UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

The Mind’s Eye Controls Attention: The Impact of Mental Imagery on Visual Search

Permalink

https://escholarship.org/uc/item/1xr395td

Author

Moon, Austin

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Mind's Eye Controls Attention: The Impact of Mental Imagery on Visual Search

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Neuroscience

by

Austin Moon

December 2024

Dissertation Committee:
Dr. Rachel Wu, Chairperson
Dr. Ilana J. Bennett
Dr. Halle R. Dimsdale-Zucker

Copyright by
Austin Moon
2024

The Dissertation of Austin Moon is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

This dissertation would not have been possible without the support, guidance, and encouragement of many individuals. First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Rachel Wu, for her unwavering support, patience, and constant encouragement in the past five years. I am truly grateful to have an advisor who was always present throughout my graduate career. I also want to thank my dissertation committee: Dr. Lani Bennett and Dr. Halle Dimsdale-Zucker. I could not have asked for a stronger team of expert scientists who could have pushed my development as a critical thinker the way they did.

I want to thank several people in my lab. I started out as the only graduate student, and now, there are five of us. I want to thank all of my graduate lab members for their support and friendship: Leah Ferguson, Tania Rodriguez, Bethany Tavenner, and Isa Farias Lopes de Queiroz. I also want to thank my incredible team of undergraduate research assistants who were integral to the dissertation. They conducted long in-lab visits that required arduous procedures with setting up EEG. In no particular order, I want to thank Yixing Chen, Preetika Ashok, Alexander Nguyen, Rae Sunny Liu, Ivy Nguyen, Jayleen Andres, Cassandra Joy Rosales, Katherine Hernandez, Isaiah Viramontes, Sarah Sutjipto, Moses P. Chen, Teresa Ngo, Kevin Kim, Thao Tran, Areg Mehrabian, and Lluvia Arana Sanchez.

I must thank my family for their incredible support. To my father, Bong-Woon Moon, who set me up for great success before he passed. Even though you are not here now, you continue to inspire me to be a better person. To my mother, Myoung-Sook

Moon, who raised her children as a single parent the majority of my life. Nothing could ever pay back the sacrifices you made to ensure my success. To my sister, Justine Moon, who was essentially a second parent growing up. Thank you for always being there for me. To my partner, Diana Ortiz, for her unconditional support in every aspect of my career. You are not just my partner but also my best friend, and I feel like the luckiest person to have you be part of this journey. Finally, to my two cats, Monchis and Lucy, thank you for giving me daily pet therapy throughout my journey in graduate school.

ABSTRACT OF THE DISSERTATION

The Mind's Eye Controls Attention: The Impact of Mental Imagery on Visual Search

by

Austin Moon

Doctor of Philosophy, Graduate Program in Neuroscience
University of California, Riverside, December 2024
Dr. Rachel Wu, Chairperson

The human mind can 'see' visual objects without directly perceiving them through mental imagery. This ability represents objects in the mind to elicit visual experiences similar to those from perception. If imagery and perception represents objects in similar forms, then the effects of perception on subsequent processes should be similar to those of imagery. For example, a typical search task requires observers to bring to mind the target object to guide attention and increase search efficiency. However, whether imagery is deployed during search often is assumed and not empirically tested for. Imagery skills also widely vary across individuals based on how vividly their imagery appears to them, but how this variation impacts search is poorly understood. Across four studies, I investigated the use of mental imagery as an explicit strategy in search tasks. I hypothesized that imagery functions as a top-down process that mediates attentional processes and behavior. In Study 1, I present a series of visual search experiments with EEG that examined search efficiency (reaction time; the N2pc ERP components as a marker for attentional selection) following imagery. In Study 2, I

present another series of EEG experiments to compare imagery-related effects between salient targets and non-salient targets (search asymmetry). In Study 3, I conducted a single experiment with the goal to dissociate imagery-related effects on early perceptual processing (N2pc) from those on late post-selection (marked by the SPCN component). In Study 4, I analyzed all previous data with a different approach to address how variations in imagery skills (subjective vividness ratings) modulated imagery-related effects. The general results revealed stronger and faster N2pc components, and faster reaction time, when searching for a target that matched a previously imagined object, compared to searching for a mismatched target. Neural signals during search were distinguishable and reaction time differed based on vividness ratings. The collective findings suggest that mental imagery is a potent strategy for completing search tasks and that the effects of imagery are bounded by individual imagery skills.

Table of Contents

Chapter 1: Mental imagery and visual attention 1

1.1. Mental imagery 1

1.1.1. Theories of imagery 2

1.1.1.1. Depictive vs. Propositional Theory 2

1.1.1.2. Reverse Hierarchy Model 5

1.2. Visual attention 7

1.2.1. Theories of attention 9

1.2.1.1. Stimulus-driven account 9

1.2.1.2. Goal-driven account 11

1.2.1.3. Post-selection 14

1.3. Relation between imagery and visual attention 16

1.3.1. Impact of imagery in visual search tasks 17

1.3.2. Current gaps 19

1.3.2.1. Does imagery actually guide attention? (Study 1) 19

1.3.2.2. What about bottom-up processing? (Study 2) 20

1.3.2.3. Does imagery impact post-selection? (Study 3) 21

1.3.2.4. Individual differences in imagery skills (Study 4) 22

1.4. Outline of dissertation 24

Chapter 2: EEG 28

2.1. Description of EEG 28

2.1.1. What is an EEG signal? 28

2.1.2. What are ERP components? 31
2.1.3. Why use ERP components here? 32
2.2. EEG set-up 33
2.2.1. Electrodes 33
2.2.2. Experimental room 35
2.2.3. Procedure 37
2.2.4. Steps for preprocessing 39
2.3. ERP components 40
2.3.1. N2pc 42
2.2.1.1. N2pc mean amplitude 42
2.2.1.2. N2pc onset latency 43
2.3.2. SPCN (Study 3, Chapter 5) 43
2.2.2.1. SPCN mean amplitude (Study 3, Chapter 5) 44
2.4. Multivariate pattern analysis (MVPA) 44
2.4.1. Classification accuracy (Study 4, Chapter 6) 45
2.5. Adequate sample size 45
2.6. Adequate number of trials 46
Chapter 3: Top-down effects of imagery on early attentional processes 48
3.1. Abstract 48
3.2. Introduction 49
3.3. Methods 52
3.3.1. Participants 52

3.3.2. Stimuli 53
3.3.3. Design 55
3.3.4. Procedures 56
3.3.5. EEG system and steps for preprocessing 58
3.4. Results 60
3.4.1. Experiment 1 60
3.4.1.1. N2pc mean amplitude 60
3.4.1.2. N2pc onset latencies 62
3.4.1.3. Reaction time 63
3.4.1.4. Summary 64
3.4.2. Experiment 2 64
3.4.2.1. N2pc mean amplitude 64
3.4.2.2. N2pc onset latencies 67
3.4.2.3. Reaction time 68
3.4.2.4. Summary 69
3.4.3. Experiment 3 70
3.4.3.1. N2pc mean amplitude 71
3.4.3.2. N2pc onset latencies 73
3.4.3.3. Reaction time 74
3.4.3.4. Summary 75
3.5. General discussion 75
3.5.1. Summary and interpretation of findings 75

3.5.2. Theoretical contributions 78
3.5.3. Limitations 80
3.5.4. Conclusion 81
Chapter 4: Interaction between top-down imagery and pop-out effects 82
4.1. Abstract 82
4.2. Introduction 83
4.3. Methods 85
4.3.1. Participants 85
4.3.2. Stimuli 86
4.3.3. Design 86
4.3.4. Procedure 87
4.3.5. EEG system and data preprocessing 89
4.4. Results 89
4.4.1. Experiment 4 89
4.4.1.1. Average reaction time 89
4.4.1.2. Slope of reaction time 90
4.4.1.3. Summary 92
4.4.2. Experiment 5 92
4.4.2.1. N2pc mean amplitude 92
4.4.2.2. N2pc onset latencies 93
4.4.2.3. Average reaction time 95
4.4.2.4. Slope reaction time 96

4.4.2.5. Summary 98
4.5. General discussion 98
4.5.1. Summary and interpretation of findings 98
4.5.2. Theoretical contributions 102
4.5.3. Limitations 103
4.5.4. Conclusion 105
Chapter 5: Selection vs. Post-selection 106
5.1. Abstract 106
5.2. Introduction 107
5.3. Methods 108
5.3.1. Participants 108
5.3.2. Stimuli and design 109
5.3.3. Procedures 110
5.3.4. EEG system and data preprocessing 110
5.4. Results 112
5.4.1. N2pc mean amplitude 112
5.4.2. SPCN mean amplitude 114
5.4.3. Reaction time 116
5.5. General discussion 117
5.5.1. Summary and interpretation of findings 117
5.5.2. Theoretical contributions 118
5.5.3. Limitations 119

5.5.4. Conclusion 120

Chapter 6: Individual differences in vividness 121

6.1. Abstract 121

6.2. Introduction 121

6.3. Methods 123

6.3.1. Participants 123

6.3.2. MVPA technique with EEG data 123

6.4. Results 124

6.4.1. Classification accuracy 124

6.4.2. Reaction time based on vividness 126

6.5. General discussion 127

6.5.1. Summary and interpretation of findings 127

6.5.2. Theoretical contributions 130

6.5.3. Limitations 132

6.5.4. Conclusion 133

Chapter 7: Conclusion 134

7.1. Overview 134

7.2. Main findings 135

7.2.1. Did imagery elicit top-down effects? 135

7.2.2. Did imagery modulate pop-out effects? 136

7.2.3. Did imagery impact post-selection? 139

7.2.4. Did vividness modulate top-down effects? 140

7.3. Contributions 142
7.3.1. Informing the debate between stimulus-driven and goal-driven account 142
7.3.2. Informing the imagery debate 143
7.3.3. Building on the Reverse Hierarchy Model 145
7.3.4. An upper limit to attentional processing 145
7.3.5. Methodological advancements 146
7.4. Implications outside current work 148
7.4.1. Are imagery and templates one and the same? The next debate to come 148
7.4.2. Target enhancement vs. distractor suppression 150
7.4.3. Causal role of imagery: a case for studying aphantasia 151
7.4.4. Is imagery the best strategy? 153
7.5. Unresolved issues to address in future studies 155
7.5.1. Active control group 155
7.5.2. Unclear effects on post-selection 156
7.5.3. Limitation of the MVPA technique 157
7.5.4. Did participants actually follow the imagery instruction? 158
7.6. Conclusion 159
References 160

List of Tables

Table 6.1 *Fixed effects for LME of vividness data* 126

Table 7.1 *Summary of findings* 137

List of Figures

Figure 1.1 *Depictive and propositional theories of imagery* 2

Figure 1.2 *Reverse hierarchy model* 6

Figure 1.3 *Example of asymmetric search design* 10

Figure 1.4 *Multistage model of search* 11

Figure 2.1 *Dipole of a neuron* 30

Figure 2.2 *Sample ERP waveform* 32

Figure 2.3 *Easycap electrode layout with 64 channels* 34

Figure 2.4 *Experimental room* 36

Figure 2.5 *Experimental materials* 39

Figure 2.6 *Example of N2pc and SPCN components* 41

Figure 3.1 *Sample Stimuli* 54

Figure 3.2 *Sample Trial Types* 56

Figure 3.3 *Sample Trial* 59

Figure 3.4 *Grand averaged waveforms (Experiment 1)* 61

Figure 3.5 *N2pc mean amplitudes (Experiment 1)* 62

Figure 3.6 *Average reaction time (Experiment 1)* 63

Figure 3.7 *Grand averaged waveforms (Experiment 2)* 66

Figure 3.8 *N2pc mean amplitudes (Experiment 2)* 67

Figure 3.9 *Average reaction time (Experiment 2)* 69

Figure 3.10 *Grand averaged waveforms (Experiment 3)* 72

Figure 3.11 *N2pc mean amplitudes (Experiment 3)* 73

Figure 3.12 *Average reaction time (Experiment 3)* 74

Figure 4.1 *Sample trial* 88

Figure 4.2 *Averaged reaction time (Experiment 4)* 90

Figure 4.3 *Slope reaction time (Experiment 4)* 91

Figure 4.4 *Grand averaged waveforms (Experiment 5)* 94

Figure 4.5 *N2pc mean amplitudes (Experiment 5)* 95

Figure 4.6 *Average reaction time (Experiment 5)* 96

Figure 4.7 *Slope reaction time (Experiment 5)* 97

Figure 5.1 *Sample trial* 111

Figure 5.2 *Grand averaged waveforms (Experiment 6)* 113

Figure 5.3 *N2pc mean amplitudes (Experiment 6)* 114

Figure 5.4 *SPCN mean amplitudes (Experiment 6)* 115

Figure 5.5 *Average reaction time (Experiment 6)* 116

Figure 6.1 *Classification accuracy of vividness* 125

Figure 6.2 *Average reaction time by vividness rating* 127

Chapter 1: Mental imagery and visual attention

1.1. Mental imagery

Close your eyes and picture a sunset over a serene beach. Most people are able to do this task by using mental imagery (seeing with the "mind's eye"). Imagery is the ability to have visual experiences without directly perceiving any external stimuli (Kosslyn, 1981). Many would consider imagery to be useful in tasks that demand attention, such as visual search. For example, pretend you are a parent at a crowded beach, and your mischievous son runs off without your notice. You are confronted with a task in which you must find your son among the crowd. Without a strategy, you need to inspect every person until you finally find your son; a taxing and slow way to search in a crowded setting. Alternatively, you may deploy imagery to visualize what your son was wearing (e.g., blue swimming trunks) to narrow your focus while ignoring non-matching items (e.g., red trunks); a less taxing and faster way to search.

It seems obvious that imagery is a good strategy for visual search. However, empirical findings to support this view shockingly are underwhelming. To understand the impact of imagery on attention, it is critical to understand two fundamental ideas of imagery. First, what is the nature of imagery? Many would describe their imagery as if they 'see' visual information (e.g., the color blue), whereas others would describe the same information purely in a symbolic sense (e.g., the concept of blue as a color). Second, does the nature of imagery serve a functional purpose for attentional processing? If imagery is analogous to 'seeing' visual information, then mechanisms of imagery should tap into related mechanisms that process visual information. Such a matter has

puzzled scientists in the 1970s, with competing theories on how imagery is internally represented. Two prominent theories dominated the discourse: Kosslyn's depictive theory and Pylyshyn's propositional theory. The depictive theory states that imagery activates 'pictorial' (visual) representations that overlap with those during perception (Kosslyn, 1975, 1981), whereas the propositional theory states that the representations are descriptive and language-like(Z. Pylyshyn, 2003). Below, I review the theoretical accounts of imagery.

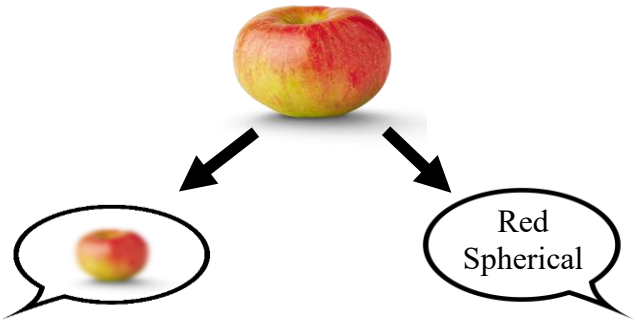
1.1.1. Theories of imagery

1.1.1.1. Depictive vs. Propositional Theory

Figure 1.1 illustrates the two competing theories. The propositional theory argues that imagery is represented by symbolic descriptions that are used in language (Z. Pylyshyn, 2003; Z. W. Pylyshyn, 1973, 1973, 1981), whereas the depictive theory argues that imagery can be structured as a visual representation metaphorically similar to a picture (Kosslyn, 1975, 1981, 2005; Kosslyn et al., 2006).

Figure 1.1

Depictive and propositional theories of imagery



Note. Representation of apple according to depictive (left) and propositional (right) accounts.

The two competing theories agree on *what* imagery can represent (the content) but disagree on *how* imagery is represented (the format). For example, visual information (e.g., content of apple) can be depicted as a low-resolution photograph (picture of apple) or conveyed through language ("red and spherical"). The imagery debate called into question whether visual information *can* be formatted as a visual representation. The propositional theory stemmed from strict behaviorist viewpoints in the 1950s that argued against the necessity of internal representations for behavior (Z. Pylyshyn, 2003). In such a case, all visual information are stored in a descriptive manner, such that imagery is akin to 'inner speech'. The depictive theory argues that the same information can be structured in other formats, such as visual representations. Why was there much discourse on the representational aspects of imagery? Perhaps one's anecdote with imagery partially drove the debate. A previous study found that scientists who viewed their own imagery with vivid details were more sympathetic to the depictive theory, while those with less vivid details ascribed more to the propositional theory (Reisberg et al., 2003).

With the advancement of neuroimaging techniques in the 1990s, the debate took on a new form with appeals to perceptual networks. Empirical findings come from functional magnetic resonance imaging (fMRI) studies that compared neural activity patterns between imagery tasks and perception tasks (Bartolomeo, 2002; Cichy et al., 2012; Dijkstra et al., 2019). Following the depictive theory, brain regions that activate visual representations during perception (e.g., visual cortex) should play a role in imagery. The primary visual cortex (V1) was a candidate site as it is the first cortical region to receive visual input from the retina (Pearson & Kosslyn, 2015). V1 creates a

retinotopic map of the visual field, meaning that areas of the retina project onto corresponding areas in V1 (Wandell et al., 2007). Damage to a local portion of V1 creates a blind spot in the corresponding part of the retina (Horton & Hoyt, 1991). Together, representations in V1 are necessarily depictive, and if activity patterns in V1 are present during imagery tasks, that may suggest imagery is also depictive. Indeed, several findings showed comparable activities between imagery and perception tasks (Pearson & Kosslyn, 2015), though not always perhaps due inter-study differences (Pearson, 2019).

Beyond demonstrating 'mere' overlaps in activity patterns, other fMRI studies implemented multivariate pattern analysis (MVPA) to identify activity patterns that were associated with imagery. One utility of MVPA is to train an algorithm to decode activity patterns in V1 based on what stimulus participants perceived (Haxby et al., 2001). For example, voxel-wise models of perception were trained to decode low-level visual features (e.g., orientation) that form the building blocks of visual perception (K. N. Kay et al., 2008). These models that were trained from data during perception tasks successfully decoded similar features during imagery tasks (Albers et al., 2013; Stokes et al., 2009; Thirion et al., 2006), meaning that perceiving a feature evoked activity patterns in V1 similar to ones when imaging a feature (Lee et al., 2012). Can the source of this similarity be explained by a depictive account or are there other contributing sources? For example, perceiving and imagining complex scenes may rely on sources beyond the ones for low-level features. A critical study addressed this concern by using a multifeatured-based model (Naselaris et al., 2011) trained to decode low-level features of complex scenes (Naselaris et al., 2015). Findings revealed similar pattern of results, with

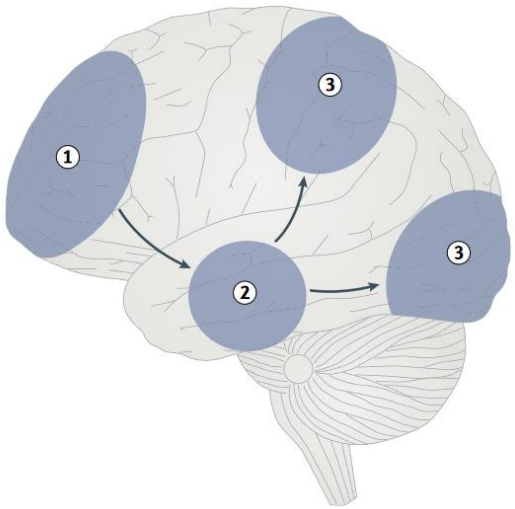
successful decoding of imagined features from the model fitted to perceptual data, further supporting the depictive theory. In summary, converging evidence from fMRI studies gives credence to the view that imagery can be structured as a visual representation.

1.1.1.2. Reverse Hierarchy Model

Given there were imagery-perception similarities in activity patterns, many have postulated that the mechanisms of imagery may act on mechanisms of perception in a 'top-down' manner (Pearson, 2019). Top-down refers to a broad category of processes that are goal-driven, meaning that the individual voluntarily engages with cognitive processes that guide thoughts and behaviors. As such, visual representations are activated voluntarily during imagery. On a neural level, a top-down process can be characterized by activity patterns in low-level sensory regions (e.g., V1) that are modulated by higher-level regions (e.g., frontal regions, Miller & Cohen, 2001). The Reverse Hierarchy Model depicts imagery in a similar way (Pearson, 2019; Figure 1.2). First, imagery triggers activity patterns in the frontal cortex. Previous fMRI studies showed activation in frontal regions when participants performed imagery tasks (Ishai et al., 2000; Ragni et al., 2021). Frontal activity patterns were decoded more accurately for task identity than for object representation (Hebart et al., 2018), suggesting that the frontal region was responsible for a general executive role (Dijkstra et al., 2019; Mechelli, 2004) as opposed to generating visual representations.

Figure 1.2

Reverse hierarchy model



Note. Taken from Pearson (2019). Imagery is initiated in frontal region (1). Activity patterns are sent to the medial temporal lobe where memory is retrieved (2). Spatial representations are activated in the parietal region, and visual representations are activated in the occipital region (3).

Second, these activity patterns pass through the medial temporal lobe (MTL) including the hippocampal region. Previous studies found that the hippocampus was active during imagery tasks (e.g., Bird et al., 2010; Ranganath & D'Esposito, 2005), and hippocampal lesions were associated with disrupted imagery skills (e.g., Hassabis et al., 2007). The main point is that imagery requires memory retrieval in order to activate relevant representations. Third, though not the focus of this dissertation, if imagery requires spatial information (e.g., location of objects), spatial representations of imagery are activated in parietal regions (Winlove et al., 2018). More importantly, the activity patterns in MTL cascade across visual regions (e.g., V1) that activate visual representations. The involvement of the primary visual cortex was discussed in the previous section, with evidence supporting the depictive account.

