L. P., & Krauzlis, R. J. (2011). Modulation of microsaccades in monkey during a covert visual attention task. *Journal of Neuroscience*,

31(43), 15219–15230. Scopus. https://doi.org/10.1523/JNEUROSCI.3106-
11.2011

Haji-Abolhassani, A., & Clark, J. J. (2014). An inverse Yarbus process: Predicting
observers' task from eye movement patterns. *Vision Research*, *103*, 127–
142. https://doi.org/10.1016/j.visres.2014.08.014

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by
instructions. *Vision Research*, *18*(10), 1279–1296.
https://doi.org/10.1016/0042-6989(78)90218-3

Hamker, F. H. (2002). *How does the ventral pathway contribute to spatial attention
and the planning of eye movements*.

Hamker, F. H. (2006). Modeling feature-based attention as an active top-down
inference process. *Biosystems*, *86*(1), 91–99.
https://doi.org/10.1016/j.biosystems.2006.03.010

Harel, J., Koch, C., & Perona, P. (2006). Graph-Based Visual Saliency. *Advances in
Neural Information Processing Systems*, *19*.
https://proceedings.neurips.cc/paper_files/paper/2006/hash/4db0f8b0fc895
da263fd77fc8aecabe4-Abstract.html

Hartikainen, K. M. (2021). Emotion-Attention Interaction in the Right Hemisphere.
Brain Sciences, *11*(8), Article 8. https://doi.org/10.3390/brainsci11081006

Heidlmayr, K., Kihlstedt, M., & Isel, F. (2020). A review on the
electroencephalography markers of Stroop executive control processes.
Brain and Cognition, *146*, 105637.
https://doi.org/10.1016/j.bandc.2020.105637

Henderson, J. M., Shinkareva, S. V., Wang, J., Luke, S. G., & Olejarczyk, J. (2013). Predicting Cognitive State from Eye Movements. *PLOS ONE*, *8*(5), e64937. https://doi.org/10.1371/journal.pone.0064937

Hendrickson, A. (2005). Organization of the Adult Primate Fovea. In P. L. Penfold & J. M. Provis (Eds.), *Macular Degeneration* (pp. 1–23). Springer. https://doi.org/10.1007/3-540-26977-0_1

Hering, E. (1879). *Über Muskelgeräusche des Auges [Concerning muscle noise of the eye]. 79*(3), 137–154.

Hermens, F., Zanker, J. M., & Walker, R. (2010). Microsaccades and preparatory set: A comparison between delayed and immediate, exogenous and endogenous pro- and anti-saccades. *Experimental Brain Research*, *201*(3), 489–498. https://doi.org/10.1007/s00221-009-2061-5

Hershler, O., & Hochstein, S. (2009). The importance of being expert: Top-down attentional control in visual search with photographs. *Attention, Perception, & Psychophysics*, *71*(7), 1478–1486. https://doi.org/10.3758/APP.71.7.1478

Hirsch, J. A., & Martinez, L. M. (2009). Visual Cortical and Subcortical Receptive Fields. In M. D. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedia of Neuroscience* (pp. 4307–4310). Springer. https://doi.org/10.1007/978-3-540-29678-2_6348

Hulleman, J., & Olivers, C. N. L. (2017a). On the brink: The demise of the item in visual search moves closer. *Behavioral and Brain Sciences*, *40*, e163. https://doi.org/10.1017/S0140525X16000364

Hulleman, J., & Olivers, C. N. L. (2017b). The impending demise of the item in visual search. *Behavioral and Brain Sciences*, *40*, e132. https://doi.org/10.1017/S0140525X15002794

Humphreys, G. W., & Muller, H. J. (1993). SEarch via Recursive Rejection (SERR): A Connectionist Model of Visual Search. *Cognitive Psychology*, *25*(1), 43–110. https://doi.org/10.1006/cogp.1993.1002

Hunziker, H. W. (2006). *Im Auge des Lesers: Vom Buchstabieren zur Lesefreude : foveale und periphere Wahrnehmung*. Hunziker Hans-Werner.