Putting these ideas together, the imagery debate led to critical findings in favor of the view that imagery can activate visual representations (Pearson & Kosslyn, 2015). These representations are housed in regions typically engaged during visual perception. FMRI studies that used MVPA techniques decoded activity patterns in the primary visual cortex (Albers et al., 2013; Bannert & Bartels, 2018; Cichy et al., 2012; Dijkstra et al., 2017; Naselaris et al., 2015). These findings support the depictive theory that imagery can be structured as visual representations similar to those during perception (Slotnick et al., 2005). Furthermore, imagery-induced activity patterns in visual perception regions can be modulated by higher-level regions, suggesting that imagery can function in a top-down manner.

1.2. Visual attention

Returning to the 'find son at beach' example, some key assumptions are made. First, it is impossible to attend to every item in a crowded setting. You cannot simultaneously attend to every clothing item or person at the beach. This obvious point was heavily discussed in the 1950s. A limited attentional capacity has been described as a 'bottleneck' (A. M. Treisman, 1969), such that incoming information from the environment must be selectively filtered. Since the 1990s, the focus shifted from how the selection process works to what factors influence the selection process (Wolfe & Horowitz, 2017), which leads to the next assumption. Some information will catch your attention more than others. If your goal is to find 'blue trunks', then selection will be directed to people in blue trunks more than people in red trunks. Like imagery, this form of selection functions in a top-down manner, as attention is driven voluntarily to goal-

matching items (Desimone & Duncan, 1995). Additionally, some items may 'pop out' from the setting and will automatically capture your attention, such as bright objects (e.g., yellow umbrella) or distinctive objects (e.g., a large sandcastle or an upright surfboard). This form of selection functions in a bottom-up manner as attention is captured without voluntary effort (Theeuwes, 1994). Both top-down and bottom-up selection have been explained by 'goal-driven' and 'stimulus-driven' accounts, respectively (Theeuwes, 2010; Wolfe et al., 2003).

The final assumption is that once items are selected, further processing is necessary to recognize the items. Once a candidate target is selected (*someone* in blue trunks), you must determine whether this item is in fact the target (*your* child in blue trunks) or not (*another* child in blue trunks). This post-selection process is separate from the selection process (Eimer, 2014). Below, I discuss the theoretical accounts of selection (stimulus-driven and goal-driven) and of post-selection. Specifically, I focused on visual search tasks, which were widely used to investigate attentional processing and the efficiency of behaviorally relevant responses (Wolfe, 1998). These tasks typically required participants to locate a target item among varying numbers of nontargets. Search paradigms can be systematically manipulated to test for between-condition effects, offering the means to study attentional processing.

1.2.1. Theories of attention

1.2.1.1. Stimulus-driven account

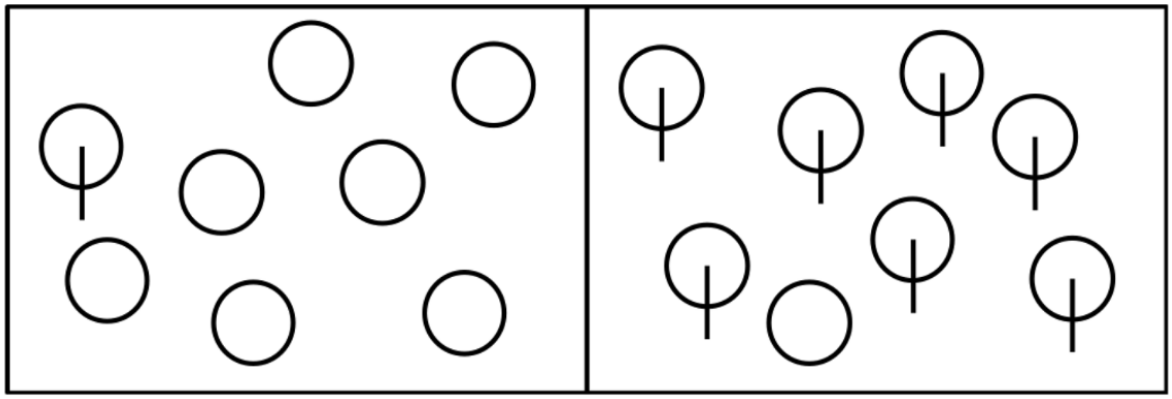
The stimulus-driven account postulates that attention is automatically captured by salient features, independent of the observer's goals (Jonides & Yantis, 1988; Theeuwes, 1991, p. 199, 1992). Salient features elicit 'pop-out' effect (Theeuwes, 2018; Theeuwes et al., 2006), characterized as the attentional capture of salient features through bottom-up processing. Influential search studies tested the pop-out effect where one item was different from others by a salient feature (Theeuwes, 1992). Previous studies used singleton search to examine pop-out effects when a nontarget possessed a salient feature (e.g., Theeuwes & Burger, 1998). Each trial presented either a salient nontarget or no salient nontarget, alongside the target. Across several studies, reaction times were slower in trials with the salient feature compared to trials without it (Belopolsky et al., 2007; Franconeri & Simons, 2003; B. Wang & Theeuwes, 2020). Collectively, these results showed that bottom-up processing is sensitive to saliency.

The previous studies examined pop-out effects when the nontarget possessed a salient feature. The same effect has been observed when the target possessed the same feature. This phenomenon was well-observed in asymmetric search studies (A. Treisman & Souther, 1985; Wolfe, 2001). Asymmetry refers to differences in search efficiency based on the presence or absence of salient target-defining features. Figure 1.3 shows a classic example of asymmetric search design. The search goal was to identify the unique stimulus among homogenous nontargets. The left panel shows a target with a salient feature, which I refer to as Q. The right shows a target that is absent of such feature,

which I refer to as O. Seminal work showed that search for ‘Q among Os’ was more efficient than ‘O among Qs’ (A. Treisman & Souther, 1985).

Figure 1.3

Example of asymmetric search design



Note. Taken from Guo et al. (2021). Search was more efficient when the target has the vertical line among circles (left) than vice versa (right).

The stimulus-driven account supports asymmetric search in that attentional capture is automatic without requiring top-down processes. This is reflected in the criterion that search should be ‘effortless’, meaning that attentional capture occurs independent of the observer’s goal. Asymmetry studies often manipulated the set size of search items to test search efficiency, measured by slope reaction time as a function of set size (Wolfe et al., 2010). Several studies showed flat slopes when salient features were present, indicating that a pop-out effect occurred. In the ‘Q among Os’ search (Figure 1.3, left panel), the vertical line of the target elicited pop-out effects. In contrast, studies showed positive slopes when the feature was absent, indicating that a pop-out effect did not occur (Dosher et al., 2004; Saiki et al., 2005; Shen & Reingold, 2001; A. Treisman & Gormican, 1988). The absence of the vertical line required a serial, or more effortful,

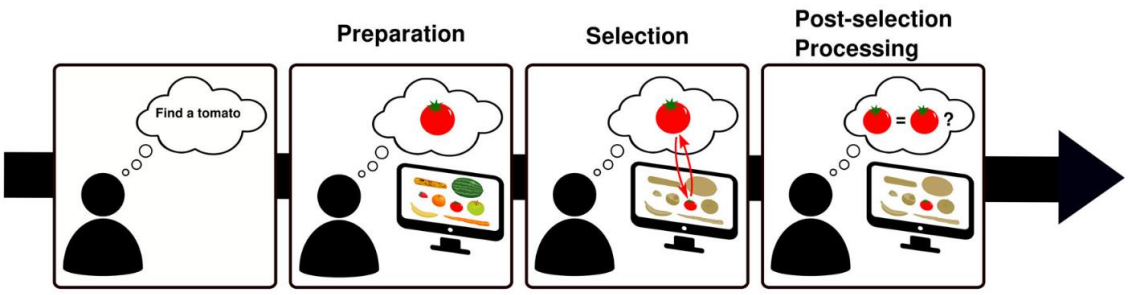
search in which each item was processed one by one until a target was found (A. Treisman & Gormican, 1988). Collectively, these results showcased an important phenomenon in search where bottom-up processing was sensitive to saliency.

1.2.1.2. Goal-driven account

The goal-driven account challenged the stimulus-driven account by proposing that attention is voluntarily controlled based on the observer's goal, rather than the physical properties of the stimuli. Top-down selection prioritizes features relevant to current tasks, whereas salient features only influence attention when they align with the goal (Folk et al., 1992, p. 199, 1994; Folk & Remington, 1998, p. 19). Figure 1.4 shows a traditional model that describes top-down selection.

Figure 1.4

Multistage model of search



Note. Taken from Ort & Olivers (2020). The observer is given a search cue (Box 1). An internally generated representation (template) of the search cue is activated (Box 2). The template guides attention by selecting matching items (Box 3). The observer determines whether the selected item is the target (Box 4).

In everyday search tasks, the observer typically knows what to search for, like blue trunks in the beach example. Likewise in experimental search tasks, participants often are cued to search for a specific item (Box 1). For example, in the famous Posner

paradigm, a left or right arrow cue was presented to prepare attention to the possible target location (Posner et al., 1980). In other cases, participants were presented with a visual cue to prepare attention to the possible target feature (Töllner et al., 2010; Vickery et al., 2005; Wolfe et al., 2004). These cue types served to generate attentional 'templates' of the target during preparation (Box 2). Similar to imagery, templates are characterized as visual representations that were mapped in the visual cortex (Desimone & Duncan, 1995; Duncan & Humphreys, 1989). Visual processing regions were observed to fire selectively based on target-defining features (Chawla et al., 1999; Giesbrecht et al., 2006). This selective firing reflects preparatory activity patterns in response to cued objects (Chawla et al., 1999; Chelazzi et al., 1998; Kok et al., 2013; Serences et al., 2004). Together, the findings suggest that visual representations of the search goal are activated during preparation.

After templates are created, the selection process starts (Box 3). The function of templates during this stage was extensively studied and modeled in famous work (e.g., Guided Search 6.0, Wolfe, 2021; attentional control model, Eimer, 2014; biased competition model, Desimone & Duncan, 1995). A common explanation was that a template (e.g., of a tomato) guides attention by enhancing visual processing of task-relevant features (red and circular) while inhibiting processing of task-irrelevant features (e.g., yellow and elongated shape). Finally, once an item is selected, the observer must identify whether the item is the target so that search can be terminated, or nontarget so that search can continue (Box 4).

If attention is sensitive to internally generated representations (Olivers et al., 2011), then behaviorally relevant responses should be influenced by top-down effects. Following the cue-search paradigms, several studies showed behavioral findings in support of the goal-driven account. For example, a given trial cued a probable target feature (e.g., 'red' word cue) in an upcoming search task, in which the cue either correctly predicted the target (red shape among blue shapes) or incorrectly (blue shape among red shapes; Müller et al., 2003). Search performance was better following correct cues (Conci et al., 2019; Müller et al., 2003; Töllner et al., 2010; Vickery et al., 2005; Wolfe et al., 2004) than incorrect cues (Müller et al., 2003; Töllner et al., 2010), suggesting there to be top-down control of attention. However, whether top-down effects specifically impact selection (Box 3) or instead post-selection (Box 4) has been a controversial topic (Eimer & Kiss, 2010). The controversy lies in whether bottom-up processing described in the stimulus-driven account is 'impenetrable' to top-down processing. A challenge to this concern was that measures of behavior do not directly capture responses that occur at specific stages of search (Figure 1.4). Any impact that either processes have on selection or post-selection ultimately would impact behavior (Eimer & Kiss, 2010). Rather, a measure with high temporal resolution was necessary to mark responses during selection.

Given that attentional processing occurs within milliseconds (ms; Woodman & Luck, 1999, 2003), electroencephalogram (EEG) is widely used with search paradigms for its high temporal resolution. The event-related potentials (ERPs) can be extracted from EEG data to measure components associated with task-relevant processes. For example, several studies measured the N2pc component to mark responses during

selection (Luck & Hillyard, 1994). The N2pc is the fastest and most temporally precise marker for selection, measured approximately 200 to 300 ms after the stimulus presentation (Eimer, 1996; Luck, 2011). Pivotal studies examined robust N2pc components in tasks with the pop-out design (Hickey et al., 2006) and cue-search paradigm (Kiss, Van Velzen, et al., 2008; Woodman et al., 2009; Woodman & Luck, 1999). This measure paved the way for countering the strict stimulus-driven view that bottom-up processing cannot be influenced by top-down processing. To disprove this hypothesis, a simple experiment was to show the N2pc component in response to pop-out effects and then examine changes to the component with the addition of cue types. Indeed, previous studies showed this pattern of results and furthermore, an absence of the N2pc when the salient features were task-irrelevant (Eimer & Kiss, 2008; Kiss et al., 2012; Kiss, Jolicœur, et al., 2008), suggesting that bottom-up processing could be modulated by top-down processing.

1.2.1.3. Post-selection

The previous conclusion regarding the N2pc findings faced rebuttals from the stimulus-driven camp. Perhaps the N2pc did not mark the selection stage and instead post-selection (Theeuwes, 2010). Therefore, any top-down effects that may have occurred took place during post-selection, after bottom-up processing during selection. This view has been countered by the fact that two separate ERP components could mark these stages independently. Separate from the N2pc component (Eimer, 1996; Luck, 2011), the SPCN component reflects the signals approximately from 400 to 600 ms after stimulus onset (Jolicœur et al., 2008; Mazza et al., 2007). Post-selection involves extracting

features of the selected item (Huang & Pashler, 2007; Xu & Chun, 2009). For example, when participants were tasked with selecting multiple items (e.g., 4 briefly appearing grey shapes) and later tested on their memories of the features, their performance was poor (Xu & Chen, 2009). One interpretation was that post-selection is necessary for analyzing features of selected items, whereas selection is necessary for allocating attention to task-relevant features (Ghorashi et al., 2010). Given their functional differences, previous studies examined the N2pc and SPCN components separately by manipulating task demands. When the search task was to simply locate the target (allocate attention to spatially relevant locations), only the N2pc was present (Dell'Acqua et al., 2006; Jolicœur et al., 2008; Rashal et al., 2023). When the search task required identifying a feature of the target (allocate spatial attention *and* analyze feature), both the N2pc and SPCN were present (Berggren & Eimer, 2018; Grubert et al., 2017; Mazza et al., 2007). In one important study, participants were instructed either to locate the target (e.g., line) with a distinctive feature (e.g., orientation) or to locate the target and then identify the feature (e.g., left or right orientation, Mazza et al., 2007). The former search required only selection, evinced by the N2pc but no SPCN component, whereas the latter search required both selection and post-selection, evinced by both N2pc and SPCN components.

Putting these ideas together, attentional selection could be influenced by bottom-up processing (e.g., pop-out effects, Theeuwes, 1994) and top-down processing (e.g., guidance of templates, Wolfe, 2021). A strict stimulus-driven account cannot address the N2pc findings from tasks that evoke both bottom-up and top-down processing. It is clear

that in some cases, bottom-up processing could be modulated by top-down processing during the selection stage (Eimer & Kiss, 2010). The top-down effects could be tracked by the N2pc components (Eimer, 1996; Luck, 2011), separate from responses during the post-selection stage, marked by the SPCN component (Mazza et al., 2007).

1.3. Relation between imagery and visual attention

The theoretical accounts of imagery state that imagery can activate visual representations in the visual cortex also engaged during perception (depictive theory). These representations are activated in a top-down manner, such that activity patterns in the visual cortex are modulated by higher-level, frontal regions (Reverse Hierarchy Model). The theoretical accounts of visual attention state that the attentional selection is a limited system ('bottleneck') that filters in relevant information and out irrelevant information. The selection process is influenced by bottom-up processing when salient features are present (stimulus-driven account) and by top-down processing when a goal is internally represented (goal-driven account). Once an object is selected, a subsequent post-selection process is necessary to recognize features of the object.

Putting these theoretical accounts together, the functional role of imagery in tasks that demand attention may be similar to what is described about top-down processing by the goal-driven account. First, both imagery and the preparation stage involve the activation of visual representations. It is understudied whether these representations overlap with one another. Here, I only point to the fact they both can be visually structured as described by the depictive theory, thus may elicit similar top-down effects. Second, a recent line of studies showed converging evidence that suggest imagery can

guide attention in search tasks, with similar results to what was observed in 'template-based' search studies. Below, I review these findings and then discuss the current gaps that motivated this dissertation.

1.3.1. Impact of imagery in visual search tasks

Given that imagery can activate visual representations similar to those of perception (Dijkstra et al., 2019), recent studies have compared behavioral effects from perceptual representation with imagery (Cochrane & Milliken, 2019; Wantz et al., 2015). For example, color imagery has been shown to influence performance in visual search tasks. Cochrane, Wang, et al. (2021) implemented a search task where in each trial, participants were cued either a displayed perceptual color (e.g., a green shape) or an imagery word indicated by the first two letters of the color (e.g., 'gr' for green). A search task followed, in which a target and homogenously colored nontarget shapes were presented. When imagery cues were presented, reaction time was faster when targets (e.g., green among red shapes) matched the imagined color (congruent) than when nontargets instead matched (incongruent; red among green shapes). When perceptual cues were presented, the same behavioral effect was present albeit smaller than the imagery effect. Related studies showed this congruency-incongruency tradeoff effects. Moriya (2018) implemented a search task where in each trial, participants were cued a word of a color (e.g., "blue") to imagine. A search task followed, in which a target and nontarget square with different colors were presented in opposing visual hemifields. Reaction time was faster when targets (e.g., blue square) matched the imagined color (congruent) than when nontargets matched (incongruent). This behavioral effect was not

observed in a control group that did not receive any imagery instruction (Moriya, 2018, 2023). The same effect was absent when participants were instructed to articulate the cue instead of imagery (Cochrane et al., 2018; Maljkovic & Nakayama, 1994). Collectively, the results suggest that imagery guides attention to task-relevant features, in line with the goal-driven account.

Previous studies often used single-colored features as the imagery cue and search items, making it unclear how the effect may scale up to real-world objects. A few studies have used more complex, real-world stimuli. Moriya (2018) used verbal imagery cues of fruit items (e.g., "lemon") strongly associated with a color (yellow) and separately, visual imagery cues of grayscale fruit items following search (see also Moriya, 2023). Similar results were found when searching for colored targets (faster reaction time in congruent than in incongruent trials). These findings suggest that imagery influenced attention to diagnostic features of the imagined objects. However, the search goal was a single-color feature of the target, not the whole target itself. In contrast, Monzel et al. (2021) used similar verbal imagery cues of real-world objects but instead presented photos of real-world objects during search, as opposed to simple-colored shapes. Participants were sampled from the aphantasic population (self-reported inability to use imagery, Zeman et al., 2015) and non-aphantasic population. Non-aphantasics were faster in congruent trials compared to aphantasics, suggesting that guidance was minimized (or non-existent) in the absence of imagery. Collectively, the results suggest that imagery elicited top-down effects in search tasks by activating visual representations that guide attention to matching items.

1.3.2. Current gaps

1.3.2.1. Does imagery actually guide attention? (Study 1)

Previous studies that implemented imagery instructions to search tasks relied on behavioral measures to suggest that imagery guides attention. Recent studies used differences in reaction time between congruent and incongruent targets as indirect supporting evidence (Cochrane et al., 2018; Cochrane, Ng, Khosla, et al., 2021, 2021; Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021; Monzel et al., 2021; Moriya, 2018, 2023). However, search involves distinct attentional processes (Figure 1.4), and behavioral responses result from the accumulation of these processes (Eimer, 2014). The behavioral effect of imagery alone cannot address whether imagery directly impacts selection. Given that imagery induced behaviorally relevant responses, could the impact of imagery be traced to effects on selection? The N2pc component is the fastest and most temporally precise ERP component associated with selection (Eimer, 1996; Luck, 2011). This component is widely used to test effects on early responses in attentional processing, measured by the magnitude and latency of the component (Luck & Kappenman, 2013). The N2pc initially was examined when searching for target features (e.g., basic shapes, Luck & Hillyard, 1994). Later studies showed similar findings when searching for whole items (Nako et al., 2014) and even categories of items (e.g., animate-inanimate items, Moon, He, et al., 2022; alphanumeric items, Wu et al., 2013). Together, the findings showed that the N2pc was sensitive to the search goal, following the goal-driven account. Importantly, if imagery also follows the goal-driven account, then imagery-related effects should be traced by this component. Currently, only

one study has tested this hypothesis. Cochrane, Townsend, et al. (2021) found that following color imagery, significant components were found when searching for congruent targets but not for incongruent targets, suggesting that imagery guided selection. Further research is warranted to determine whether effects on selection can be observed.

1.3.2.2. What about bottom-up processing? (Study 2)

Previous studies showed that bottom-up processing (e.g., pop-out effect) could be modulated by top-down processing (guidance by templates). Assuming that imagery elicits top-down effects on selection, how do these effects interact with those from bottom-up processing? Related imagery studies implemented an 'odd-one-out' search paradigm in which the target was distinguished from homogeneous nontargets by a salient feature (e.g., green target among red nontargets, Cochrane, Wang, Pratt, et al., 2021, 2021). If pop-out effects are insensitive to top-down influence, then imagery should not impact search (i.e., no difference between congruent and incongruent trials). However, imagery-related effects were observed based on the task-relevant features and not on the salient feature (Cochrane, Townsend, Lee, et al., 2021), again supporting the goal-driven account. To further examine the interaction between top-down imagery and pop-out effects, I implemented the asymmetric search design, which was widely used to compare search efficiency based on the presence or absence of salient features (Wolfe, 2001). If imagery guides attention, then search efficiency should increase in search tasks with present salient feature (Figure 1.3, left panel), consistent with previous studies that showed top-down effects modulated bottom-up processing (Eimer & Kiss, 2010). For

search tasks with absent salient feature (Figure 1.3, right panel), perhaps imagery also would increase search efficiency, following the goal-driven account. On the other hand, without a pop-out effect, perhaps search may be conducted in a serial manner, following the stimulus-driven account.

1.3.2.3. Does imagery impact post-selection? (Study 3)

As previously mentioned, it was assumed from behavioral effects that imagery guides attention. However, it may be possible that these effects stemmed from exclusively the post-selection stage, or perhaps both selection and post-selection. Isolating these effects from specific stages in search has been a non-trivial matter in the attention literature (Eimer, 2014). The N2pc and SPCN components are two distinct neural markers that can be used to distinguish effects on selection and post-selection. Previous studies differentiated the effects by implementing search tasks that differed in recognition demands (Mazza et al., 2007). Tasks that only required selection but not post-selection elicited the N2pc components but not SPCN components. In contrast, tasks that required both selection and post-selection elicited both N2pc and SPCN components (Berggren & Eimer, 2018; Grubert et al., 2017; Mazza et al., 2007). Several imagery-search studies that showed behavioral effects used tasks that required both selection and post-selection (e.g., Cochrane, Townsend, et al., 2021), which further raises the question of how these effects arise from attentional processing. Further research is needed to discern such effects to specific attentional processes.

1.3.2.4. Individual differences in imagery skills (Study 4)

People naturally vary in their imagery skills (Marks, 1973), such that some can 'see' an actual image vividly and others see a less vivid image or nothing at all (Zeman et al., 2020). Such anecdotal variation in vividness was one suggestion for the discourse behind the depictive vs. propositional debate (Reisberg et al., 2003). Several studies used subjective measures of vividness (M. Runge et al., 2015; M. S. Runge et al., 2017), such as the Vividness of Visual Imagery Questionnaire (VVIQ; Mark, 1973). Later fMRI studies found a correlation between the VVIQ scores and neural activity patterns in the visual cortex (Cui et al., 2007). Related studies observed a relationship between the size of V1 and the vividness of imagery (Bergmann et al., 2016). MVPA techniques were used to demonstrate overlap in activity patterns between imagery and perception (Albers et al., 2013; Stokes et al., 2009; Thirion et al., 2006). Furthermore, these overlaps were correlated with VVIQ scores, such that those with higher vividness displayed stronger overlaps in hemodynamic responses (Albers et al., 2013; Lee et al., 2012). Using similar techniques, other studies showed that trial-by-trial vividness ratings, as opposed to overall vividness scores like from VVIQ, also predicted neural overlaps between imagery and perception (Dijkstra et al., 2017).

Putting it together, previous studies examined the vividness of imagery using subjective and objective measures and found significant results using data-driven techniques to infer neural correlates of imagery. These findings were used to support the depictive theory, but it is unclear whether the variation in vividness implicates any limitation on other processes. Some studies found evidence to warrant such concern. For

example, previous studies measured sensory strength of imagery using the binocular rivalry task (Pearson et al., 2008). Participants were given an imagery cue of a color (e.g., green) and then were rapidly presented with one of two separate colors for each eye (e.g., red for left eye, green for right eye). They responded to the color that they perceived, with the prediction that perception would be biased to the previous cue. At the end of the trial, participants reported the vividness of the imagined color. The number of trials with responses in favor of the cue reflected the sensory strength. Several findings showed stronger sensory strength for trials with higher vividness ratings than lower ratings (Keogh & Pearson, 2017, 2018, 2021; Pearson et al., 2011), suggesting that perceptual awareness is biased to imagery but that this bias is limited to vividness. Additionally, some studies reported correlation between the sensory strength of imagery and performance in cognitive tasks, such as visual working memory (Jacobs et al., 2018; Keogh et al., 2021). Together, the results warrant further research into whether vividness systematically drives other processes.

Much less is known about whether vividness of imagery imposes limitation on attentional processing. Current models of attention often assume a uniform ability to generate visual representations (Wolfe, 2021), potentially limiting explanatory power. To address this concern, I explored the top-down effects of imagery on search based on trial-by-trial vividness ratings. First, previous fMRI studies used MVPA techniques to examine the neural correlates of imagery based on vividness (Dijkstra et al., 2017). For search, there consists of multistage processes (Figure 1.4) that are temporally distinct (Eimer, 2014), and as opposed to fMRI, EEG has high temporal resolution appropriate for

measuring attention-related responses (Woodman & Luck, 1999, 2003). Recent studies adapted MVPA techniques from fMRI studies to fit EEG data (Bae & Luck, 2018, 2019), such that neural activity patterns could be distinguished based on experimental conditions across timepoints. If activities during search are systematically driven by imagery and if imagery-related activities differ based on vividness, then perhaps MVPA techniques could detect differences in activities during search based on vividness ratings. Second, some imagery-search studies examined the typical congruent-incongruent effects with color imagery, with larger differences when vividness ratings were high than when ratings were low (Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021). I explored these effects with more varied stimulus sets and a larger dataset.

1.4. Outline of dissertation

This dissertation presents three experimental studies (Studies 1, 2, and 3) and one analytical study (Study 4) that addressed the broad question of how imagery impacts attentional processing in visual search. Conventional theories of imagery and attention were siloed but provided important accounts for how imagery may be understood as a top-down process. I hypothesized that imagery elicits top-down effects on attentional processing, such that imagery would guide attention by selecting imagery-matching items. To test this hypothesis, a visual search task with EEG was necessary to collect temporally precise measures related to selection (N2pc) and post-selection (SPCN), in addition to traditional behavioral measures widely used to measure overall search performance (average reaction time) and search efficiency (slope reaction time).

Study 1 (Chapter 3) tested the hypothesis that imagery guides attention to imagery-matching items. Three EEG experiments (Experiments 1, 2, and 3) were conducted. Experiment 1 set up the search paradigm for later experiments and was inspired by previous imagery-search paradigms (Cochrane, Townsend, Lee, et al., 2021). Imagery instructions were imbedded between two search tasks per trial. The search task following imagery displayed search items with either a target that matched with imagery (congruent) or a target that mismatched but the nontargets matched (incongruent). Experiment 2 compared the effects in Experiment 1 with a control condition, in which participants were given an alternative mental task in lieu of imagery. Experiment 3 compared the effects in Experiment 1 by adding a neutral trial type in which neither the target nor nontargets matched.

Study 2 (Chapter 4) tested the hypothesis that pop-out effects would be modulated by the top-down effects of imagery. A behavioral experiment without EEG (Experiment 4) and an EEG experiment (Experiment 5) were conducted. For both experiments, I adapted the search paradigm from Experiment 1 based on asymmetric search designs (similar to Figure 1.3). The purpose of this design was to compare search efficiency based on the presence of a salient feature (Q among Os; pop-out effect) and the absence of it (O among Qs; no pop-out). Experiment 4 examined relevant behavioral effects based on the trial types (congruent and incongruent) and the salient feature (present or absent). Experiment 5 aimed to replicate the behavioral effects and also examined relevant neural effects using the N2pc component.

Study 3 (Chapter 5) tested the hypothesis that imagery impacts selection and/or post-selection stage of search. I predicted that selection would be influenced, as predicted by the goal-driven account, but it was unclear whether there would be any impact on post-selection. One EEG experiment (Experiment 6) with two search tasks was conducted. The two tasks were identical in design but differed in search instructions. A 'locate' search task instructed participants to respond to the location of the target. A 'identify' search task instructed participants to respond to a feature of the target. Both tasks had selection demands, but only the latter had recognition demands. These differences were necessary to differentiate the N2pc and SPCN components, which were commonly used to mark selection and post-selection separately (Mazza et al., 2007).

Study 4 (Chapter 6) tested the hypothesis that vividness (subjective ratings of imagery) modulated the top-down effects of imagery. Two exploratory analyses were conducted with a collective dataset from previous experiments. In each experiment, vividness rating of imagery was collected on a trial-by-trial basis. These ratings were used to organize trials in the EEG and behavioral data based on how strongly imagery was experienced. For the EEG data, trials were organized between highly and less vivid ratings based on median-split responses per participant. Previous fMRI studies used MVPA to examine the spatial dynamics of imagery based on vividness (e.g., Dijkstra et al., 2017). Here, I used MVPA using procedures appropriate for EEG data (Bae & Luck, 2018, 2019) to examine the temporal dynamics of imagery during search. If neural responses were influenced by imagery, then perhaps these responses may be distinguishable based on measures of imagery like vividness. For the behavioral data

(averaged reaction time), trials were organized based on individual vividness ratings. Previous imagery studies showed that differences in search performance between congruent and incongruent trials were larger for trials with higher vividness ratings than for trials with lower ratings (Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021). Here, I conducted a linear mixed-effects analysis to test differences in reaction time based on vividness ratings. If vividness predicts search performance, then vividness should predict behaviorally relevant responses according to the trial-by-trial ratings.

Chapter 2: EEG

This chapter summarizes the experimental protocols for collecting and handling EEG data. EEG is a non-invasive method that records electrophysiological signals from the scalp (Luck, 2014). Extracranial EEG was used to examine the impact of imagery on time-sensitive measures of attentional processing. It is particularly well-suited for my goal of extracting temporally precise data given that attentional processing occurs within milliseconds (Woodman & Luck, 1999, 2003). Behavioral measures alone are insufficient for such a goal. EEG data was used to extract data related to the N2pc component (Studies 1, 2, and 3), the SPCN component (Study 3), and classification accuracy (Study 4). Protocols were based on guidelines provided in Luck (2014).

2.1. Description of EEG

2.1.1. What is an EEG signal?

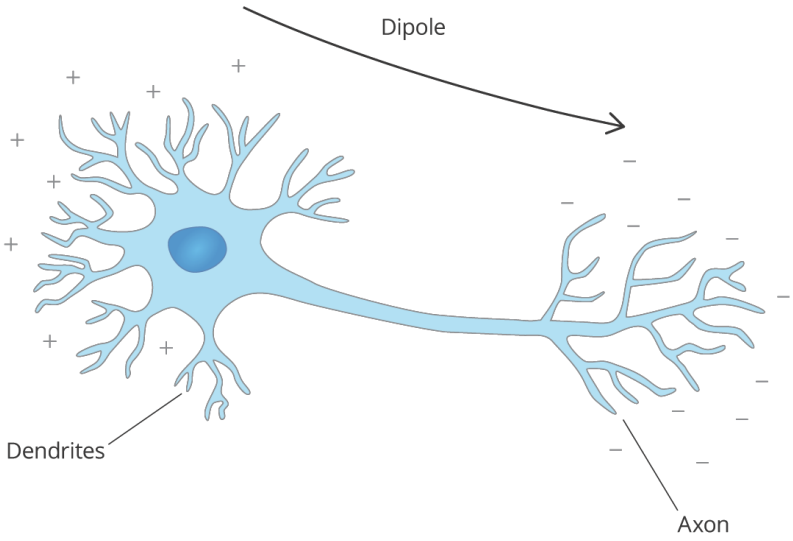
Underneath the scalp, there are billions of brain cells called neurons, each with three basic structures. The dendrites are short, branched extensions of the neuron that receive input from neighboring neurons (Stuart et al., 2007). The axon is a long nerve fiber that transmits electrical signals known as action potentials (Debanne, 2004). At the end of the axon, axon terminals release chemical messengers called neurotransmitters, which bind to receptors on neighboring neurons (Südhof & Malenka, 2008). These neural interactions take place at synapses, the junctions between the presynaptic neuron (the one releasing neurotransmitters) and the postsynaptic neuron (the one receiving the neurotransmitters via receptors). Electrical activity in the brain can be measured in two ways. One is through the movement of positive and negative ions across the neuron's

membrane, or the outer layer of neurons. The charged ions pass selectively through protein channels that open and close. The difference in electrical charge between the inside and outside of the neuron is called the membrane potential (Carnevale & Hines, 2006). When positive ions flow into the neuron, the membrane potential becomes more positive, potentially leading to an action potential. However, the electrical changes in individual neurons are too small to be detected by EEG (Luck, 2014).

When neurotransmitters bind to the receptors of the post-synaptic neuron, ion channels on the dendrites open, allowing ions to enter the neuron, which leads to changes in the post-synaptic potential (PSP). An excitatory postsynaptic potential (EPSP) occurs when positive ions move into the neuron, causing a positive change in the PSP. In contrast, an inhibitory postsynaptic potential (IPSP) occurs when negative ions moves into the neuron, leading to a negative change in the PSP (Carnevale & Hines, 2006). When positive ions flow inward during an EPSP, the space outside the dendritic membrane becomes more negative, while the space at the opposite end of the neuron becomes relatively more positive. Conversely, during an IPSP, the inflow of negative ions causes the outside dendritic space to become more positive, and the opposite end becomes more negative. This ionic movement creates a dipole (Niedermeyer, 1996), which refers to the separation of charges between the two ends of the neuron (Figure 2.1). The dipole generates an electrical current that flows from the positive end to the negative end. The direction of this current is referred to as the polarity. An EEG signal reflects the sum of PSP across millions of neurons based on their polarity.

Figure 2.1

Dipole of a neuron



Note. Example of dipole from IPSP. +/- reflect the charge of the PSP at each end of the dipole (Source: EEGEdu, https://eegedu.com/).

However, EEG signals do not capture every electrical activity across the brain. If the orientation of the dipole (i.e., the direction of the polarity) across neurons is inconsistent, the PSPs across neurons will cancel each other out. This is particularly true for neurons located deeper within the brain, such as those in the brainstem, where the orientation of dipoles is more variable (Niedermeyer, 1996). In contrast, pyramidal cells, which are found throughout the cerebral cortex, have a consistent orientation, with their dendrites aligned perpendicular to the cortical surface. This alignment allows for the summation of PSPs, which produces a measurable electrical charge that EEG can detect through scalp electrodes (Buzsáki et al., 2012). The rapid nature of changes in potential,

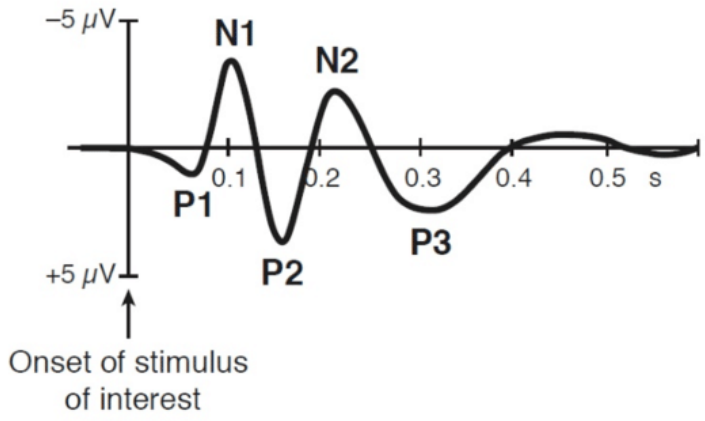
occurring within milliseconds, makes EEG particularly advantageous for capturing the precise timing of neural events. This high temporal resolution is a significant advantage over neuroimaging techniques, such as fMRI, which have slower temporal dynamics. EEG's ability to track neural signals in real time makes it especially useful for studying cognitive tasks that involve brief trials, such as visual search, where attentional selection occur rapidly (Luck, 2014).

2.1.2. What are ERP components?

The general aim of an ERP study is to capture specific neural responses during a cognitive event. These changes in electrical potential are termed 'event-related' because they are directly associated with and driven by a relevant cognitive process. However, individual EEG trials contain a mix of signal and noise, the latter of which is unrelated to the cognitive event. To isolate the signal, EEG data are averaged across multiple trials, resulting in a clean ERP waveform that cancels out the random noise. An ERP waveform is composed of distinct peaks and troughs (Figure 2.2), known as ERP components, which are cognitively meaningful (Luck & Kappenman, 2013). These components are named based on their electrical polarity—positive (P) or negative (N)—and the order in which they occur (e.g., P1 for the first positive peak; N1 for the first negative peak). ERP waveforms are time-locked to the onset of a stimulus, meaning the timing of each component reflects neural responses at specific moments following the stimulus presentation.

Figure 2.2

Sample ERP waveform



Note. From Teixidó et al. (2018). Waveform time-locked to stimulus onset. Bolded labels refer to ERP components. P/N refer to positive and negative components. Numbers refer to the order of peaks or toughs.

2.1.3. Why use ERP components here?

The goal of the present studies was to examine top-down effects of imagery on attentional selection and post-selection. Both of these cognitive events occur rapidly (<600 ms), which cannot be adequately captured by standard behavioral measures. ERP components are particularly well-suited for this investigation because they offer precise temporal information about neural processes related to attention. Unlike neuroimaging techniques such as fMRI, which provide excellent spatial resolution but lack the temporal precision required for studying rapid cognitive events, ERP offers the millisecond-level resolution necessary for tracking the fast dynamics of attention (Luck, 2014). Furthermore, due to logistical and resource constraints, neuroimaging techniques were beyond what was feasible in my lab. Since all experiments involved human participants,

EEG was the ideal method for capturing electrophysiological responses safely and non-invasively, without the need for incisions or other invasive procedures. This made it possible to observe the brain's activity during attentional tasks while minimizing risk to participants.

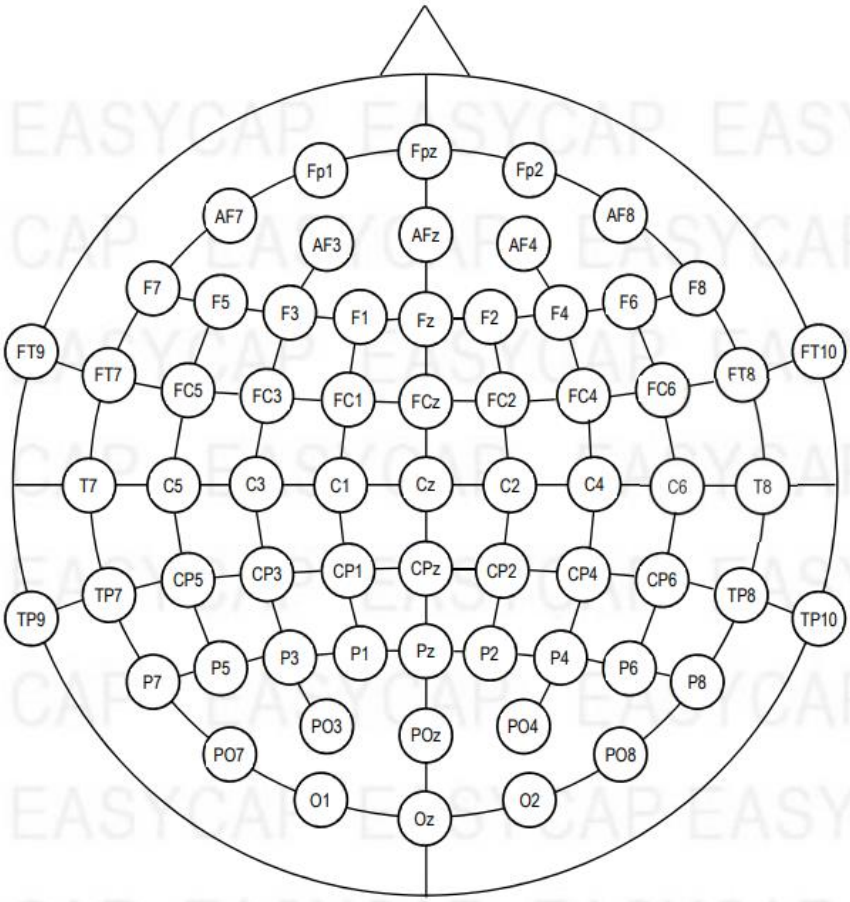
2.2. EEG set-up

2.2.1. Electrodes

EEG was recorded using Brain Products' actiCHamp active 32-electrode system with the international 10/20 placement (Figure 2.3). One end of the electrode set consists of electrodes that were clamped onto EasyCap electrode caps located at Fp1/2, Fz, F3/4, F7/8, FC5/6, FC1/2, FT9/10, C3/4, T7/8, CP1/2, Pz, P3/4, P7/8, PO7/PO8, Oz, O1/2. CP5/6, Cz. Two of the electrodes (FT9/10) were used to record left and right horizontal eye movements, labeled HEOGL and HEOGR, respectively. Two other electrodes (Fp1/2) were used to monitor eye blinks. Additional two electrodes were placed on the earlobes and used as reference electrodes. A final ground (GND) electrode was placed on the cap approximately at the center of the forehead. The other end of the electrode set consists of a flat-ribbon cable that connects to the actiCHamp base module. This module amplified signals from each electrode and digitized the signals onto a computer. The module was charged by the actiPOWER. The GND electrode was connected to the GND socket of the module.