Irons, J. L., Folk, C. L., & Remington, R. W. (2012). All set! Evidence of simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 758–775. https://doi.org/10.1037/a0026578

Irons, J. L., & Leber, A. B. (2016). Choosing attentional control settings in a dynamically changing environment. *Attention, Perception, & Psychophysics*, *78*(7), 2031–2048. https://doi.org/10.3758/s13414-016-1125-4

Ischebeck, A., Hiebel, H., Miller, J., Höfler, M., Gilchrist, I. D., & Körner, C. (2021). Target processing in overt serial visual search involves the dorsal attention network: A fixation-based event-related fMRI study. *Neuropsychologia*, *153*, 107763. https://doi.org/10.1016/j.neuropsychologia.2021.107763

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*. https://doi.org/10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*. https://doi.org/10.1038/35058500

Jacobs, A. M. (1986). Eye-movement control in visual search: How direct is visual span control? *Perception & Psychophysics*, *39*(1), 47–58. https://doi.org/10.3758/BF03207583

Janzen, G., & van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, *7*(6), Article 6. https://doi.org/10.1038/nn1257

Japkowicz, N., Hanson, S. J., & Gluck, M. A. (2000). Nonlinear Autoassociation Is Not Equivalent to PCA. *Neural Computation*, *12*(3), 531–545. https://doi.org/10.1162/089976600300015691

Jefferies, L. N., Smilek, D., Eich, E., & Enns, J. T. (2008). Emotional Valence and Arousal Interact in Attentional Control. *Psychological Science*, *19*(3), 290–295. https://doi.org/10.1111/j.1467-9280.2008.02082.x

Jia, S., & Bruce, N. D. B. (2020). EML-NET: An Expandable Multi-Layer NETwork for saliency prediction. *Image and Vision Computing*, *95*, 103887. https://doi.org/10.1016/j.imavis.2020.103887

Jóhannesson, Ó. I., Haraldsson, H. M., & Kristjánsson, Á. (2013). Modulation of antisaccade costs through manipulation of target-location probability: Only under decisional uncertainty. *Vision Research*, *93*, 62–73. https://doi.org/10.1016/j.visres.2013.10.010

Jóhannesson, Ó. I., Tagu, J., & Kristjánsson, Á. (2018). Asymmetries of the visual system and their influence on visual performance and oculomotor dynamics. *European Journal of Neuroscience*, *48*(11), 3426–3445. https://doi.org/10.1111/ejn.14225

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–Hand Coordination in Object Manipulation. *Journal of Neuroscience*, *21*(17), 6917–6932. https://doi.org/10.1523/JNEUROSCI.21-17-06917.2001

Johnston, K., & Everling, S. (2006). Monkey Dorsolateral Prefrontal Cortex Sends Task-Selective Signals Directly to the Superior Colliculus. *Journal of Neuroscience*, *26*(48), 12471–12478. https://doi.org/10.1523/JNEUROSCI.4101-06.2006

Judd, T., Durand, F., & Torralba, A. (2012). *A Benchmark of Computational Models of Saliency to Predict Human Fixations*. https://dspace.mit.edu/handle/1721.1/68590

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). Learning to predict where humans look. *2009 IEEE 12th International Conference on Computer Vision*, 2106–2113. https://doi.org/10.1109/ICCV.2009.5459462

Kagan, I., & Hafed, Z. M. (2013). Active Vision: Microsaccades Direct the Eye to Where It Matters Most. *Current Biology*, *23*(17), R712–R714. https://doi.org/10.1016/j.cub.2013.07.038

Kanan, C., Tong, M. H., Zhang, L., & Cottrell, G. W. (2009). SUN: Top-down saliency using natural statistics. *Visual Cognition*, *17*(6–7), 979–1003. https://doi.org/10.1080/13506280902771138

Katsuki, F., & Constantinidis, C. (2014). Bottom-Up and Top-Down Attention: Different Processes and Overlapping Neural Systems. *The Neuroscientist*, *20*(5), 509–521. https://doi.org/10.1177/1073858413514136

Kaye, D. B., & Ruskin, E. M. (1990). The Development of Attentional Control Mechanisms. In J. T. Enns (Ed.), *Advances in Psychology* (Vol. 69, pp. 227–244). North-Holland. https://doi.org/10.1016/S0166-4115(08)60459-8

Killian, N. J., Potter, S. M., & Buffalo, E. A. (2015). Saccade direction encoding in the primate entorhinal cortex during visual exploration. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(51), 15743–15748. https://doi.org/10.1073/pnas.1417059112

Kimmel, D. L., Mammo, D., & Newsome, W. T. (2012). Tracking the eye non-invasively: Simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Frontiers in Behavioral Neuroscience*, *6*, 49. https://doi.org/10.3389/fnbeh.2012.00049

Kimmig, H., Greenlee, M., Gondan, M., Schira, M., Kassubek, J., & Mergner, T. (2001). Relationship between saccadic eye movements and cortical activity as measured by fMRI: Quantitative and qualitative aspects. *Experimental Brain Research*, *141*(2), 184–194. https://doi.org/10.1007/s002210100844