Figure 2.3

Easycap electrode layout with 64 channels



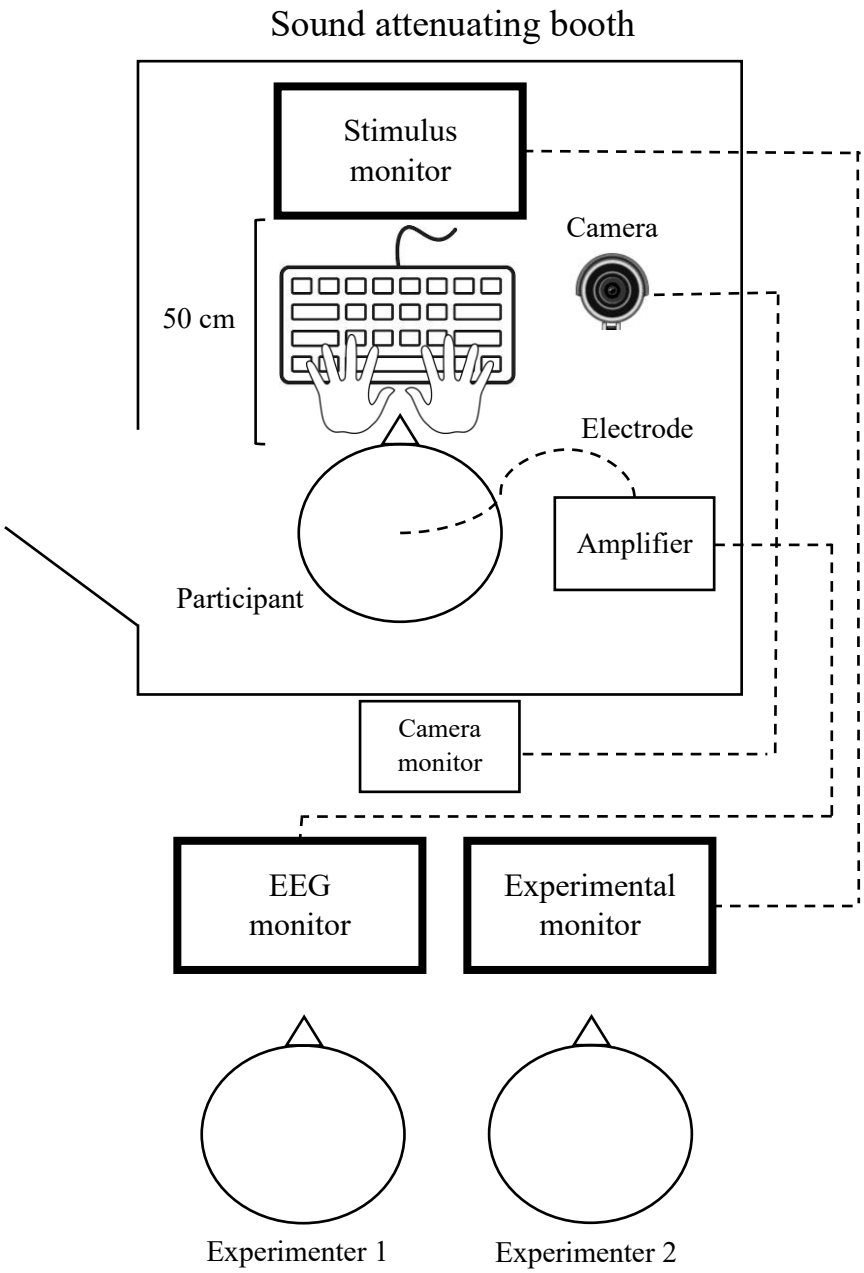
Note. Used 32 channels located at Fp1/2, Fz, F3/4, F7/8, FC5/6, FC1/2, FT9/10, C3/4, T7/8, CP1/2, Pz, P3/4, P7/8, PO7/PO8, Oz, O1/2. CP5/6, Cz.

2.2.2. Experimental room

Figure 2.4 shows a schematic of the experimental room. Participants were seated comfortably inside a sound attenuating booth to minimize outside noises. On the right of the participant, the amplifier was supported by an adjustable mounted stand. The flat-ribbon cable of the electrode set and the GND electrode were attached to the amplifier without blocking the participants' field of vision. The amplifier was connected to the EEG monitor outside of the booth to digitize the signals. The EEG monitor was used to operate the software for recording EEG data and to observe EEG-related artifacts. In front of the participant, a 24-inch Dell stimulus monitor was positioned approximately 50 cm away from the participant's eyes. A keyboard was placed at a comfortable distance from the participant, minimizing movement. The stimulus monitor was connected to a separate experimental monitor outside of the booth. The experimental monitor was used for one experimenter to operate the tasks of each experiment and project these tasks onto the stimulus monitor (1024 × 768 resolution). Next to the stimulus monitor, a video camera was placed. A continuous video (not recorded) was projected onto a monitor outside the booth for experimenters to view during the experimental run. This monitor was used to check for movement or frequent blinking that could have disrupted the EEG signals.

Figure 2.4

Experimental room



Note. Schematic of the experimental room.

2.2.3. Procedure

Figure 2.5 shows the materials used in each session. Two undergraduate research assistants were assigned to each session. First, to select an appropriately sized EasyCap, a measuring tape was wrapped around the participant's head, between the center of the eyebrows and the inion (most prominent projection at the back of skull). This technique was to estimate the head circumference. The cap was carefully placed on the participant's head, and then, the electrodes including the GND were clamped onto the relevant electrode sites using the 10/20 placement. Each earlobe was gently scrubbed with alcohol wipes to remove oil or debris, and then the reference electrodes were placed on the earlobes using double-sided stickers. If participants wore earrings, they were asked to remove them and place them aside.

Plastic syringes were filled with conductive, electrolyte gel and capped with a blunt needle. Each electrode had an opening that fit the needle. The scalp has natural oil and layers of dead skin cells that can impede the flow of electrical signals from the brain to electrodes. A small trace of gel was inserted into each electrode to increase conductivity between the scalp and the electrode. Next, the electrode set and the GND electrode were connected to the amplifier. The signals were digitized onto the EEG monitor. Brain Vision Recorder 2.0 was used to monitor electrodes with poor signals before recording the experimental run. The software required only clicking buttons and has a simple user interface that the experimenters familiarized with. Electrodes with signals greater than 10 Hz were considered to have high impedance and was correct by applying additional gel. Once each electrode showed acceptable signals, the

experimenters instructed the participants with the procedures of the task. The experimenters emphasized minimizing eye-blinks before giving a response and reducing physical movement to not interfere with the signals. The participants were made aware that the camera was on to monitor movement but was not recording.

All experiments included a practice run of the task for participants to get familiar with the procedures. Before the experimental run began, experimenters clicked the record button and monitored for excessive eye-blinks or physical movements. Additionally, the EEG monitor displayed live markers of correct and incorrect responses for experimenters to monitor. The experimental run consisted of multiple breaks between each block. During this time, experimenters gave additional clarifying instructions if incorrect responses were given frequently. The accuracy of each block was displayed during the break for participants to self-monitor their progress and ask additional questions.

After completing the experimental run, participants carefully removed the caps from the participants' heads and assisted the participants in cleaning off the gels on their hair using the lab's sink. Once cleaned, the participants were given a debriefing form giving a general overview of the experiment and then excused from the session. To clean the electrodes, an electrode remover tool (clear cylinder objects next to toothbrush) was used to detach each electrode. The caps and electrodes were washed with baby shampoo. The gels from each electrode were scrubbed using a toothbrush and then washed repeatedly. Once cleaned, the caps and electrode set were placed in a drying rack.

Figure 2.5

Experimental materials



Notes. Materials used in each EEG session.

2.2.4. Steps for preprocessing

Data was recorded with 500-Hz sampling rate and DC-40 Hz bandpass. I used left-earlobe reference online and re-reference offline to averaged earlobes (Nunez & Srinivasan, 2006; Luck, 2014). The continuous EEG data were preprocessed using the toolbox EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) on MATLAB 2021a. Steps for preprocessing were adapted from the recommended protocols in ERP CORE (Kappenman et al., 2021). First, a 0.1 Hz high-pass filter (12 dB/octave) was applied to the continuous data. Bipolar HEOG (HEOGL – HEOGR) signals was computed to use for artifact rejection. Second, independent component analysis (ICA) was performed on the filtered data. Components associated with ocular artifacts (horizontal eye movements and eyeblinks) were removed based on visual

inspection. Corrected bipolar HEOG (HEOGL – HEOGR) signals were computed from the ICA-corrected data (separate from HEOG computed before correction). Third, the continuous data were segmented from -100 to 600 ms relative to the search array onset (baseline correction: -100 to 0 ms), consistent with segments used from studies in our lab (e.g., Ferguson et al., 2021; Moon, He, et al., 2022; Moon, Zhao, et al., 2022; Wu et al., 2015, 2017). Only segments with correct responses were included in the final analyses. Fourth, channels with excessive levels of noise, excluding electrodes of interests (P3, P4, P7, P8, PO7 and PO8), were visually inspected and interpolated using EEGLAB's spherical interpolation algorithm. I then applied artifact rejection criteria using individualized thresholds based on visual inspection of each participant's data (Kappenman et al., 2021). For each participant's data, trials were excluded if there were evidence of the following: large voltage excursions in any channels, large eye movements (greater than 4° of visual angle) in the corrected HEOG, eye movements (greater than 2° of visual angle) in the uncorrected HEOG (before ICA), and eyeblinks in Fp1 and Fp2. Following standard protocols from my lab (e.g., Moon, Zhao, et al., 2022; Wu et al., 2015, 2017), I excluded participants who had more than 30% of their total trials removed after artifact rejection.

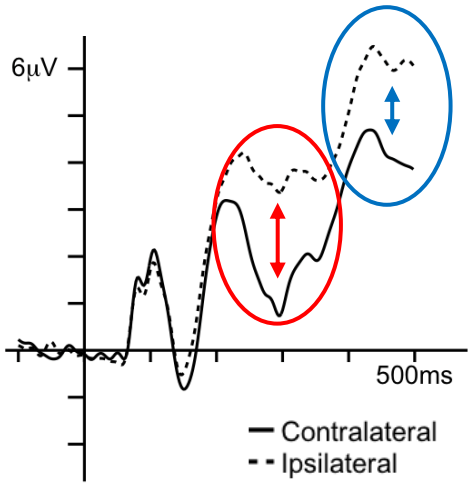
2.3. ERP components

Figure 2.6 shows an example of the N2pc and SPCN components. The two lines reflect averaged ERP waveforms recorded from the electrodes of interests. The waveforms are time-locked to the onset of stimulus presentation. The N2pc and SPCN were measured from the difference in signals between the ipsilateral and contralateral

electrodes, based on the lateralized position of the attended object (Luck, 2012). For example, if the observer attended to a target on the left visual field, the components would be stronger for the contralateral electrodes than for the ipsilateral. Importantly, the magnitude and the timing of the components were used to examine effects on a cognitive process (Luck, 2014). In imagery-search studies, a critical manipulation was the trial type, with effects measured by comparisons between congruent and incongruent trials (Cochrane et al., 2018; Cochrane, Ng, Khosla, et al., 2021; Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021; Monzel et al., 2021; Moriya, 2018). Therefore, all experiments included trial type as the main independent variable.

Figure 2.6

Example of N2pc and SPCN components



Note. N2pc (red) and SPCN (blue) components from the same waveforms. Contralateral and ipsilateral waveforms (time-locked to stimulus presentation) were based on lateralized position of the target.

2.3.1. N2pc

The N2pc component was used to investigate top-down effects of imagery on signals related to selection. This component is characterized by a negative deflection approximately 200 ms (Figure 2.6, red circle) after the stimulus onset (Luck, 2012, 2014). It is observed over the posterior scalp, contralateral to the location of the attended object. This component is associated with the selection of target items in a visual search task (Eimer, 1996; Luck, 2011), making it an ideal neural marker for observing top-down effects.

EEG continuous data were segmented 100 ms before and 600 ms after the onset of the post-imagery array. Segments for the post-imagery array were split by congruent and incongruent trials. All segments were baseline-corrected from 100 ms to onset. The ERP was recorded from EEG electrodes P3, P4, P7, P8, PO7 and PO8, contralateral and ipsilateral to the target location. Trials in which the target appeared above or below the fixation cross were excluded. Segments were averaged and pooled across electrodes P3, P7, and PO7 and across electrodes P4, P8, and PO8. The grand average was computed for each pooled electrode and then subtracted to get the grand average difference (contralateral minus ipsilateral waveform) waveform.

2.3.1.1. N2pc mean amplitude

The N2pc mean amplitude was used to test between-condition effects of imagery on the magnitude of the component. The amplitude was computed from 200 to 300 ms after stimulus onset, a standard approach used in our lab (Moon, Zhao, et al., 2022; Wu et al., 2015, 2017) and from critical studies (Brisson & Jolicœur, 2007; Kiss, Van Velzen, et

al., 2008). Significant N2pc components were determined by a one-sample *t*-test compared to zero.

2.3.1.2. N2pc onset latency

The N2pc onset latency was used to test between-condition effects of imagery on the timing of the component. The onset was measured using jackknife latency analysis, following well-established procedures (Kiesel et al., 2008). Briefly, one participant's data were excluded at a time, and the grand average ERP waveforms were repeatedly calculated for each subsample. For each trial type, latencies from the difference waveform were found at timepoints that exceeded 50% of the N2pc peak amplitude. This approach reduces variability and provides a more stable estimate of the onset latency, as opposed to extracting onset data per participant (Luck, 2014)

2.3.2. SPCN (Study 3, Chapter 5)

The SPCN (sustained posterior contralateral negativity) component was used to distinguish the top-down effects of imagery between the selection and post-selection stage. Current imagery-search findings are limited to behavioral effects (Cochrane et al., 2018; Moriya, 2018), which cannot be used to address any potential effects on attentional processing. This component is characterized by a negative deflection approximately 400 ms (Figure 2.6, blue circle) after the stimulus onset (Berggren & Eimer, 2018; Grubert et al., 2017; Mazza et al., 2007) and shares the same characteristics of N2pc (negative deflection) except that the duration lasts ~200 ms (Joliocoeur et al., 2008). This component was extracted from the same grand average difference waveform as the N2pc.

2.3.2.1. SPCN mean amplitude (Study 3, Chapter 5)

The SPCN mean amplitude was measured from the same waveforms as the N2pc. Following previous methods (Berggren & Eimer, 2018; Grubert et al., 2017; Mazza et al., 2007), I averaged the signals from 400 to 600 ms after stimulus onset.

2.4. Multivariate pattern analysis (MVPA)

Previous fMRI studies used MVPA techniques to examine the neural correlates of imagery based on vividness (e.g., Dijkstra et al., 2017). Here, I used a similar technique adapted for EEG data (Bae & Luck, 2018, 2019) to examine differences in activity patterns during search based on vividness. The pre-processed EEG data from all studies, except for Experiment 2's control group, were analyzed collectively to increase power.

The protocols were adapted from Bae & Luck (2018) and have been used in my own research (Moon, Zhao, et al., 2022). For each participant, congruent and incongruent trials were analyzed separately to measure one accuracy score for each trial type. Segments were bin-organized by individual median-split vividness rating (highly vivid vs. less vivid). A median split, rather than individual vividness rating, was used to include a sufficient number of trials per condition. Support vector machines (SVMs) were trained on activity patterns from all channels (excluding HEOG, Fp1, Fp2, and reference electrodes) at each time point of the segment. Using a 3-fold cross-validation procedure (10 iterations), the segments for each bin were randomly divided into equal thirds and then averaged. Two of the thirds were randomly selected for training, while the last third was used for testing. For each iteration (= 10), the procedure was repeated, with new randomized subsets of averaged trials (split into thirds) randomly selected for training

and testing. The training data was labeled as highly vivid or less vivid. One classifier was trained for binary classification, between highly vivid and less vivid trials. During testing, the SVMs predicted the median split of vividness for the unused (last third) data without labels. For each time point, classification accuracy was calculated as a percentage of correctly predicted labels, collapsed across all 10 iterations (chance level at 1/2 = 50%), for a total of 60 decoding attempts (2 bins x 3 cross validations x 10 iterations).

2.4.1. Classification accuracy (Study 4, Chapter 6)

The classification accuracy reflected the proportion of trials in which the SVMs predicted the median split (highly vivid vs. less vivid). For each participant, the accuracy was measured at each time point from the stimulus onset. Accuracies were calculated for each trial type separately. A final classification accuracy for each trial type was collapsed across all participants.

2.5. Adequate sample size

Each experiment targeted a sample size of 20 participants with two exceptions. In Experiment 2, a between-subjects design was used with one experimental group (n = 20) and one active control group (n = 20), resulting in a total sample size of 40 participants. Experiment 4 (no EEG) included 32 participants. Experiment 1 was conducted before finalizing a target sample size for later experiments. Previous studies consistently showed faster reaction times for congruent targets compared to incongruent targets (Cochrane et al., 2018; Cochrane, Ng, et al., 2021a; Cochrane, Townsend, et al., 2021b; Cochrane, Wang, et al., 2021; Monzel et al., 2021; Moriya, 2018). A previous study reported that six participants were sufficient to detect this effect (power = .90, alpha = .05; Cochrane et al.,

2021). My own preliminary data showed behavioral effect ($p < .001$, effect size ~ .90) with 52 participants. Experiment 1 also revealed this effect ($p < .001$, effect size ~ .90) with 20 participants.

Importantly, Experiment 1 aimed to demonstrate the N2pc component, which was expected to be less robust than reaction time due to noise in neural data. Previous studies in our lab (Ferguson et al., 2021; Moon, He, et al., 2022; Moon, Zhao, et al., 2022; Wu et al., 2013, 2015, 2017) and others (Eimer, 1996; Luck & Hillyard, 1994; Woodman & Luck 1999) with the same or smaller sample size revealed significant N2pc components. Experiment 1 revealed significant components ($p = .002$, effect size ~ .55) with 20 participants. Given these results, the same sample size was deemed appropriate for later experiments.

In Experiment 2, practical challenges associated with EEG data collection limited feasibility. While this sample size was likely sufficient for detecting between-condition effects, which was the primary goal, it may not have been adequate for detecting between-group effects. The potential for reduced power in between-group comparisons was acknowledged as a limitation, and any interpretation of such comparisons were made cautiously. In Experiment 4, there were no practical constraints as EEG was not used, allowing for a larger sample size. A minimum of 30 participants was decided a priori.

2.6. Adequate number of trials

The number of trials per condition (i.e., trial types) was selected with the goal of examining 1) significant N2pc components and 2) between-condition N2pc effects. For the former, pivotal studies suggested that a modest number of trials is sufficient with a

relatively small sample size (Huffmeijer et al., 2014; Rietdijk et al., 2014; Thigpen et al., 2017). For example, a previous simulation study showed significant components (not specific to N2pc but computed in a similar manner) with 30 trials per condition and 12 participants (Boudewyn et al., 2018). Each EEG experiment here consisted of 20 participants (40 for Experiment 2).

For the latter (between-condition), some challenges were met. First, because imagery inherently is a covert task, there were concerns over whether participants would be compliant. It was assumed that fatigue might make participants less compliant, and the experimental sessions already took two hours. Second, the numerical value of the N2pc (e.g., mean amplitude) typically is small, as the component is derived from the difference between two waveforms (Luck, 2014). Therefore, it was expected that the difference between trial types also would be small. For example, Experiment 1 revealed a significant difference of 1 μV with the N2pc mean amplitudes (p = .02; effect size ~ .40) between trial types. From the same simulation study, effects were observed with the same difference (power = .90) with 45 trials per condition and 20 participants (Boudewyn et al., 2018). Additionally, to test for between-group effects (important for Experiment 2), the study found significant effects (power ~ .80) with 45 trials per condition and 16 participants per group. Following these parameters, a minimum of 50 trials per trial type, after artifact rejection (70% of total trials), were expected for each participant. Approximately 160 to 200 total trials were used across experiments.

Chapter 3: Top-down effects of imagery on early attentional processes

3.1. Abstract

Recent studies showed behavioral advantages (faster reaction time) for imagery-matched targets but disadvantages (slower reaction time) for imagery-mismatched targets. However, these findings have predominantly relied on behavioral measures, whereas visual search relies on distinct attentional processes that precede behavior. Study 1 examined how mental imagery influences these early stages of attentional processing by measuring the N2pc component (neural marker for attentional selection). Across Experiments 1, 2, and 3, participants completed visual search tasks in which they searched for one pop-out target object among homogenous nontarget items. Between search trials, participants were instructed to imagine the previous target object and then rate the vividness of their mental images. Subsequent search trials presented either imagery-matched targets among mismatched nontargets (congruent trials) or imagery-mismatched targets among matched nontargets (incongruent trials). In all three experiments, behavioral effects were replicated, with faster reaction times in congruent than in incongruent trials, and ERP findings showed larger N2pc components in congruent than in incongruent trials. The neural effects were minimized when participants engaged in task-irrelevant mental tasks instead of mental imagery (Experiment 2) and were partially attenuated in neutral trials where neither targets nor nontargets were matched to imagery (Experiment 3). Together, these results suggest that imagery directly influences early attentional processes in a search setting and serves as a useful strategy for guiding attention.

3.2. Introduction

Early research in the 1970s sparked a debate on whether mental imagery is structured as a visual representation, as proposed by the depictive theory (Kosslyn, 1975, 1981, 2005; Kosslyn & Alper, 1977), or strictly as symbolic descriptions akin to language, as suggested by the propositional theory (Pylyshyn, 1973, 1981, 2002, 2003). The development of fMRI techniques in the 1990s provided critical evidence in support of the depictive theory, with findings showing that mental imagery activated neural patterns in the visual cortex, similar to those observed during perception (Pearson, 2019). Given this overlap between imagery and perception (Dijkstra et al., 2019), recent studies have explored whether these visual representations influence performance in visual search tasks (Cochrane & Milliken, 2019, 2020; Wantz et al., 2015).

One line of research has focused on search paradigms that incorporate imagery instructions. For example, participants may be asked to imagine a specific color (e.g., "blue") before performing a visual search task where they must locate a target item among distractors of varying colors. Studies have consistently found that reaction times are faster when the imagined color matches the target (congruent trials) compared to when it does not (incongruent trials; Moriya, 2018, 2023). In another study, Cochrane et al. (2021a) showed that reaction times were faster when the imagined color matched the target, with the effect being more pronounced with imagery cues compared to perceptual cues. Additionally, Monzel et al. (2021) found that individuals without aphantasia (poorly vivid imagery or none at all; Zeman et al., 2015) were quicker to locate targets in congruent trials, suggesting that imagery enhances attention in visual search tasks—a

process that may be impaired in individuals with aphantasia. Collectively, these findings suggest that imagery activates visual representations that guide attention toward matching items, thereby influencing search performance.

The effects of imagery on visual search suggest that it may enhance the selection of task-relevant features, functioning as a top-down process that guides attention based on internally generated representations. However, this idea has rarely been tested directly. The N2pc ERP component, known for its sensitivity to top-down attentional selection, offers a promising measure to explore the early effects of imagery on attention (Eimer, 1996; Luck & Hillyard, 1994; Luck, 2011). Prior studies have consistently shown robust N2pc components during visual search for various stimuli, including basic features (Luck & Hillyard, 1994), whole objects (Nako, Wu, Smith, et al., 2014), and object categories (Moon et al., 2022; Wu et al., 2013). The goal-driven account of attention states that attentional templates, which are internally generated representations of target features, exert top-down control over selection (Desimone & Duncan, 1995; Duncan & Humphreys, 1989). Seminal work examined this top-down control with EEG (Snyder & Foxe, 2010; Jokisch & Jensen, 2007), fMRI (Chawla et al., 1999; Giesbrecht et al., 2006; Kok et al., 2012; Serences et al., 2004; Soon et al., 2013), and single-cell recordings (Chelazzi et al., 1993, 1998). However, the role of imagery in this process remains underexplored.

Given that imagery activates internal representations according to the depictive theory, it is plausible that the N2pc component could capture the top-down effects of imagery on attentional selection. Behavioral studies have shown that when individuals

formed templates based on a cue, their performance improved when the targets matched these templates (Conci et al., 2019; Müller et al., 2003; Töllner et al., 2010; Vickery et al., 2005). Conversely, performance decreased when there was a mismatch between the template and the target (Müller et al., 2003; Töllner et al., 2010). The role of imagery in these processes has not been directly measured in N2pc studies, where it is often assumed rather than explicitly tested. One study explored this question using the N2pc component to measure the effects of color imagery on attentional selection (Cochrane, Townsend, Lee, et al., 2021). The findings showed that the N2pc was larger for congruent targets following an imagery cue but absent for incongruent targets, supporting the goal-driven account of attentional selection.

In summary, the integration of visual search paradigms with imagery instructions is a relatively new area of research (Cochrane et al., 2018, 2021; Monzel et al., 2021; Moriya, 2018) that warrants further investigation. While recent studies have demonstrated the impact of imagery on search performance, little is known about its early effects on attentional mechanisms. Following the goal-driven account, the top-down effects of imagery should be detectable during the selection process, offering insights into the temporal dynamics of attentional processing. Building on previous research showing behavioral effects of imagery, this study examines whether the impact of imagery can be detected during attentional selection via the N2pc ERP component.

Three EEG experiments were conducted. Experiment 1 tested whether the N2pc component could be traced during selection following imagery. In addition to replicating previous behavioral effects (Moriya, 2018, 2023), I hypothesized that imagery would

enhance selection when target objects matched the imagery but hinder selection when nontargets matched instead. Experiment 2 probed these effects further by comparing an imagery group with an active control group, who completed a separate mental task instead of imagery, hypothesizing that the imagery-related effects would be absent in the control group. Experiment 3 manipulated stimulus parameters by adding neutral trials, where neither the target nor nontargets matched the imagery. If imagery influences search for matching items, trials with non-matching items should not be affected by imagery.

3.3. Methods

3.3.1. Participants

Twenty undergraduates participated in Experiment 1. Three additional participants were excluded in the final analyses due to excessive eye-movement artifacts (> 30% of total trials removed due to artifacts and/or incorrect responses). Forty undergraduates participated in Experiment 2, with 20 participants in the imagery condition and 20 in the non-imagery active control condition. Four additional participants in the imagery condition and three additional participants in the active control condition were excluded in the final analyses due to excessive eye-movement artifacts. Group conditions were counterbalanced across participants. Twenty undergraduates participated in Experiment 3. Three additional participants were excluded in the final analyses due to excessive eye-movement artifacts. All participants were recruited via UC Riverside's Human Subject Research Pool. All participants self-reported normal or corrected-to-normal vision. General pre-screening process included normal or corrected vision, no color blindness, and fluent in English. Following protocols from my lab's previous EEG studies (Moon,

He, et al., 2022; Moon, Zhao, et al., 2022; Wu et al., 2013, 2015, 2017, 2018), I excluded participants if more than 30% of their EEG data were removed and/or accuracy was less than 70%. For each experiment, no analyses were conducted until data collection was complete. Participants received course credit for their participation.

3.3.2. Stimuli

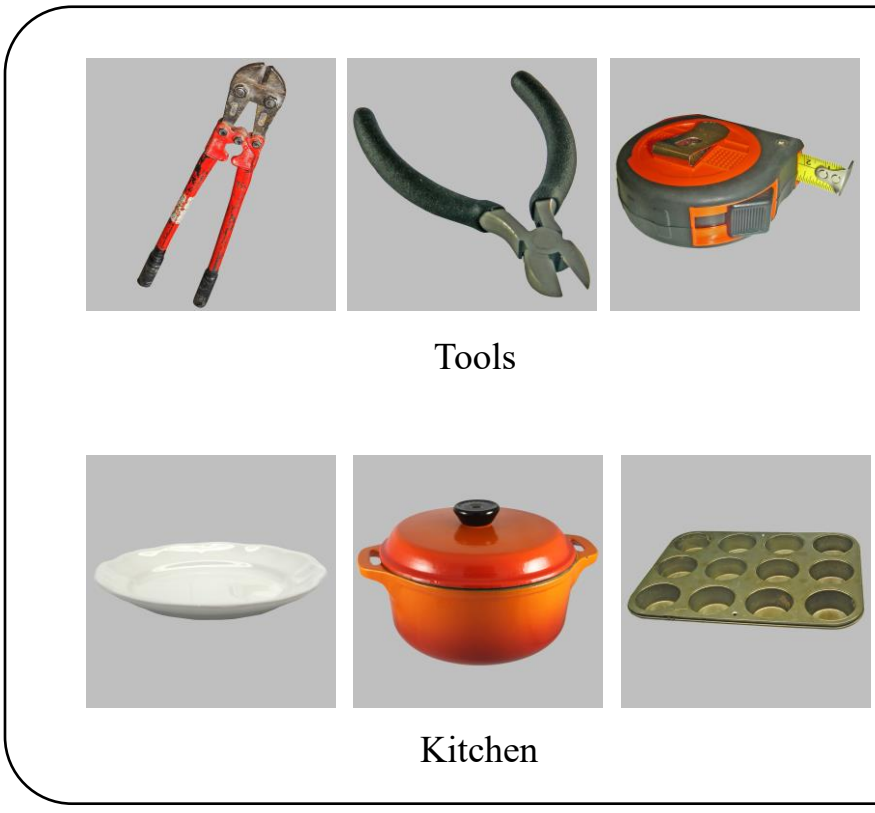
All experiments used the same stimulus set. Figure 3.1 shows sample stimuli. Stimuli were colored images of real-world objects from the Bank of Standardized Stimuli (Brodeur et al., 2014). Previous imagery-search studies primarily examined color imagery (Cochrane et al., 2018; Cochrane, Ng, Khosla, et al., 2021, 2021; Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021; Monzel et al., 2021; Moriya, 2018, 2023), making it unclear whether the observed effects may scale up to real-world objects. As a secondary goal, I examined whether effects would scale up to more complex objects.

Participants reported subjective ratings of familiarity and shape complexity for each image using a 5-point Likert scale (5 = highly familiar/complex, 1 = not familiar/complex). Images were assigned a category label at the superordinate level (electronics, toys, kitchen items, sport items, and tools) based on the proportion of category agreement across participants. These specific categories were chosen because they consisted of common objects likely recognizable for the participants. Images were excluded if less than 70% of participants agreed on a category label and if the average familiarity rating was less than 3. If more than one image contained items from the same basic-level category (e.g., multiple images of cups), the image with the highest familiarity

rating was used. Of the remaining images, we selected five from each category with the highest familiarity rating (= 25 visually distinct images) and at least 90% of category agreement. The average shape complexity rating was consistent across categories (Electronics: 2.19; Toys: 2.54; Kitchen items: 1.98; Sport items; 1.89; Tools: 2.65). Each item in the image was isolated on a grey background (RGB: 96, 96, 96). All stimuli were presented on a 24" 60Hz Dell monitor, using E-Prime 2.0 software.

Figure 3.1

Sample Stimuli



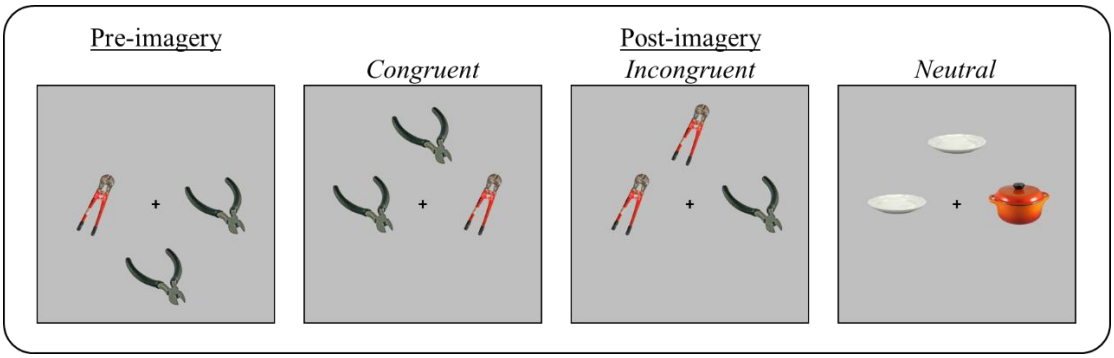
Note. Examples of stimulus from the 'tools' and 'kitchen' categories. The same stimuli were used in all experiments.

3.3.3. Design

Designs were identical across all experiments. Each trial consisted of two search arrays: pre-imagery and post-imagery. Both arrays presented three items arranged in a cross configuration surrounding a fixation cross in the center. All items shared the same category label. Categories were randomly selected across trials, with the restriction that the same category did not repeat back-to-back between trials. The items to be displayed were randomly selected. Two of the three items were positioned to the left and right of the fixation cross, each subtending 3.6° x 3.6°. The third item was located randomly either above or below the fixation cross. The pre-imagery array presented one item (the target) from a specific category and two identical items (nontargets) from the same category. The post-imagery array presented items in one of two ways (trial types): the congruent trial showed the same target and nontargets as in the pre-imagery array but located in different positions; the incongruent trial showed the previous target as two nontargets and the previous nontargets as the target. In Experiment 3, a third trial type with neutral trials was added: search items in the post-imagery array were randomly selected, with the exception that they do not repeat the same items from the pre-imagery array. Figure 3.2 shows a sample of each trial type.

Figure 3.2

Sample Trial Types



Note. Search goal was to locate the 'odd-one-out'. Targets in this example are bolt cutter (pre-imagery and congruent), wire cutter (incongruent), and cooking pot (neutral). Only Experiment 3 used neutral trial type. Other trial types were used in all experiments.

3.3.4. Procedures

Procedures were similar across all experiments. Figure 3.3 shows a sample trial. Before starting the experimental run, participants completed a practice run with 10 trials to get familiar with the task. For Experiments 1 and 2, the experimental run included 5 blocks of 32 trials (160 total trials), evenly split between congruent and incongruent trials. For Experiment 3, the experimental run included 6 blocks of 33 trials (198 total trials), evenly split across congruent, incongruent, and neutral trials.

At the start of each trial, participants saw a grey screen with black text, 'Press SPACE to start.' This screen served to separate the pre-imagery and post-imagery arrays between trials. Subsequently, a blank screen with a black fixation cross was presented for 500 ms, followed by the pre-imagery array displayed for 800 ms. Participants were instructed to locate the 'odd-one-out' object that did not resemble the other two items. Participants gave their responses using the arrow keys corresponding to the location of

the target (e.g., pressing the left arrow key when the target was on the left). The odd-one-out search instruction in pre-imagery was used to elicit pop-out effects without impact of imagery. Only Experiment 1 had a stimulus presentation of 2000 ms, as it was unclear how fast participants would respond to complex items without a pre-defined target. However, behavioral results showed an average response time of approximately < 800 ms with > 90% accuracy. Several N2pc studies, specifically those that used simple features, often restricted the presentation to a short duration (~200-600 ms) to reduce ocular artifacts (Luck, 2011; Luck & Hillyard, 1994). Given the high search performance and to maximize trials without EEG-related artifacts, the stimulus presentation was reduced to 800 ms for all other experiments.

The subsequent display was another blank screen with a fixation cross for 4000 ms. During this time, participants were instructed to vividly imagine a specific item. In Experiment 2, participants in the active control group were instructed to think of a number between one and four and later respond what number they thought of. There is no universally accepted duration to complete a basic imagery task. Previous studies instructed participants to imagine simple features for 2000 ms (e.g.,(Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021), 6000 ms (e.g., L. Kay et al., 2022; Keogh & Pearson, 2014), and even 10 seconds (e.g., Koenig-Robert & Pearson, 2019). To limit the already lengthy trial (pre-imagery, imagery, and post-imagery), 4000 ms was chosen and assumed to be sufficient for completing the imagery task.

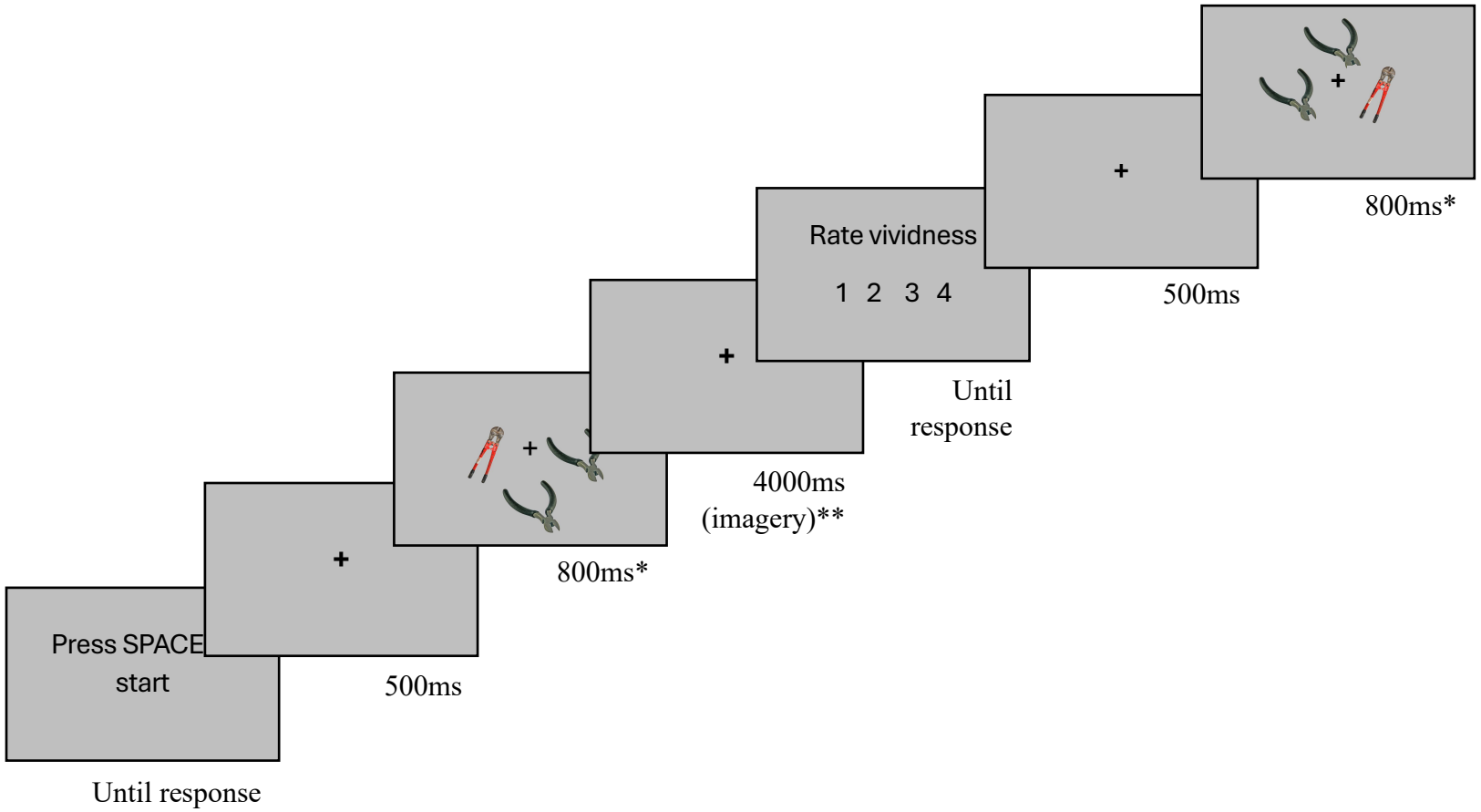
After, participants were prompted to rate the vividness of their imagery by selecting a number key from 1 (‘not vivid at all’) to 4 (‘very vivid’). The prompt remained on the screen until a response was provided. Following this, another blank screen with a fixation cross was shown for 500 ms. The final display showed the post-imagery array for 800 ms. Participants received the same instructions as the pre-imagery array (find the odd-one-out), using the arrow keys to give their responses. Unlike pre-imagery search, post-imagery search was susceptible to imagery-related effects. According to the stimulus-driven account, selection may be influenced by the salient features of the target, in which case effects between trial types should not be observed. According to the goal-driven account, selection may be influenced by features that matched with imagery, in which case effects between trial types should be observed. Post-imagery search allows for testing these two accounts.

3.3.5. EEG system and steps for preprocessing

All experiments used the same EEG system and steps for preprocessing (see Chapter 2). After artifact correction, incorrect trials and trials with other artifacts were removed from analyses using individualized thresholds for each participant. The following number of participants were excluded from final analyses due to excessive artifacts and/or low accuracy: three participants in Experiment 1, seven in Experiment 2, and three in Experiment 3.

Figure 3.3

Sample Trial



Note. Same task for all experiments. *For Experiment 1, the duration of pre- and post-imagery were 2000 msec. **For Experiment 2, active control group thought of a number from 1 to 4.

3.4. Results

3.4.1. Experiment 1

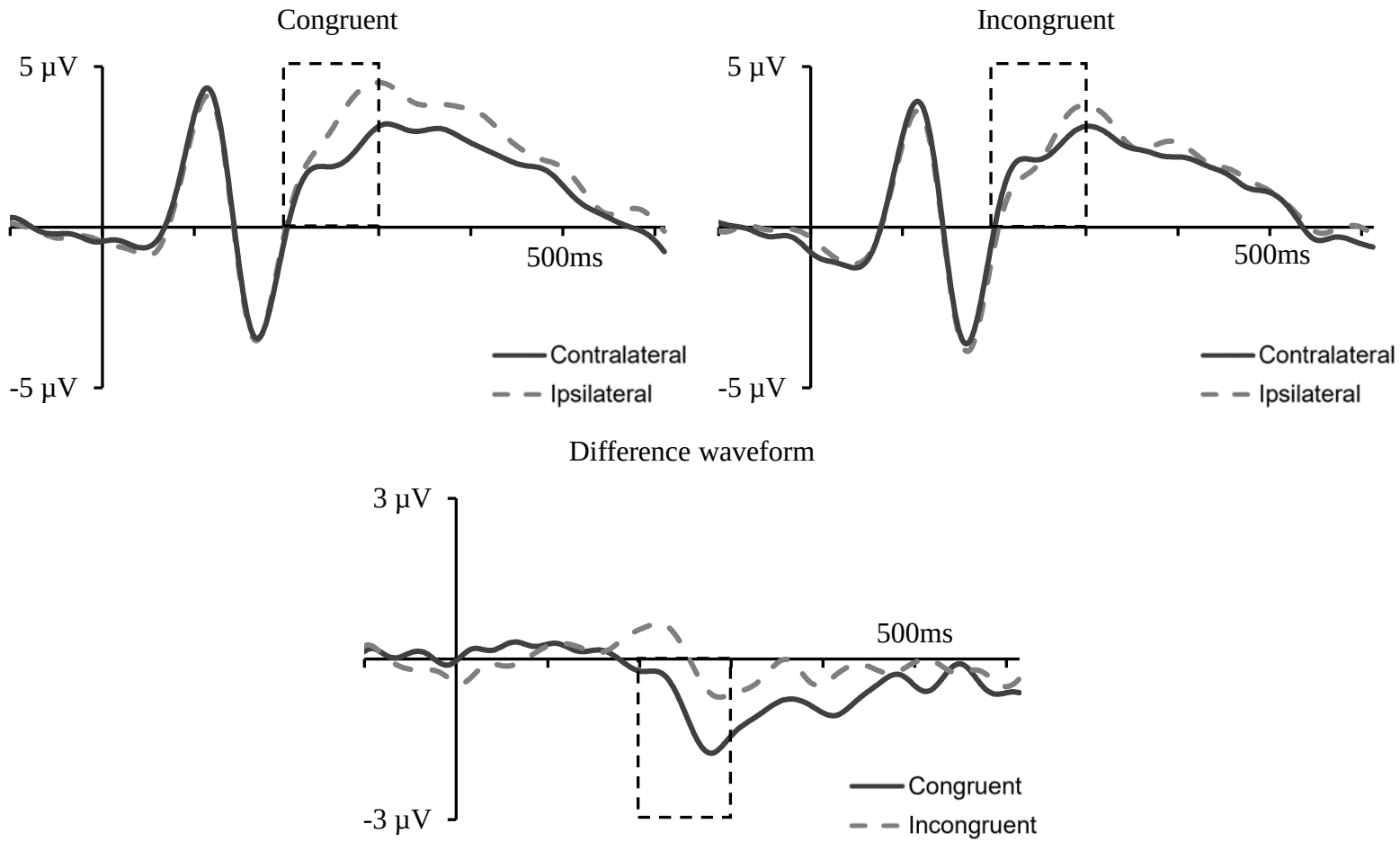
The goal was to examine whether imagery impacts selection and search performance. There were two predictions. First, using the N2pc mean amplitudes and onset latencies, I predicted that the components would be stronger and faster for congruent targets than incongruent targets, consistent with a previous study that used a similar paradigm with color imagery (Cochrane et al., 2021). Second, using the averaged reaction time, I predicted that reaction time would be faster for congruent targets than incongruent targets, aligning with the N2pc results and previous behavioral findings (Cochrane et al., 2018; Cochrane, Ng, Khosla, et al., 2021; Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021; Monzel et al., 2021).

3.4.1.1. N2pc mean amplitude

Figure 3.4 shows the grand-averaged parent and difference waveforms. Figure 3.5 shows the N2pc mean amplitudes. To compare the strength of selection between trial types (congruent and incongruent), a pairwise *t*-test using N2pc mean amplitudes was conducted between trial types (congruent and incongruent). The N2pc components were larger in the congruent (M = -1.07 μV, SD = 0.93) than in the incongruent trials (M = -0.07 μV, SD = 1.22), $t(19) = -2.49$, $p = .02$, $d = .40$. To examine the strength of selection for each trial type, a one-sample *t*-test compared to zero revealed significant components in the congruent trials, $t(19) = -3.66$, $p = .002$, but not in the incongruent trials, $t(19) = -0.12$, $p = .91$.