Ko, H., Poletti, M., & Rucci, M. (2010). Microsaccades precisely relocate gaze in a high visual acuity task. *Nature Neuroscience*, *13*(12), 1549–1553. https://doi.org/10.1038/nn.2663

Koch, K., McLean, J., Segev, R., Freed, M. A., Berry, M. J., Balasubramanian, V., & Sterling, P. (2006). How Much the Eye Tells the Brain. *Current Biology : CB*, *16*(14), 1428–1434. https://doi.org/10.1016/j.cub.2006.05.056

Koene, A. R., & Zhaoping, L. (2007). Feature-specific interactions in salience from combined feature contrasts: Evidence for a bottom–up saliency map in V1. *Journal of Vision*, *7*(7), 6. https://doi.org/10.1167/7.7.6

Kowler, E. (2011). Eye movements: The past 25years. *Vision Research*, *51*(13), 1457–1483. https://doi.org/10.1016/j.visres.2010.12.014

Kowler, E., & Collewijn, H. (2010). The eye on the needle. *Nature Neuroscience*, *13*(12), 1443–1444. https://doi.org/10.1038/nn1210-1443

Kowler, E., & Steinman, R. M. (1980). Small saccades serve no useful purpose: Reply to a letter by R. W. Ditchburn. *Vision Research*, *20*(3), 273–276. https://doi.org/10.1016/0042-6989(80)90113-3

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, *37*(2), 233–243. https://doi.org/10.1002/aic.690370209

Krasovskaya, S., & MacInnes, W. J. (2019). Salience Models: A Computational Cognitive Neuroscience Review. *Vision*, *3*(4), Article 4. https://doi.org/10.3390/vision3040056

Krasovskaya, S., Zhulikov, G., & MacInnes, J. (2021). *Deep Learning Neural Networks as a Component of a Model of Saccadic Generation*. PsyArXiv. https://doi.org/10.31234/osf.io/zr6ma

Krauskopf, J., Cornsweet, T. N., & Riggs, L. A. (1960). Analysis of Eye Movements during Monocular and Binocular Fixation*. *Journal of the Optical Society of America*, *50*(6), 572. https://doi.org/10.1364/JOSA.50.000572

Krauzlis, R. J., Goffart, L., & Hafed, Z. M. (2017). Neuronal control of fixation and fixational eye movements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160205. https://doi.org/10.1098/rstb.2016.0205

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience*, *12*(4), Article 4. https://doi.org/10.1038/nrn3008

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*(1), 26–49. https://doi.org/10.1016/j.tics.2012.10.011

Krebs, R. M., Boehler, C. N., Egner, T., & Woldorff, M. G. (2011). The Neural Underpinnings of How Reward Associations Can Both Guide and Misguide Attention. *Journal of Neuroscience*, *31*(26), 9752–9759. https://doi.org/10.1523/JNEUROSCI.0732-11.2011

Kristjánsson, Á. (2011). The intriguing interactive relationship between visual attention and saccadic eye movements. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0025

Kristjánsson, Á., Chen, Y., & Nakayama, K. (2001). Less attention is more in the preparation of antisaccades, but not prosaccades. *Nature Publishing Group*. https://doi.org/10.1038/nn723

Kristjánsson, Á., Vandenbroucke, M. W. G., & Driver, J. (2004). When pros become cons for anti- versus prosaccades: Factors with opposite or common effects on different saccade types. *Experimental Brain Research*, *155*(2), 231–244. https://doi.org/10.1007/s00221-003-1717-9

Kroell, L. M., & Rolfs, M. (2022). Foveal vision anticipates defining features of eye movement targets. *ELife*, *11*, e78106. https://doi.org/10.7554/eLife.78106

Kruthiventi, S. S. S., Ayush, K., & Babu, R. V. (2017). DeepFix: A Fully Convolutional Neural Network for Predicting Human Eye Fixations. *IEEE Transactions on Image Processing*, *26*(9), 4446–4456. https://doi.org/10.1109/TIP.2017.2710620

Kuen, J., Wang, Z., & Wang, G. (2016). *Recurrent Attentional Networks for Saliency Detection*. 3668–3677. https://openaccess.thecvf.com/content_cvpr_2016/html/Kuen_Recurrent_Attentional_Networks_CVPR_2016_paper.html

Kümmerer, M., Bylinskii, Z., Judd, T., Borji, A., Itti, L., Durand, F., Oliva, A., & Torralba, A. (n.d.). *MIT/Tuebingen Saliency Benchmark*. Retrieved July 1, 2023, from https://saliency.tuebingen.ai/

Kummerer, M., Wallis, T. S. A., Gatys, L. A., & Bethge, M. (2017). *Understanding Low- and High-Level Contributions to Fixation Prediction*. 4789–4798. https://openaccess.thecvf.com/content_iccv_2017/html/Kummerer_Understanding_Low-_and_ICCV_2017_paper.html

Lamare, M. (1892). *Des mouvements des yeux dans la lecture [Eye movements in reading]*. *10*, 354–364.