Figure 3.4

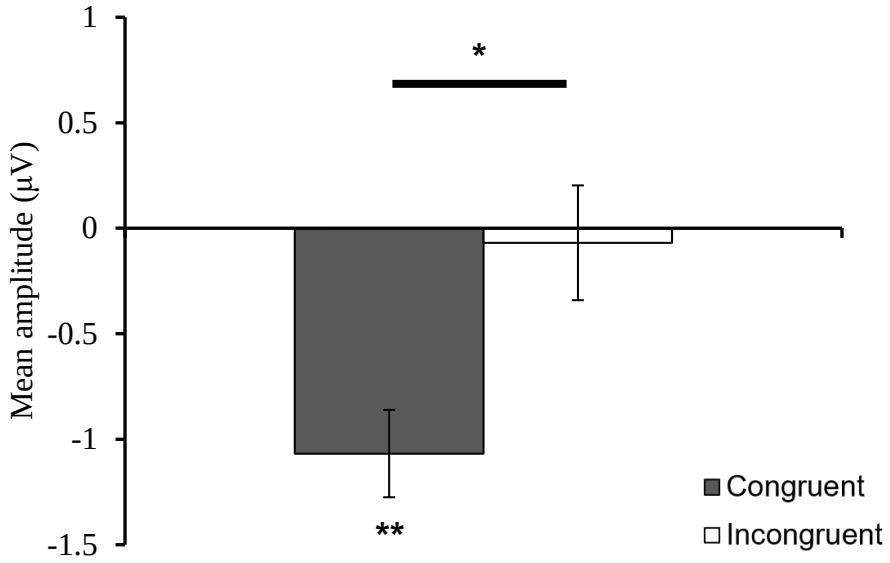
Grand averaged waveforms (Experiment 1)



Note. Parent and difference waveforms in post-imagery search. N2pc time window at 200 and 300 msec (dashed rectangle).

Figure 3.5

N2pc mean amplitudes (Experiment 1)



Note. Average of signals from 200 ms to 300 ms. Asterisks represent significance (*p < .05; **p < .01). Error bars represent ± 1 SE.

3.4.1.2. N2pc onset latencies

To compare the speed of selection based on whether imagery was instructed, a mixed-factor ANOVA using the N2pc onset latencies (jackknife approach) was conducted, with trial type as a within-subject variable and group condition as a between-subjects variable. There was a main effect of trial type, where the onset was faster in the congruent (M = 240.30 ms, SD = 7.20) than the incongruent (M = 259.50 ms, SD = 8.83) trials, $F(1,38) = 1638.40$, $p < .001$, $\eta_p^2 = .98$. There was an interaction between trial type and condition, $F(1,38) = 1013.38$, $p = <.001$, $\eta_p^2 = .96$. To investigate this interaction, Bonferroni-corrected comparisons (corrected alpha = .05/2 = .025) were conducted between trial types for the group conditions separately. In the imagery condition, the

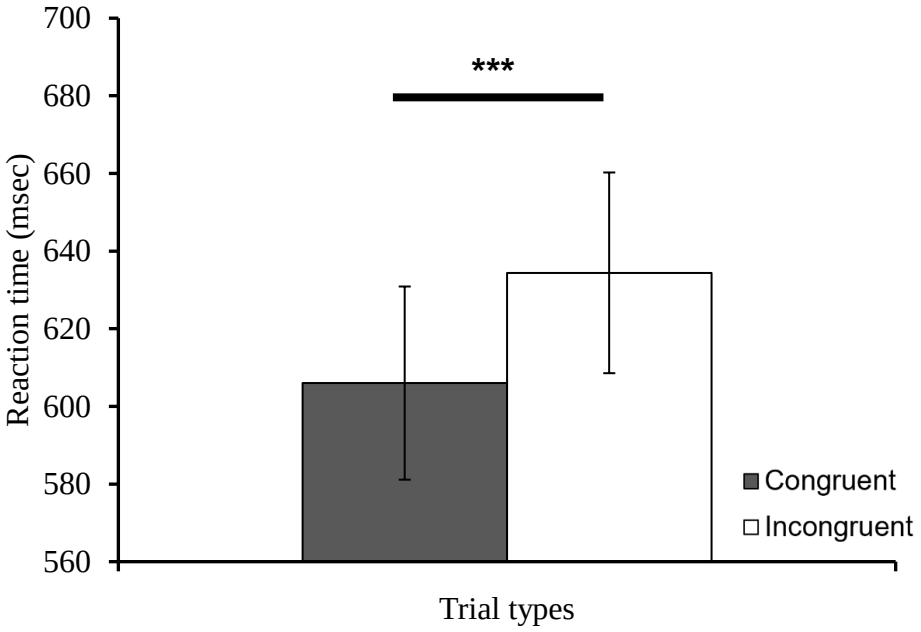
onset was faster in the congruent (M = 233.80 ms, SD = 1.58) than in the incongruent (M = 268.10 ms, SD = 1.02) trials, $t(19) = -51.13$, $p < .001$, $d = -14.77$. In the active control condition, the onset was faster in the congruent (M = 246.80 ms, SD = 3.86) than in the incongruent (M = 250.90 ms, SD = 1.77) trials, $t(19) = -6.11$, $p < .001$, $d = -1.77$.

3.4.1.3. Reaction time

Figure 3.6 shows the average reaction time. To compare the reaction times between trial types, a pairwise t-test using mean reaction time (ms) was conducted between trial types. Reaction time was faster in the congruent (M = 606.00 ms, SD = 111.26) than in the incongruent trials (M = 634.40 ms, SD = 115.59), $t(19) = -4.03$, $p < .001$, $d = .90$.

Figure 3.6

Average reaction time (Experiment 1)



Note. Asterisks represent significance (***$p < .001$). Error bars represent ± 1 SE.

3.4.1.4. Summary

Findings revealed both behavioral and neural effects following imagery. Participants were faster at responding to imagery-matched targets than to imagery-mismatched targets. Attentional selection to imagery-matched targets was larger and faster than to imagery-mismatched targets, showing that imagery can prioritize attention to items beyond just basic features. However, although imagery seemed to impact attention, it cannot be ruled out that an alternative process influenced search as there was no control condition. Some studies used an active control condition in which participants had to complete other mental operations (Cochrane et al., 2018; Maljkovic & Nakayama, 1994). This concern was addressed in the next experiment by adding an active control group.

3.4.2. Experiment 2

There were two goals. First, I aimed to replicate the effects from Experiment 1. Second, I compared these effects by including an active control group. Participants were randomly assigned to either an imagery group (identical procedure to Experiment 1) or a non-imagery active control condition (mental task in lieu of imagery). I predicted that the effects observed in the imagery group would be absent in the active control group.

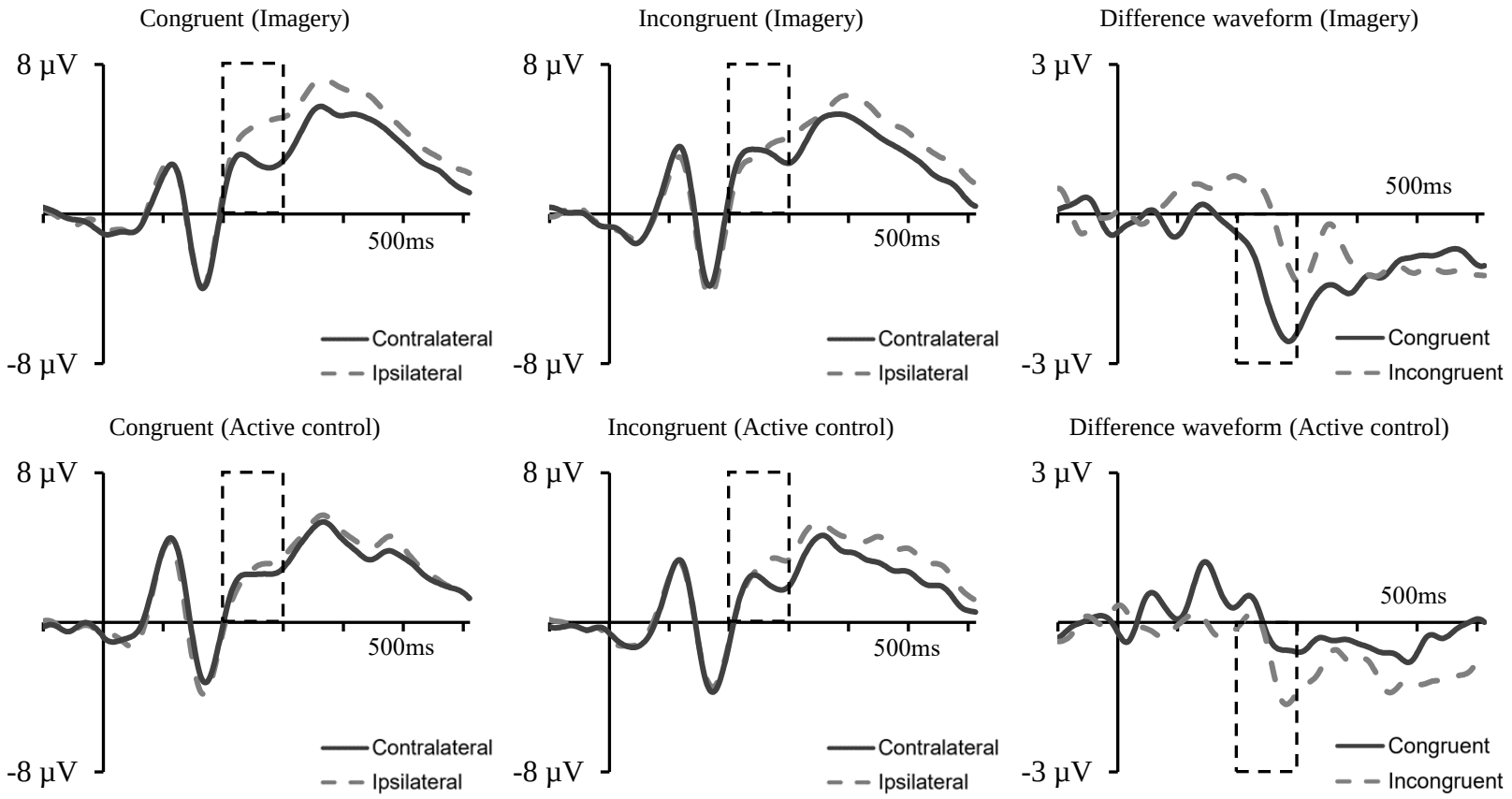
3.4.2.1. N2pc mean amplitude

Figure 3.7 shows the grand-averaged parent and difference waveforms. Figure 3.8 shows the N2pc mean amplitudes. To compare the strength of selection between trial types (congruent and incongruent) for each group condition (imagery and active control), a mixed-factor ANOVA using the N2pc mean amplitudes was conducted with trial type as

a within-subject variable and group condition (imagery and active control) as a between-subjects variable. There was a main effect of trial type, where the N2pc component was larger in the congruent trials (M = -0.95 μV, SD = 1.25) than incongruent trials (M = -0.44 μV, SD = 1.25), $F(1,38) = 11.41$, $p = .002$, $\eta_p^2 = .23$. There was an interaction between trial type and condition, $F(1,38) = 46.78$, $p = <.001$, $\eta_p^2 = .55$. To investigate this interaction, Bonferroni-corrected comparisons (corrected alpha = .05/2 = .025) were conducted between trial types for the group conditions separately. In the imagery condition, the N2pc component was larger in the congruent (M = -1.69 μV, SD = 0.72) than in the incongruent trials (M = -0.14 μV, SD = 0.87), $t(19) = -4.12$, $p < .001$, d = -1.30. In the active control condition, the N2pc component was not significantly different between congruent (M = -0.22 μV, SD = 1.25) and incongruent trials (M = -0.74, SD = 1.50), $t(19) = 2.45$, $p = .12$, $d = .47$. To examine the strength of selection for each trial type, a one-sample t-test compared to zero revealed significant components in the congruent trials, $t(19) = -10.51$, $p < .001$, but not in the incongruent trials, $t(19) = -0.71$, $p = .49$, for the imagery condition. In contrast, for the active control condition, significant components were found in the incongruent trials, $t(19) = -2.22$, $p = .04$, but not in the congruent trials, $t(19) = -0.78$, $p = .45$.

Figure 3.7

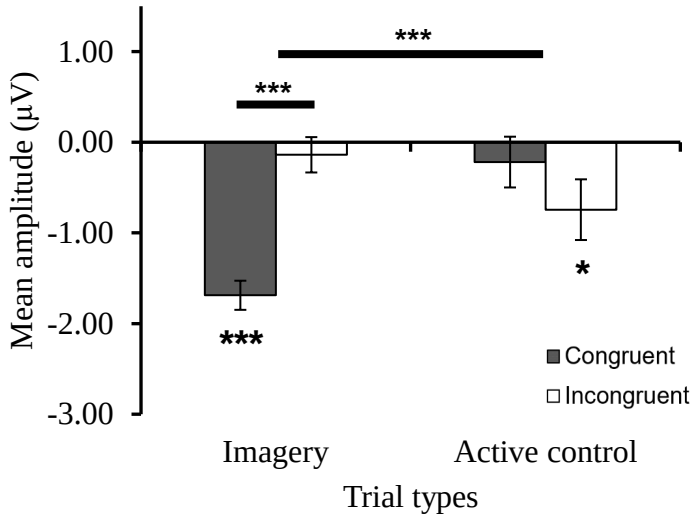
Grand averaged waveforms (Experiment 2)



Note. Parent and difference waveforms in post-imagery search. N2pc time window at 200 and 300 msec (dashed rectangle).

Figure 3.8

N2pc mean amplitudes (Experiment 2)



Note. Average of signals from 200 ms to 300 ms. Asterisks represent significance (*p < .05; ***p < .001). Error bars represent ± 1 SE.

3.4.2.2. N2pc onset latencies

To compare the speed of selection based on whether imagery was instructed, a mixed-factor ANOVA using the N2pc onset latencies (jackknife approach from Experiment 1) was conducted with trial type as a within-subject variable and group condition as a between-subjects variable. There was a main effect of trial type, where the onset was faster in the congruent (M = 240.30 ms, SD = 7.20) than the incongruent (M = 259.50 ms, SD = 8.83) trials, $F(1,38) = 1638.40$, $p < .001$, $\eta_p^2 = .98$. There was an interaction between trial type and condition, $F(1,38) = 1013.38$, $p = < .001$, $\eta_p^2 = .96$. To investigate this interaction, Bonferroni-corrected comparisons (corrected alpha = .05/2 = .025) were conducted between trial types for the group conditions separately. In the imagery condition, the onset was faster in the congruent (M = 233.80 ms, SD = 1.58) than

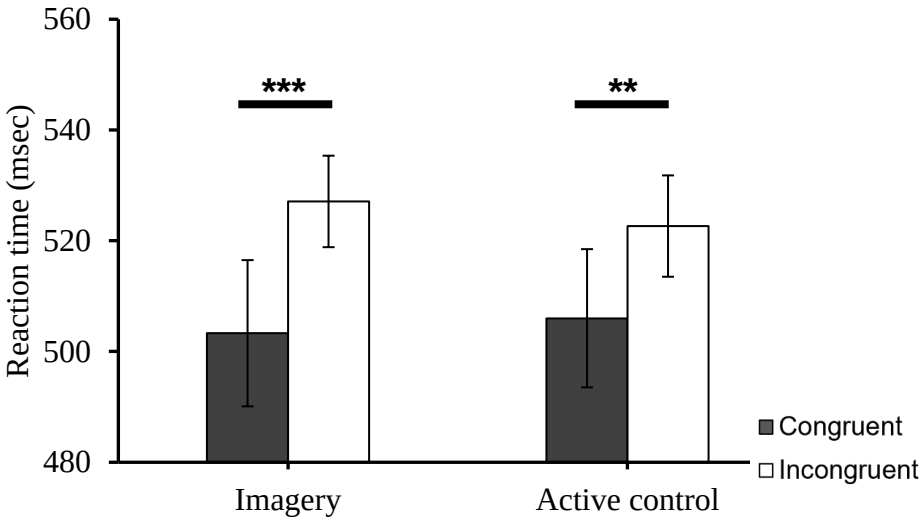
in the incongruent (M = 268.10 ms, SD = 1.02) trials, $t(19) = -51.13$, $p < .001$, $d = -14.77$. In the active control condition, the onset was faster in the congruent (M = 246.80 ms, SD = 3.86) than in the incongruent (M = 250.90 ms, SD = 1.77) trials, $t(19) = -6.11$, $p < .001$, $d = -1.77$.

3.4.2.3. Reaction time

Figure 3.9 shows the average reaction time. To compare the reaction times between trial types for each group condition, a mixed-factor ANOVA was conducted using mean reaction times with trial type as a within-subject variable and group condition (imagery and active control) as a between-subjects variable. There was a main effect of trial type, where reaction time was faster in the congruent (M = 504.65 ms, SD = 48.63) than in the incongruent trials (M = 524.88 ms, SD = 48.36), $F(1, 38) = 26.91$, $p < .001$, $\eta_p^2 = .42$. No other main effect nor interaction was found, $F < 0.85$. To test whether the main effect of trial type was driven by one condition or both, Bonferroni-corrected comparisons (corrected alpha = .05/2 = .025) were conducted between trial types for the imagery condition and for the active control condition, separately. For the imagery condition, reaction time was faster in the congruent (M = 503.30 ms, SD = 59.05) than in the incongruent trials (M = 527.10 ms, SD = 55.82), $t(19) = -4.08$, $p < .001$. Similarly for the active control condition, reaction time was faster in the congruent (M = 506.00 ms, SD = 36.94) than in the incongruent trials (M = 522.65 ms, SD = 9.15), $t(19) = -3.22$, $p = .005$.

Figure 3.9

Average reaction time (Experiment 2)



Note. Asterisks represent significance (**$p < .01$; ***$p < .001$). Error bars represent ± 1 SE.

3.4.2.4. Summary

Behavioral and neural effects from Experiment 1 were replicated for the imagery condition. These results supported the hypothesis that imagery impacts behavior and that this impact aligned with effects on selection. For the active control group, there were no differences in the magnitude of the N2pc components, but there was a small difference in onset. This difference was quite small (~4 ms) to raise concerns over alternative effects. More notably, behavioral effects were present in the control group. One potential explanation is that there were other advantages besides imagery that were present in both group conditions. In both conditions, the repetition of targets may have induced intertrial priming effects (Maljkovic & Nakayama, 1994). Additionally, a pop-out effect still could have occurred in the incongruent trial as the presentation structure (one item always

popped out) did not differ from congruent trials. One interpretation is that the pop-out in incongruent trials was 'overwritten' by imagery, such that only top-down guidance by imagery occurred (Cochrane et al., 2018). Another interpretation is that the pop-out remained present and competed with imagery for attentional guidance. Together, the inability to isolate effects from imagery and pop-out questions whether it was imagery that drove the observed effects. This concern was addressed in the next experiment by including a neutral trial type.

3.4.3. Experiment 3

There were two goals. First, I aimed to replicate the effects from Experiments 1 and 2. Second, I aimed to separate imagery-related effects from pop-out effects by adding a neutral trial type. Given that imagery was involved but pop-out may have been involved in the incongruent trial types, a third trial type was necessary, in which imagery could not impact search but the questionable pop-out effect could. In previous studies, this trial type served as a 'middle ground' between congruent and incongruent trials, such that response time was faster in congruent than in neutral, and slower in incongruent than in neutral (Cochrane et al., 2018; Monzel et al., 2021; Moriya, 2018, 2023). Such comparisons allowed for isolating imagery-related effects, as imagery was instructed in all trial types. Only the congruent trials would benefit from imagery, incongruent trials would be hindered by imagery, and neutral trials would neither benefit from nor be hindered by imagery. All trial types would benefit from pop-out effect, and the neutral trial type served to isolate this effect. I predicted that the N2pc would be strongest and fastest for

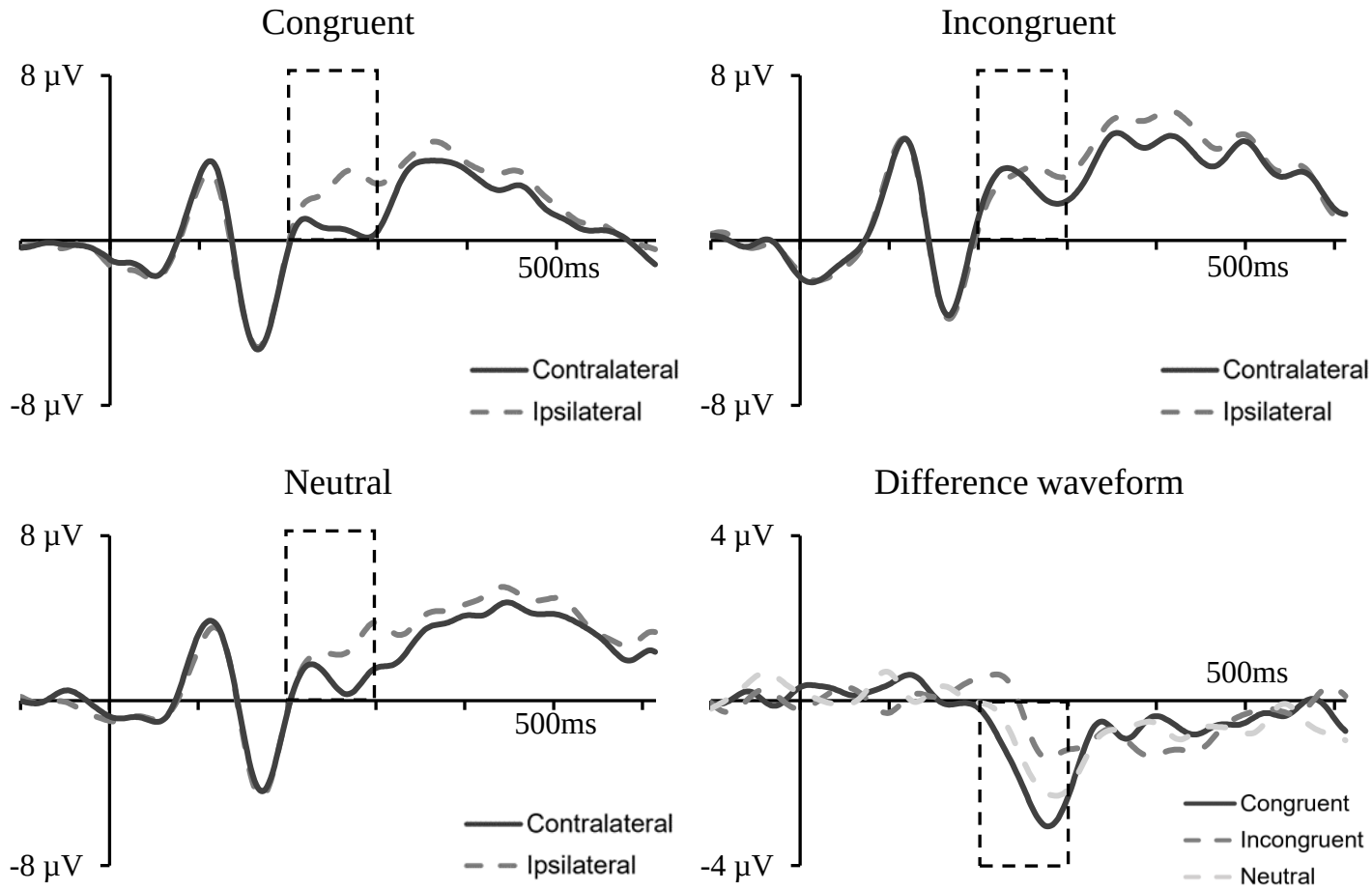
congruent, relative to the neutral and incongruent trials. I also predicted that the reaction time would be fastest for congruent trials and slowest for incongruent trials.