Lee, S. C. S., Martin, P. R., & Grünert, U. (2019). Topography of Neurons in the Rod Pathway of Human Retina. *Investigative Ophthalmology & Visual Science*, *60*(8), 2848–2859. https://doi.org/10.1167/iovs.19-27217

Lee, Y.-C., Lee, J. D., & Ng Boyle, L. (2007). Visual Attention in Driving: The Effects of Cognitive Load and Visual Disruption. *Human Factors*, *49*(4), 721–733. https://doi.org/10.1518/001872007X215791

Lee, Y.-C., Lee, J. D., & Ng Boyle, L. (2009). The Interaction of Cognitive Load and Attention-Directing Cues in Driving. *Human Factors*, *51*(3), 271–280. https://doi.org/10.1177/0018720809337814

Liu, N., & Han, J. (2018). A Deep Spatial Contextual Long-Term Recurrent Convolutional Network for Saliency Detection. *IEEE Transactions on Image Processing*, *27*(7), 3264–3274. https://doi.org/10.1109/TIP.2018.2817047

Liu, T., Slotnick, S. D., Serences, J. T., & Yantis, S. (2003). Cortical Mechanisms of Feature-based Attentional Control. *Cerebral Cortex*, *13*(12), 1334–1343. https://doi.org/10.1093/cercor/bhg080

Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, *103*(4), 603–649. https://doi.org/10.1037/0033-295x.103.4.603

Ludwig, C. J. H., & Gilchrist, I. D. (2002). Stimulus-driven and goal-driven control over visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 902–912. https://doi.org/10.1037/0096-1523.28.4.902

Machner, B., Lencer, M. C., Möller, L., von der Gablentz, J., Heide, W., Helmchen, C., & Sprenger, A. (2020). Unbalancing the Attentional Priority Map via Gaze-Contingent Displays Induces Neglect-Like Visual Exploration. *Frontiers in Human Neuroscience*, *14*. https://www.frontiersin.org/article/10.3389/fnhum.2020.00041

Macknik, S. L., & Martinez-Conde, S. (2015). The Age of Illusion. *Scientific American Mind*, *27*(1), 18–19. https://doi.org/10.1038/scientificamericanmind0116-18

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Dyar, T. A. (2006). Microsaccades counteract visual fading during fixation. *Neuron*, *49*(2), 297–305. https://doi.org/10.1016/j.neuron.2005.11.033

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: A neurophysiological analysis. *Trends in Neurosciences*, *32*(9), 463–475. https://doi.org/10.1016/j.tins.2009.05.006

McIntyre, M. M., & Graziano, W. G. (2016). Seeing People, Seeing Things: Individual Differences in Selective Attention. *Personality and Social Psychology Bulletin*, *42*(9), 1258–1271. https://doi.org/10.1177/0146167216653937

Merchant, J., Morrissette, R., & Porterfield, J. L. (1974). Remote Measurement of Eye Direction Allowing Subject Motion Over One Cubic Foot of Space. *IEEE Transactions on Biomedical Engineering*, *BME-21*(4), 309–317. https://doi.org/10.1109/TBME.1974.324318

Merzon, L., Malevich, T., Zhulikov, G., Krasovskaya, S., & MacInnes, W. J. (2020). Temporal Limitations of the Standard Leaky Integrate and Fire Model. *Brain Sciences*, *10*(1), Article 1. https://doi.org/10.3390/brainsci10010016

Miller, B. T., & D'Esposito, M. (2005). Searching for "the top" in top-down control. *Neuron*, *48*(4), 535–538. https://doi.org/10.1016/j.neuron.2005.11.002

Mirpour, K., Bolandnazar, Z., & Bisley, J. W. (2019). Neurons in FEF Keep Track of Items That Have Been Previously Fixated in Free Viewing Visual Search. *Journal of Neuroscience*, *39*(11), 2114–2124. https://doi.org/10.1523/JNEUROSCI.1767-18.2018

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417. https://doi.org/10.1016/0166-2236(83)90190-X

Møller, F., Laursen, M., Tygesen, J., & Sjølie, A. (2002). Binocular quantification and characterization of microsaccades. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *240*(9), 765–770. https://doi.org/10.1007/s00417-002-0519-2

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Moore, T., & Fallah, M. (2004). Microstimulation of the Frontal Eye Field and Its Effects on Covert Spatial Attention. *Journal of Neurophysiology*, *91*(1), 152–162. https://doi.org/10.1152/jn.00741.2002

Moran, R., Zehetleitner, M., Müller, H. J., & Usher, M. (2013). Competitive guided search: Meeting the challenge of benchmark RT distributions. *Journal of Vision*, *13*(8), 24. https://doi.org/10.1167/13.8.24

Motter, B. C., & Simoni, D. A. (2008). Changes in the functional visual field during search with and without eye movements. *Vision Research*, *48*(22), 2382–2393. https://doi.org/10.1016/j.visres.2008.07.020

Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*(3), 218–228. https://doi.org/10.1038/nrn1345

Nakayama, K., Maljkovic, V., & Kristjansson, A. (2004). Short-Term Memory for the Rapid Deployment of Visual Attention. In *The cognitive neurosciences, 3rd ed* (pp. 397–408). Boston Review.

Nobre, A. C., Sebestyen, G. N., Gitelman, D. R., Frith, C. D., & Mesulam, M. M. (2002). Filtering of Distractors during Visual Search Studied by Positron Emission Tomography. *NeuroImage*, *16*(4), 968–976. https://doi.org/10.1006/nimg.2002.1137

Noton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science (New York, N.Y.)*, *171*(3968), 308–311. https://doi.org/10.1126/science.171.3968.308

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*(3), 135–143. https://doi.org/10.1016/j.tics.2008.12.006

Opstal, A. J. van, & Kasap, B. (2019). Electrical stimulation in a spiking neural network model of monkey superior colliculus. *Progress in Brain Research*, *249*, 153. https://doi.org/10.1016/bs.pbr.2019.04.008

O'Regan, J. K., Lévy-Schoen, A., & Jacobs, A. M. (1983). The effect of visibility on eye-movement parameters in reading. *Perception & Psychophysics*, *34*(5), 457–464. https://doi.org/10.3758/BF03203061

Pan, J., Ferrer, C. C., McGuinness, K., O'Connor, N. E., Torres, J., Sayrol, E., & Giro-i-Nieto, X. (2018). *SalGAN: Visual Saliency Prediction with Generative Adversarial Networks* (arXiv:1701.01081). arXiv. https://doi.org/10.48550/arXiv.1701.01081

Paneri, S., & Gregoriou, G. G. (2017). Top-Down Control of Visual Attention by the Prefrontal Cortex. Functional Specialization and Long-Range Interactions. *Frontiers in Neuroscience*, *11*, 545. https://doi.org/10.3389/fnins.2017.00545

Paraskevoudi, N., & Pezaris, J. S. (2021). Full gaze contingency provides better reading performance than head steering alone in a simulation of prosthetic vision. *Scientific Reports*, *11*, 11121. https://doi.org/10.1038/s41598-021-86996-4

Park, S., Wiliams, L., & Chamberlain, R. (2022). Global Saccadic Eye Movements Characterise Artists' Visual Attention While Drawing. *Empirical Studies of the Arts*, *40*(2), 228–244. https://doi.org/10.1177/02762374211001811

Park, S.-J., Shin, J.-K., & Lee, M. (2002). Biologically Inspired Saliency Map Model for Bottom-up Visual Attention. In H. H. Bülthoff, C. Wallraven, S.-W. Lee, & T. A. Poggio (Eds.), *Biologically Motivated Computer Vision* (pp. 418–426). Springer. https://doi.org/10.1007/3-540-36181-2_42

Peelen, M. V., & Kastner, S. (2014). Attention in the real world: Toward understanding its neural basis. *Trends in Cognitive Sciences*, *18*(5), 242–250. https://doi.org/10.1016/j.tics.2014.02.004

Perry, V. H., Oehler, R., & Cowey, A. (1984). Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. *Neuroscience*, *12*(4), 1101–1123. https://doi.org/10.1016/0306-4522(84)90006-X

Pierce, J., McCardel, B., & Mcdowell, J. (2014). Trial-type probability and task-switching effects on behavioral response characteristics in a mixed saccade task. *Experimental Brain Research*, *233*. https://doi.org/10.1007/s00221-014-4170-z

Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Current Opinion in Neurology*, *17*(1), 17–25. https://doi.org/10.1097/00019052-200402000-00005

Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W.-H. C., Cohen, L., Guazzelli, M., & Haxby, J. V. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences*, *101*(15), 5658–5663. https://doi.org/10.1073/pnas.0400707101

Płużyczka, M. (2018). The First Hundred Years: A History of Eye Tracking as a Research Method. *Applied Linguistics Papers*, *4/2018*, 101–116. https://doi.org/10.32612/uw.25449354.2018.4.pp.101-116

Poletti, M., Listorti, C., & Rucci, M. (2013). Microscopic Eye Movements Compensate for Nonhomogeneous Vision within the Fovea. *Current Biology*, *23*(17), 1691–1695. https://doi.org/10.1016/j.cub.2013.07.007

Posner, M. I. (1988). Structures and function of selective attention. In *Clinical neuropsychology and brain function: Research, measurement, and practice* (pp. 173–202). American Psychological Association. https://doi.org/10.1037/10063-005

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*, 160–174. https://doi.org/10.1037/0096-3445.109.2.160

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), Article 1. https://doi.org/10.1038/4580

Rao, R. P. N., & Ballard, D. H. (2005). Probabilistic Models of Attention Based on Iconic Representations and Predictive Coding. In L. Itti, G. Rees, & J. K.

Tsotsos (Eds.), *Neurobiology of Attention* (pp. 553–561). Academic Press. https://doi.org/10.1016/B978-012375731-9/50095-1

Rauschecker, J. P. (2018). Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex*, *98*, 262–268. https://doi.org/10.1016/j.cortex.2017.10.020

Redden, R. S., MacInnes, W. J., & Klein, R. M. (2021). Inhibition of return: An information processing theory of its natures and significance. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *135*, 30–48. https://doi.org/10.1016/j.cortex.2020.11.009

Richardson, D. C., & Spivey, M. J. (2008). Eye Tracking: Characteristics and Methods. In *Encyclopedia of Biomaterials and Biomedical Engineering* (2nd ed.). CRC Press.

Rolfs, M. (2009). Microsaccades: Small steps on a long way. *Vision Research*, *49*(20), 2415–2441. https://doi.org/10.1016/j.visres.2009.08.010

Rolfs, M., Kliegl, R., & Engbert, R. (2008). Toward a model of microsaccade generation: The case of microsaccadic inhibition. *Journal of Vision*, *8*(11), 5. https://doi.org/10.1167/8.11.5

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: A Database and Web-Based Tool for Image Annotation. *International Journal of Computer Vision*, *77*(1–3), 157–173. https://doi.org/10.1007/s11263-007-0090-8

Rybak, I. A., Gusakova, V. I., Golovan, A. V., Podladchikova, L. N., & Shevtsova, N. A. (2005). Attention-Guided Recognition Based on "What" and "Where": Representations: A Behavioral Model. In L. Itti, G. Rees, & J. K. Tsotsos

(Eds.), *Neurobiology of Attention* (pp. 663–670). Academic Press. https://doi.org/10.1016/B978-012375731-9/50112-9

Saalmann, Y. B., Pigarev, I. N., & Vidyasagar, T. R. (2007). Neural Mechanisms of Visual Attention: How Top-Down Feedback Highlights Relevant Locations. *Science*, *316*(5831), 1612–1615. https://doi.org/10.1126/science.1139140

Sanders, A. F. (1970). Some Aspects of the Selective Process in the Functional Visual Field. *Ergonomics*, *13*(1), 101–117. https://doi.org/10.1080/00140137008931124

Schafer, R. J., & Moore, T. (2011). Selective Attention from Voluntary Control of Neurons in Prefrontal Cortex. *Science*, *332*(6037), 1568–1571. https://doi.org/10.1126/science.1199892

Schill, K. (2005). A Model of Attention and Recognition by Information Maximization. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 671–676). Academic Press. https://doi.org/10.1016/B978-012375731-9/50113-0

Schlag-Rey, M., & Schlag, J. (1984). Visuomotor functions of central thalamus in monkey. I. Unit activity related to spontaneous eye movements. *Journal of Neurophysiology*, *51*(6), 1149–1174. https://doi.org/10.1152/jn.1984.51.6.1149

Schlesinger, M., & McMurray, B. (2012). *The past, present, and future of computational models of cognitive development—ScienceDirect*. *27*(4). https://www.sciencedirect.com/science/article/pii/S088520141200041X?via%3Dihub

Sparks, D. L., & Hartwich-Young, R. (1989). The deep layers of the superior colliculus. *Reviews of Oculomotor Research*, *3*, 213–255.