3.4.3.1. N2pc mean amplitude

Figure 3.10 shows the grand-averaged parent and difference waveforms. Figure 3.11 shows the N2pc mean amplitudes. To compare the strength of selection across trial types (congruent, incongruent, and neutral), a one-way ANOVA using the N2pc mean amplitudes was conducted with trial type as the factor. Analysis revealed a significant difference in trial types, $F(2,38) = 4.07$, $p = .03$, $\eta_p^2 = .18$. Bonferroni-corrected comparisons (corrected alpha = .05/3 = .017) revealed larger N2pc components in the congruent (M = -2.35 μV, SD = 1.90) than incongruent trials (M = -0.89 μV, SD = 1.65), $t(19) = -2.78$, $p = .008$. However, there were no significant differences between congruent and neutral (M = -1.91 μV, SD = 1.66), $t(19) = -0.85$, $p = .33$, nor between neutral and incongruent, $t(19) = 1.94$, $p = .06$. To examine the strength of selection for each trial type, a one-sample t-test compared to zero revealed significant components in the congruent, $t(19) = -5.52$, $p < .001$, incongruent, $t(19) = -2.42$, $p = .026$, and neutral trials, $t(19) = -5.15$, $p < .001$.

Figure 3.10

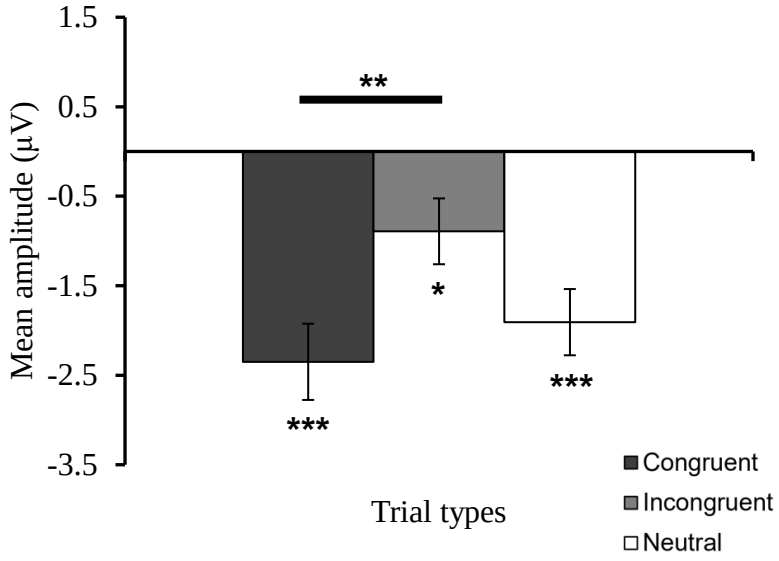
Grand averaged waveforms (Experiment 3)



Note. Parent and difference waveforms in post-imagery search. N2pc time window at 200 and 300 msec (dashed rectangle).

Figure 3.11

N2pc mean amplitudes (Experiment 3)



Note. Average of signals from 200 ms to 300 ms. Asterisks represent significance (*p < .05; **p < .01; ***p < .001). Error bars represent ± 1 SE.

3.4.3.2. N2pc onset latencies

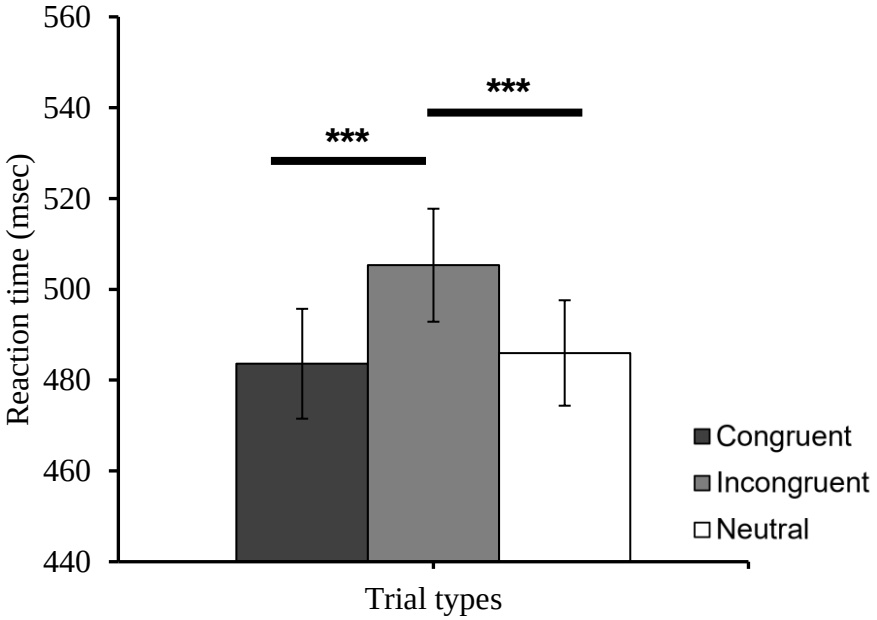
To compare the speed of selection across trial types, a one-way ANOVA using the N2pc onset latencies (jackknife approach from Experiments 1 and 2) was conducted with trial type as the factor. Analysis revealed a significant difference in trial types, $F(2,38) = 15.49$, $p < .001$, $\eta_p^2 = .45$. Bonferroni-corrected comparisons (corrected alpha = .05/3 = .017) revealed faster onset in the congruent (M = 232.30 ms, SD = 2.08), relative to incongruent (M = 264.00 ms, SD = 33.34), $t(19) = -4.30$, p <.001, d = -1.64, and neutral (M = 238.10 ms, SD = 1.89) trials, $t(19) = -14.22$, p <.001, d = -0.30. Onset was slower in the incongruent than neutral trials, $t(19) = 3.47$, p <.003, d = 1.34.

3.4.3.3. Reaction time

Figure 3.12 shows the average reaction time. To compare the reaction times across trial types, a one-way ANOVA using reaction times was conducted with trial type as the factor. Analysis revealed a significant difference in trial types, $F(2,38) = 10.20$, $p < .001$, $\eta_p^2 = .35$. Bonferroni-corrected comparisons (corrected alpha = .05/3 = .017) revealed faster reaction time in the congruent ($M = 483.60$ ms, $SD = 54.11$) than incongruent trials ($M = 505.30$ ms, $SD = 55.65$), $t(19) = -4.12$, $p < .001$. Reaction time also was faster in the neutral ($M = 485.98$ ms, $SD = 51.92$) than incongruent trials, $t(19) = 3.67$, $p < .001$. There was no significant difference between congruent and neutral trials, $t(19) = -2.38$, $p = .33$.

Figure 3.12

Average reaction time (Experiment 3)



Note. Asterisks represent significance (***$p < .001$). Error bars represent ± 1 SE.

3.4.3.4. Summary

Behavioral and neural effects were present when comparing measures between congruent and incongruent trials, as expected. The N2pc onset but not the mean amplitude was significantly different between congruent and neutral, and neutral and incongruent. One possibility is that the differences in mean amplitude were too small to detect subtle differences. Another possibility is that in congruent and neutral trials, selection benefited from pop-out effect. This view aligns with the behavioral findings, with no difference between the two trial types. However, the lack of difference does not rule out the possibility that imagery had additive benefit in congruent trials, as suggested by the N2pc onset. Another observation was that both the N2pc onset and reaction time were faster in neutral trials than incongruent trials, consistent with the prediction that imagery hindered search only for incongruent trials (Moriya 2018, 2023; though see Monzel et al., 2021).

3.5. General discussion

3.5.1. Summary and interpretation of findings

I examined the question of whether imagery elicited top-down effects on selection and ultimately, behavior. To address this question, I used the N2pc component in addition to reaction time as relevant measures. Effects were assessed based on comparisons between congruent and incongruent trial types. Experiment 1 showed that following imagery instructions, the N2pc was larger and faster for congruent than incongruent trials, as predicted (Cochrane, Townsend, Lee, et al., 2021). Experiment 2 showed that the N2pc was not different in magnitude when participants did not receive imagery

instructions. Experiment 3 showed that the N2pc was faster for congruent than neutral trials and for neutral than incongruent trials. Finally, across all experiments, reaction time was faster for congruent than incongruent trials, replicating standard findings from previous studies (Cochrane, Ng, et al., 2021; Cochrane, Townsend, et al., 2021; Cochrane, Wang, et al., 2021; Moriya, 2018, 2023).

Some unpredicted results require further discussion. In Experiment 2, the active control group was not completely absent of effects. Although there were no differences in the N2pc mean amplitude, the N2pc onset was slightly faster for the congruent than incongruent trials. This difference was particularly small (~4 ms), which may not reflect meaningful differences. Instead, a more surprising result was that behavioral effects were present. One possibility was that participants formed expectations of the upcoming target. This was unlikely, as trial types were evenly distributed and randomized, thus there was no way to predict the upcoming target. If expectation alone was sufficient, then passive conditions in which no mental tasks were assigned should reveal similar results. Previous studies did not find such findings (Moriya, 2018, 2023), and other lines of studies also reported similar results (Keogh & Pearson, 2017; Kosslyn & Thompson, 2003; Pearson et al., 2008). Other studies included more congruent trials than incongruent trials to increase expectation, yet showed similar results (e.g., Cochrane et al., 2018). Another possibility was that participants engaged in imagery despite not being instructed to do so. However, if they did, then their effects should have aligned with the neural effects observed in the group that were given imagery instructions.

Finally, effects may have been partially driven by the repetition of targets, or intertrial priming (Maljkovic & Nakayama, 1994, 1996). The current view is that an implicit memory system maintains information of recently focused items to facilitate attention if the same items appeared again. Pivotal studies showed a specific form of intertrial priming called priming of pop-out (PoP; Maljkovic & Nakayama, 1994, 1996), which is the repetition effect of pop-out targets, similar to how congruent trials were presented. Related findings showed that the N2pc was faster to repeated targets compared to switched targets, presented similarly to congruent and incongruent trials, respectively (Eimer et al., 2010). However, what exactly PoP does to attentional processing remain unclear (Ramgir & Lamy, 2022). If PoP systematically drove the behavioral effects, then there is no reason why the imagery group also would not benefit from PoP as the task design was not different. A simple explanation is that the active control group benefited from PoP, whereas the imagery group benefited from both PoP and imagery-related effects. Because only the imagery group showed N2pc effects, it was unlikely that PoP influenced target selection but perhaps another stage in attentional processing (Eimer, 2014).

Another unpredicted result was from Experiment 3. The N2pc mean amplitude was not different between the congruent and neutral trials, nor between the neutral and incongruent trials. However, the N2pc onset did differ, with fastest onset for congruent and slowest for incongruent. One explanation is that imagery had more influence on the temporal dynamics of selection than influence on the enhancement of selection, consistent with findings in Experiments 1 and 2. Another explanation is that the N2pc

mean amplitude was too small to detect subtle differences when comparing to neutral trials (~0.5 to 1.5 μV). Regardless, pop-out effect may have enhanced selection for congruent and neutral trials, though imagery benefited congruent trials with faster selection. The opposite was true with incongruent trials. If pop-out effect enhanced selection for incongruent, then the N2pc onset and reaction time should have been comparable to the neutral trials. In contrast, if imagery hindered selection for incongruent, then the N2pc onset and reaction time should be slower (Cochrane et al., 2018), as observed. Putting it together, the top-down effects of imagery were best captured by the differences in N2pc onset, suggesting that pop-out effects were modulated with favorable effects when aligned with imagery and hindering effects when misaligned.

3.5.2. Theoretical contributions

This study provides several key insights regarding imagery as a top-down process. First, the findings support the view that imagery is one type of top-down process that can influence selection. While the search literature has extensively discussed factors that guide attention (Wolfe & Horowitz, 2017), this discussion has been less explored in imagery studies (Cochrane et al., 2018). Importantly, the results contribute to the ongoing debate on attentional mechanisms, reinforcing the perspective that top-down processes, such as imagery, can influence attentional selection. This view aligns with theories suggesting that top-down processes are critical in directing attention toward relevant stimuli (Wolfe, 2021).

Second, what place does imagery have in attentional models? One suggestion is that imagery shares underlying mechanisms with those used during the preparation stage of search (Eimer, 2014). During this stage, a search template is formed by activating visual representations of the target (Desimone & Duncan, 1995; Duncan & Humphreys, 1989), a process similar to the depictive theory of imagery, where mental images serve as functional equivalents to visual perception (Pearson et al., 2015). The findings suggest that imagery elicits top-down effects similar to those observed when using templates, supporting the goal-driven account of attention (Töllner et al., 2010; Vickery et al., 2005; Wolfe et al., 2004). However, while the effects appear similar, the extent to which imagery and search templates overlap remains unclear. This study highlights the potential for imagery to guide attention toward matching features, as evidenced by the observed N2pc effects, yet further research is needed to clarify the exact nature of this relationship (Cabbai et al., 2023; Keogh & Pearson, 2021).

Third, the findings raise an important consideration for individual differences in imagery skills. Specifically, the study opens new avenues for comparing attentional mechanisms between individuals with intact imagery and those with aphantasia, a condition characterized by an inability to visualize mental images (Zeman et al., 2015). Behavioral findings showed faster reaction time from individuals with intact imagery than those with aphantasia, likely due to the advantages conferred by imagery (Monzel et al., 2021). However, there is a significant gap in the neural evidence comparing the selection processes between these groups. Investigating how imagery, or the lack thereof,

influences early attentional processes before behavioral responses could offer valuable insights into the cognitive and neural differences underlying this condition.

3.5.3. Limitations

There were some limitations to the study. First, all experiments (excluding active control group) had identical imagery instructions, but it was unclear whether imagery actually was used at each individual trial. Ultimately, the engagement of imagery was assumed by the behavioral and neural effects. This concern partially was addressed in Experiment 2, which revealed neural effects (N2pc mean amplitude) after imagery instruction but not after other mental instruction. However, this manipulation only concerned the *disengagement* of imagery for the control group. Related studies used passive conditions (Moriya, 2018, 2023) or verbal articulation instructions (Cochrane et al., 2018) to achieve the same goal, though here, procedures were constrained by the ERP design (Huffmeijer et al., 2014). Future studies should use ERP designs to test for appropriate paradigms with a control condition.

Second, as previously mentioned, it cannot be ruled out that intertrial priming influenced the results (Maljkovic & Nakayama, 1994, 1996), as congruent and incongruent targets were inevitably repeated and switched targets, respectively. A previous study paired two search arrays within one trial to elicit intertrial priming, an automatic effect in which search was faster when the targets (e.g., red square) and nontargets (green square) repeated in successive trials compared to when targets (green square) and nontargets (red square) switched (Cochrane et al., 2018). This effect was reversed (faster reaction time for switched than repeated) by instructing participants to

imagine the color opposite of the target (if target was a red square, imagine the color green). This manipulation was not possible with the current design, as each trial presented unique items and no 'opposite' variants were present. To address this limitation, Study 2 implemented similar imagery instruction in which participants imagined the stimulus not previously selected.

Finally, pop-out effects were assumed by the experimental design, as the target was distinct from homogenous nontargets. However, I did not test that there actually were pop-out effects. Previous studies systematically manipulated the difference in targets and nontargets by one feature (e.g., Dosher et al., 2004; Saiki et al., 2005; Shen & Reingold, 2001; Treisman & Gormican, 1988). Because complex items were used, the targets and nontargets unsystematically varied across several features (e.g., shape and color). This limitation was examined more closely in Study 2, which systematically manipulated the difference between target and nontargets by one salient feature using asymmetric search designs (Treisman and Souther,1985; Wolfe, 2001). Furthermore, the slope reaction time was measured to directly examine pop-out effects (Wolfe et al., 2010).

3.5.4. Conclusion

The present study advanced current understanding of how top-down imagery influences attentional processing. The enhancement of the N2pc component in congruent trials provided direct neural evidence supporting the goal-driven account of attention and underscores the role of imagery in early attentional selection.

Chapter 4: Interaction between top-down imagery and pop-out effects

4.1. Abstract

Influential work on asymmetric search showed behavioral advantages when targets were presented with salient features. Search was more efficient when the target had more features than the nontargets (Q among Os) than when the target had less features than the nontargets (O among Qs). While stimulus-driven accounts suggest that salient features elicit pop-out effects through bottom-up processing, remaining unaffected by top-down influences, goal-driven accounts propose that bottom-up processing can be modulated by top-down processes. Study 2 examined how top-down imagery interacts with pop-out effects using similar asymmetric search design. In Experiments 4 and 5, participants completed visual search tasks in which they searched for either 'Q among Os' or 'O among Qs'. Unlike Study 1, participants were instructed to imagine the letter not previously selected to address confounding effects of intertrial priming. Experiment 4 revealed faster reaction time for congruent targets than incongruent targets when the feature was present but the opposite when the feature was absent. Experiment 5 replicated these effects and additionally reported stronger and faster N2pc components for congruent targets than incongruent targets. Both experiments showed that search was more efficient (flatter slope reaction time) when the feature was present than when absent. These results suggest that top-down imagery can modulate bottom-up processing in line with the goal-driven account. The findings contribute to understanding the interaction between imagery and pop-out effects and their implications for attentional processes.

4.2. Introduction

Visual search asymmetry refers to the phenomenon where search efficiency varies depending on the presence or absence of target-defining features (Treisman, 1985). Seminal research demonstrated that searching for a target defined by a salient feature among homogeneous nontargets lacking this feature (e.g., Q among Os) was more efficient than when the roles were reversed (O among Qs; Treisman & Souther, 1985). To investigate pop-out effects, researchers often manipulated the set size of nontargets and measure slope reaction times as a function of set size (Wolfe et al., 2010). In a 'Q among Os' task, a flat slope indicated that search was insensitive to set size, suggesting a pop-out effect (Dosher et al., 2004; Saiki et al., 2005; Shen & Reingold, 2001; A. Treisman & Gormican, 1988). Conversely, in an 'O among Qs' task, a positive slope suggested a more serial search, indicative of a more effortful and less efficient process (Dosher et al., 2004; Saiki et al., 2005; Shen & Reingold, 2001; Treisman & Gormican, 1988).

Given that mental imagery is understood to exert top-down effects on attentional processing, a key question arises: How does imagery interact with the effects observed in asymmetric search studies? This question is underexplored but touches on a broader theoretical issue concerning the interaction between bottom-up and top-down processing in attention-driven tasks. The stimulus-driven account posits that pop-out effects in asymmetric search are driven by bottom-up processing that is unaffected by top-down influences (Theeuwes, 2010). In contrast, the goal-driven account suggests that pop-out effects can be modulated by top-down processing (Eimer & Kiss, 2010). The debate is

not about whether these processes contribute to overall attentional processing (Eimer & Kiss, 2010; Theeuwes, 2010) but rather about how they interact.

Supported by Study 1 and recent literature (Cochrane et al., 2018; Monzel et al., 2021; Moriya, 2018, 2023), imagery likely evokes top-down effects on attentional processing. However, these effects have not been observed in asymmetric search paradigms. Previous studies have often used an 'odd-one-out' paradigm, where the target was distinguished by a salient feature (e.g., a green target among red nontargets), similar to a 'Q among Os' search. Imagery effects observed in these studies aligned with the goal-driven account (Cochrane et al., 2018). If pop-out effects were truly immune to top-down influences, imagery should not alter search performance, leading to no differences between congruent and incongruent trials. However, evidence suggests that effects were sensitive to trial types rather than merely the saliency of features (Cochrane, Townsend, Lee, et al., 2021).

To further investigate the interaction between top-down imagery and pop-out effects, asymmetric search designs can be used to distinguish top-down effects based on the presence (Q among Os) or absence (O among Qs) of salient features. In other words, how do top-down effects compare when pop-out effects are evoked and when absent? The phenomenon of visual search asymmetries provides a valuable framework for exploring this question. The stimulus-driven account predicts that top-down imagery would not modulate pop-out effects, whereas the goal-driven account suggests that it would. Specifically, if traditional asymmetric search findings are observed regardless of trial type, this would support the stimulus-driven account. Conversely, if effects of

imagery are observed regardless of the salient features, this would support the goal-driven account.