St-Cyr, G. J., & Fender, D. H. (1969). Nonlinearities of the human oculomotor system: Time delays. *Vision Research*, *9*(12), 1491–1503. https://doi.org/10.1016/0042-6989(69)90065-0

Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, *11*(5), 13–13. https://doi.org/10.1167/11.5.13

Stuyven, E., Van der Goten, K., Vandierendonck, A., Claeys, K., & Crevits, L. (2000). The effect of cognitive load on saccadic eye movements. *Acta Psychologica*, *104*(1), 69–85. https://doi.org/10.1016/S0001-6918(99)00054-2

Sun, Y., & Fisher, R. (2003). Object-based visual attention for computer vision. *Artificial Intelligence*, *146*(1), 77–123. https://doi.org/10.1016/S0004-3702(02)00399-5

Suzuki, M., & Gottlieb, J. (2013). Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nature Neuroscience*, *16*(1), Article 1. https://doi.org/10.1038/nn.3282

Tallon, M., Greenlee, M. W., Wagner, E., Rakoczy, K., & Frick, U. (2021). How Do Art Skills Influence Visual Search? – Eye Movements Analyzed With Hidden Markov Models. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.594248

Tang, Y., Wu, X., & Bu, W. (2016). Deeply-Supervised Recurrent Convolutional Neural Network for Saliency Detection. *Proceedings of the 24th ACM International Conference on Multimedia*, 397–401. https://doi.org/10.1145/2964284.2967250

Tari, B., Fadel, M., & Heath, M. (2019). Alternating Between Stimulus-Driven and Minimally-Delayed Prosaccades: Switch-Costs Manifest via Response Suppression. *Journal of Vision*, *19*(10), 83c–83c. https://doi.org/10.1167/19.10.83c

Tari, B., & Heath, M. (2019). Pro- and antisaccade task-switching: Response suppression—and not vector inversion—contributes to a task-set inertia. *Experimental Brain Research*, *237*(12), 3475–3484. https://doi.org/10.1007/s00221-019-05686-w

Taylor, E., Rajsic, J., & Pratt, J. (2016). Object-based selection is contingent on attentional control settings. *Attention Perception & Psychophysics*, *78*. https://doi.org/10.3758/s13414-016-1074-y

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review*, *11*(1), 65–70. https://doi.org/10.3758/bf03206462

Tipper, S. P., & Baylis, G. C. (1987). Individual differences in selective attention: The relation of priming and interference to cognitive failure. *Personality and Individual Differences*, *8*(5), 667–675. https://doi.org/10.1016/0191-8869(87)90064-X

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

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

Tsotsos, J. K., Culhane, S. M., Kei Wai, W. Y., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, *78*(1), 507–545. https://doi.org/10.1016/0004-3702(95)00025-9

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592. https://doi.org/10.1037/0033-295X.108.3.550

van Gompel, R. P. G. (2007). *Eye Movements: A Window on Mind and Brain*. Elsevier.

Vecera, S. P., Cosman, J. D., Vatterott, D. B., & Roper, Z. J. J. (2014). The Control of Visual Attention: Toward a Unified Account. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 60, pp. 303–347). Academic Press. https://doi.org/10.1016/B978-0-12-800090-8.00008-1

Veniero, D., Gross, J., Morand, S., Duecker, F., Sack, A. T., & Thut, G. (2021). Top-down control of visual cortex by the frontal eye fields through oscillatory realignment. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21979-7

Vernon, M. D. (1928). Methods of Recording Eye Movements. *British Journal of Ophthalmology*, *12*(3), 113–130. https://doi.org/10.1136/bjo.12.3.113

Wade, N. J. (2010). Pioneers of Eye Movement Research. *I-Perception*, *1*(2), 33–68. https://doi.org/10.1068/i0389

Wade, N. J., & Tatler, B. W. (2011). Origins and applications of eye movement research. In *The Oxford handbook of eye movements* (pp. 17–43). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.001.0001

Wade, N. J., Tatler, B. W., & Heller, D. (2003). Dodge-ing the issue: Dodge, Javal, Hering, and the measurement of saccades in eye-movement research. *Perception*, *32*(7), 793–804. https://doi.org/10.1068/p3470

Wade, N., Ono, H., Mapp, A., & Lillakas, L. (2011). The Singular Vision of William Charles Wells (1757-1817). *Journal of the History of the Neurosciences*, *20*, 1–15. https://doi.org/10.1080/09647040903552764

Walther, D., Itti, L., Riesenhuber, M., Poggio, T., & Koch, C. (2002). Attentional Selection for Object Recognition—A Gentle Way. In H. H. Bülthoff, C. Wallraven, S.-W. Lee, & T. A. Poggio (Eds.), *Biologically Motivated Computer Vision* (pp. 472–479). Springer. https://doi.org/10.1007/3-540-36181-2_47

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*(9), 1395–1407. https://doi.org/10.1016/j.neunet.2006.10.001

Wandell, B. A. (1995). *Foundations of vision* (pp. xvi, 476). Sinauer Associates.

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual Field Maps in Human Cortex. *Neuron*, *56*(2), 366–383. https://doi.org/10.1016/j.neuron.2007.10.012

Wardak, C., Ibos, G., Duhamel, J.-R., & Olivier, E. (2006). Contribution of the Monkey Frontal Eye Field to Covert Visual Attention. *Journal of Neuroscience*, *26*(16), 4228–4235. https://doi.org/10.1523/JNEUROSCI.3336-05.2006

Watanabe, M., Matsuo, Y., Zha, L., Munoz, D. P., & Kobayashi, Y. (2013). Fixational saccades reflect volitional action preparation. *Journal of Neurophysiology*, *110*(2), 522–535. https://doi.org/10.1152/jn.01096.2012

Wells-Gray, E. M., Choi, S. S., Bries, A., & Doble, N. (2016). Variation in rod and cone density from the fovea to the mid-periphery in healthy human retinas using adaptive optics scanning laser ophthalmoscopy. *Eye*, *30*(8), Article 8. https://doi.org/10.1038/eye.2016.107

Willeke, K. F., Tian, X., Buonocore, A., Bellet, J., Ramirez-Cardenas, A., & Hafed, Z. M. (2019). Memory-guided microsaccades. *Nature Communications*, *10*(1), 3710. https://doi.org/10.1038/s41467-019-11711-x

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/BF03200774

Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In *Integrated models of cognitive systems* (pp. 99–119). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195189193.003.0008

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/S13423-020-01859-9

Yantis, S. (1998). Control of visual attention. In *Attention* (pp. 223–256). Psychology Press/Erlbaum (UK) Taylor & Francis.

Yarbus, A. L. (1967). Eye Movements During Perception of Complex Objects. In A. L. Yarbus (Ed.), *Eye Movements and Vision* (pp. 171–211). Springer US. https://doi.org/10.1007/978-1-4899-5379-7_8

Yellott, J. I., Wandell, B. A., & Cornsweet, T. N. (1980). The Beginnings of Visual Perception: The Retinal Image and its Initial Encoding. In *Handbook of Physiology* (pp. 257–301). https://doi.org/10.1002/cphy.cp010307

Young, A. H., & Hulleman, J. (2013). Eye movements reveal how task difficulty moulds visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 168–190. https://doi.org/10.1037/a0028679

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience*, *14*(5), Article 5. https://doi.org/10.1038/nn.2773

Zemliak, V., & MacInnes, W. J. (2022). The Spatial Leaky Competing Accumulator Model. *Frontiers in Computer Science*, *4*. https://www.frontiersin.org/articles/10.3389/fcomp.2022.866029

Zhang, J., & Sclaroff, S. (2013). *Saliency Detection: A Boolean Map Approach*. 153–160. https://www.cv-foundation.org/openaccess/content_iccv_2013/html/Zhang_Saliency_Detection_A_2013_ICCV_paper.html

Zhou, H., & Desimone, R. (2011). Feature-Based Attention in the Frontal Eye Field and Area V4 during Visual Search. *Neuron*, *70*(6), 1205–1217. https://doi.org/10.1016/j.neuron.2011.04.032

Zuber, B. L., & Stark, L. (1966). Saccadic suppression: Elevation of visual threshold associated with saccadic eye movements. *Experimental Neurology*, *16*(1), 65–79. https://doi.org/10.1016/0014-4886(66)90087-2

Zuber, B. L., Stark, L., & Cook, G. (1965). Microsaccades and the Velocity-Amplitude Relationship for Saccadic Eye Movements. *Science*, *150*(3702), 1459–1460. https://doi.org/10.1126/science.150.3702.1459

Original Publications

This section contains the original publications, Papers I, II and III, that are the essential parts of Studies I, II and III., respectively.