Two experiments were conducted to test these hypotheses: one behavioral experiment (Experiment 4) and one EEG experiment (Experiment 5). Both experiments followed identical procedures, except that Experiment 5 included EEG recording. I hypothesized that the top-down effects of imagery would modulate pop-out effects in line with the goal-driven account. The search paradigm from Experiment 1 was adapted to an asymmetric search design, which allowed for manipulating the presence or absence of salient features (Wolfe, 2001) and trial types (congruent and incongruent). Additionally, the set size of search items was manipulated to measure slope reaction time, which is crucial for assessing whether search efficiency differs based on the presence of a salient feature. The aim of Experiment 4 was to investigate the interaction between top-down imagery and pop-out effects through behavioral measures, including search performance and search efficiency. The aim of Experiment 5 was to replicate the behavioral effects and additionally test the neural effects on selection, using the N2pc component.

4.3. Methods

Experiments 4 and 5 implemented identical protocols, with the exception that only Experiment 5 included EEG recording.

4.3.1. Participants

Thirty-two undergraduates participated in Experiment 4. One additional participant was excluded in the final analyses due to low accuracy (<70%). Twenty undergraduates participated in Experiment 5. Four additional participants were excluded

in the final analyses due to excessive eye-movement artifacts. All participants were recruited via UC Riverside's Human Subject Research Pool and were compensated with Psychology course credit. All participants self-reported normal or corrected-to-normal vision. Participants were excluded if more than 30% of their EEG data were removed and/or accuracy was less than 70%. No analyses were conducted until data collection was complete.

4.3.2. Stimuli

Stimuli were Os (circles with a diameter of 0.7 of visual angle) and Qs (Os with a 0.35 vertical bar on the lower point, with half the bar inside the circle and other half outside the circle). These letters were used to distinguish the target and nontargets by a single salient feature, which is difficult to do with complex items in Study 1. Similar stimulus parameters were used in pivotal studies that demonstrated search asymmetry (A. Treisman & Gormican, 1988; A. Treisman & Souther, 1985). Seminal work on search asymmetry implemented similar stimuli to examine how easily participants could detect a target with a distinguishing feature, such as the slanted bar in Q (A. Treisman & Souther, 1985).

4.3.3. Design

Similar to Study 1, each trial consisted of two search arrays: pre-imagery and post-imagery. Both arrays presented either three or six letters (set size) arranged in a circular formation around a central fixation cross. This manipulation was necessary to measure slope reaction time. For the six-letter array, letters were positioned at 60° intervals around the fixation cross, forming a hexagon configuration. For the three-letter

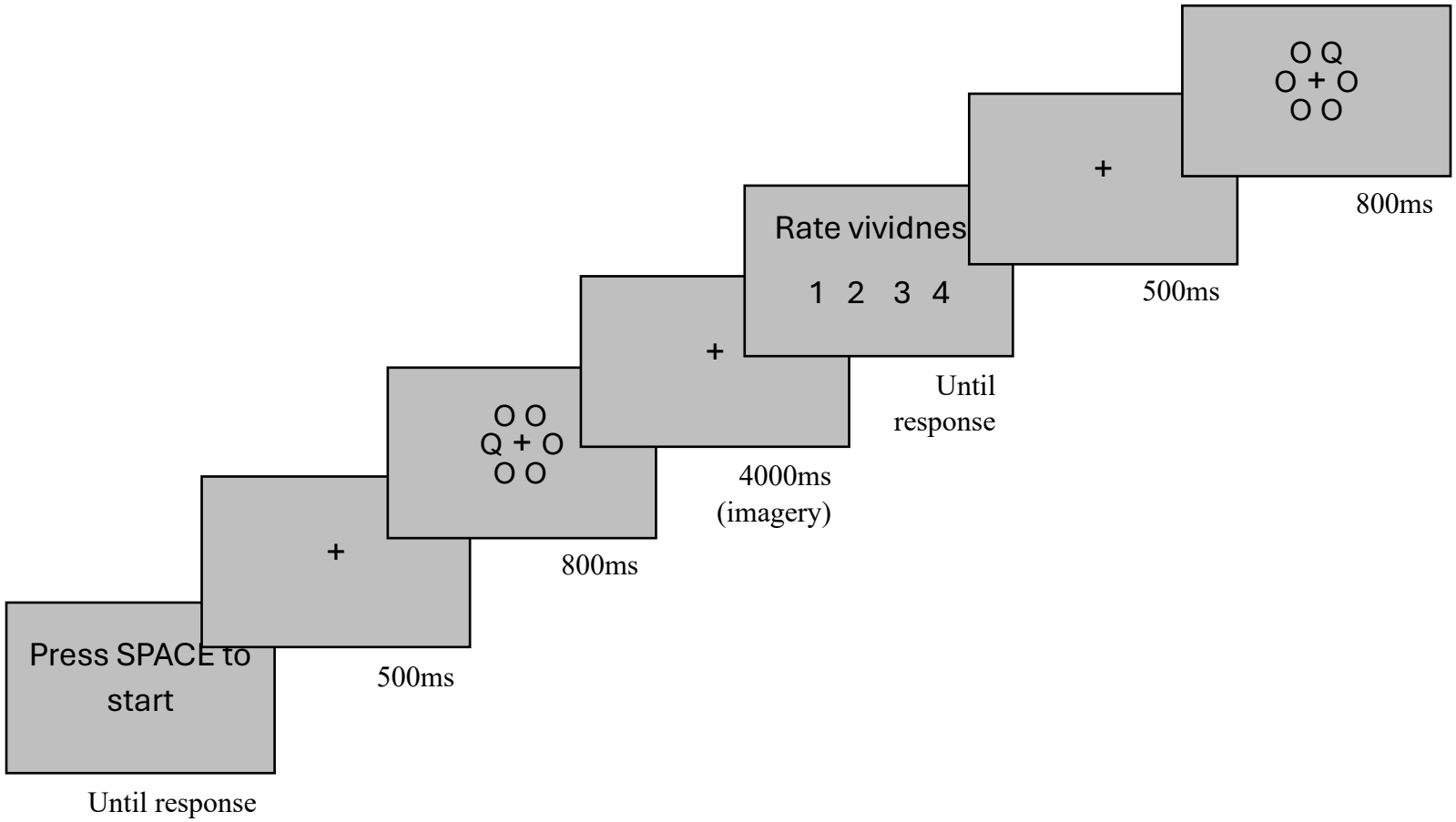
array, letters were positioned randomly at three of the six locations in the six-letter array. The pre-imagery array presented the target as a unique letter (e.g., Q) and all nontargets as the other letter (O). In contrast to Study 1, the post-imagery array presented items in one of two ways (trial types): the congruent trial switched the target and nontargets from pre-imagery; the incongruent trial repeated the target and nontargets from pre-imagery.

4.3.4. Procedure

Figure 4.1 shows a sample trial. Before starting the experimental run, participants completed a practice run with 10 trials to get familiar with the task. The experimental run included 8 blocks of 24 trials (192 trials), evenly split between trial types (congruent and incongruent). Trials for each trial type were further split depending on whether the target in post-imagery array had a present feature (target Q) or an absent feature (target O). The sequence of displays for each trial was identical to those in Experiments 2 and 3. Because the arrangement of letters differed from the cross configuration done in Experiment 1 (Figure 3.2), the search instruction was modified, such that participants were instructed to respond whether the odd-one-out target appeared on the left or right side of the array, using the left or right arrow keys. Additionally, participants were instructed to imagine the letter opposite to the target in pre-imagery (Cochrane, Ng, & Milliken, 2021). For example, if the target was O, then participants would imagine Q, and vice versa.

Figure 4.1

Sample trial



Note. Same task for Experiments 4 and 5. Imagery instruction was to imagine the opposite letter in pre-imagery (e.g., imagine O if Q was detected). Set size was three or six letters.

4.3.5. EEG system and data preprocessing

For Experiment 5, the recording and steps for data preprocessing were identical to those in Study 1. Four participants were excluded from final analyses due to excessive artifacts.

4.4. Results

4.4.1. Experiment 4

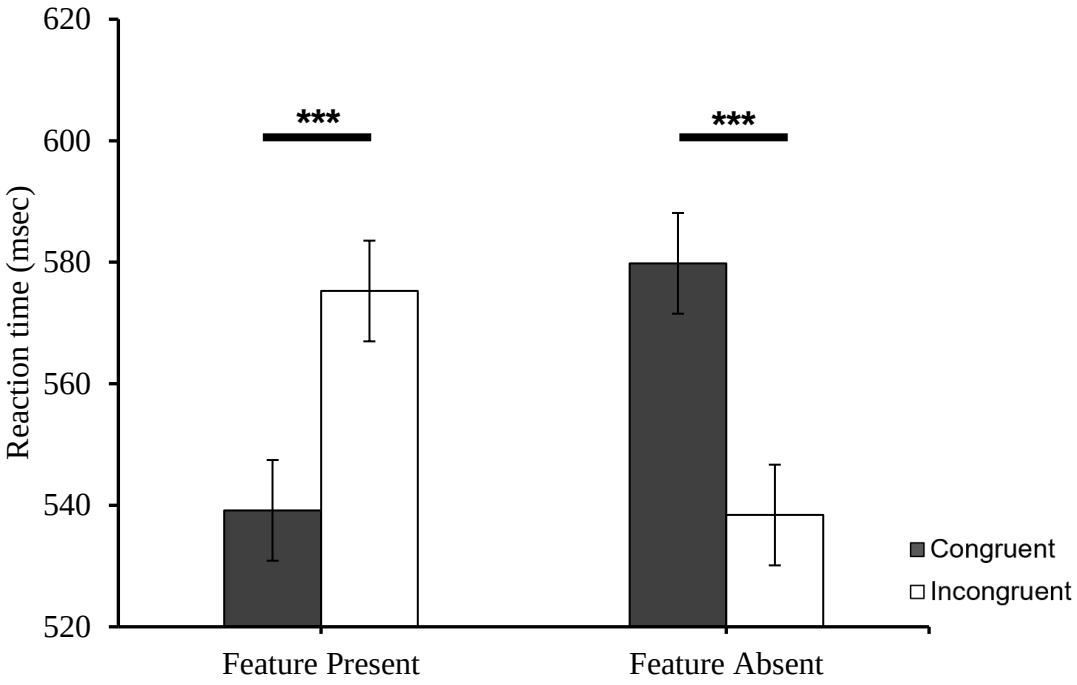
The goal was to examine how the top-down effects of imagery interacted with the pop-out effects. Only behavioral data were recorded. There were two predictions. First, using the average reaction time measure, I predicted faster responses for congruent targets than incongruent. Second, using the slope reaction time measure, I predicted better search efficiency when the feature was present than when it was absent.

4.4.1.1. Average reaction time

Figure 4.2 shows the average reaction time. A two-way ANOVA was conducted using average reaction time data, with trial type (congruent vs. incongruent) and feature (present vs. absent) as factors. There was an interaction between trial type and feature, $F(1,31) = 116.96$, $p < .001$, $\eta_p^2 = 0.79$. To investigate this interaction, Bonferroni-corrected comparison (corrected alpha = .05/2 = .025) were conducted, with comparisons between trial types based on feature. When the feature was present, search was faster for congruent targets ($M = 539.16$ ms, $SD = 52.32$) than for incongruent targets ($M = 575.27$ ms, $SD = 52.41$), $t(31) = 7.54$, $p < .001$, $d = 0.71$. When the feature was absent, search was faster for incongruent targets ($M = 538.40$ ms, $SD = 49.34$) than for congruent targets ($M = 579.82$ ms, $SD = 49.90$), $t(31) = -8.65$, $p < .001$, $d = -0.81$.

Figure 4.2

Averaged reaction time (Experiment 4)



Note. Asterisks represent significance (***$p < .001$). Error bars represent ± 1 SE.

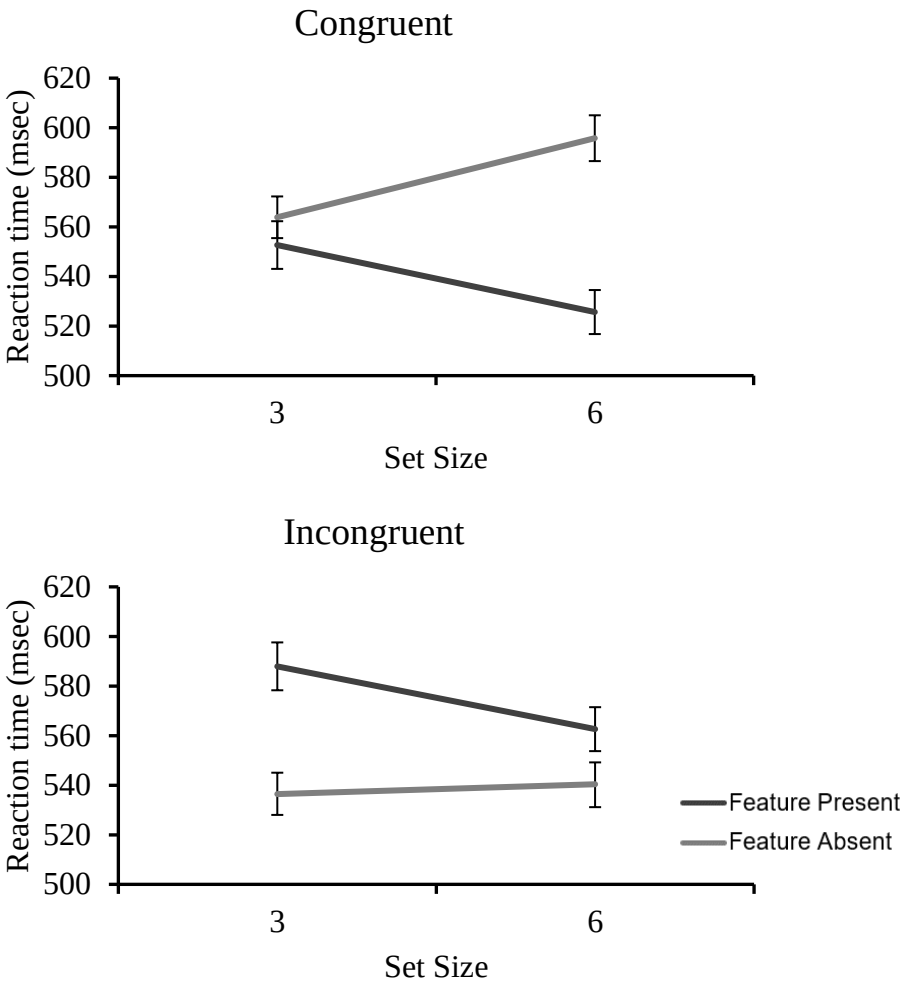
4.4.1.2. Slope of reaction time

Figure 4.3 shows the slope reaction time. Slope reaction time was calculated based on the set size. A two-way ANOVA was conducted using slope reaction time data, with feature and trial type as factors. There was a main effect of feature, $F(1,31) = 26.82$, $p < .001$, $\eta_p^2 = 0.46$, with a smaller slope for present feature (M = -8.72 ms/item) than for absent feature (M = 5.98 ms/item). There was a main effect of trial type, $F(1,31) = 7.14$, $p = .01$, $\eta_p^2 = 0.19$, with a larger slope for congruent targets (M = 0.81 ms/item) than for incongruent targets (M = -3.55 ms/item). There was an interaction between feature and trial type, $F(1,31) = 4.84$, $p = .035$, $\eta_p^2 = 0.14$. Bonferroni-corrected comparisons (corrected alpha = .05/2 = .025) were conducted based on features for each trial type. For

congruent targets, the slope was larger when the feature was absent (M = 10.63 ms/item), relative to when the feature was present (M = -9.01 ms/item), $t(31) = 5.43$, $p < .001$, d = 1.46. For incongruent targets, the slope also was larger when the feature was absent (M = 1.32 ms/item), relative to when the feature was present (M = -8.43 ms/item), $t(31) = 2.70$, $p = .009$, $d = 0.72$.

Figure 4.3

Slope reaction time (Experiment 4)



Note. Slope reaction time as a function of set size (3 or 6 search items). Error bars represent ± 1 SE.

4.4.1.3. Summary

Findings revealed faster reaction time for congruent targets than incongruent targets when the salient feature was present, consistent with previous imagery-search studies (e.g., Cochrane et al., 2018). Surprisingly, the opposite was observed when the salient feature was absent. Additionally, the slope of reaction time in relation to set size varied based on the feature. The slope was steeper (indicating increased reaction time per additional item) when the feature was absent compared to when it was present. Building on these findings, Experiment 5 implemented the same paradigm with the addition of EEG to examine neural effects (via N2pc) on attentional selection.

4.4.2. Experiment 5

There were two goals. First, I aimed to replicate the behavioral effects from Experiment 4. Second, I measured neural effects using the N2pc component to examine the interaction between top-down imagery and pop-out effects during selection. I predicted that the neural effects would share a similar pattern of results with the reaction time data, with larger components for congruent targets than incongruent.

4.4.2.1. N2pc mean amplitude

Figure 4.4 shows the grand-average parent (contralateral and ipsilateral) and difference (contralateral minus ipsilateral) waveforms. Figure 4.5 shows the N2pc mean amplitudes. A two-way ANOVA was conducted using N2pc mean amplitude data, with trial type (congruent vs. incongruent) and feature (present vs. absent) as factors. There was an interaction between trial type and feature, $F(1,19) = 5.94$, $p = .025$, $\eta_p^2 = 0.24$. To investigate this interaction, Bonferroni-corrected comparison (corrected alpha = .05/2 =

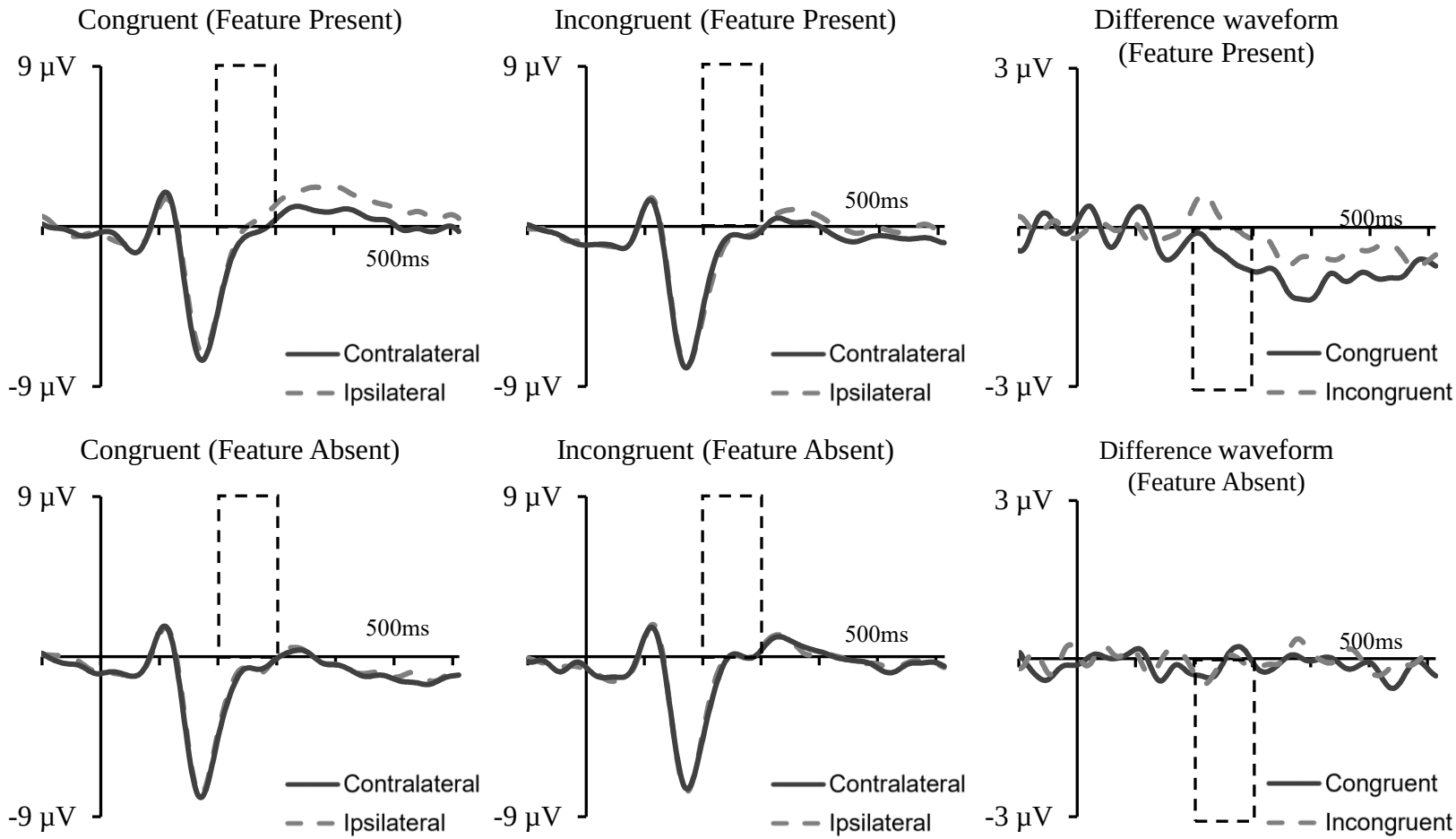
.025) were conducted, with comparisons between trial types for each feature. When the feature was present, the amplitude was larger for congruent targets (M = -0.55 μV, SD = 1.08) than incongruent targets (M = 0.15 μV, SD = 0.68), $t(19) = 2.81$, $p = .008$, $d = 0.86$. When the feature was absent, the amplitude did not differ between congruent (M = -0.09 μV, SD = 0.72) and incongruent targets (M = -0.16 μV, SD = 0.71), $t(19) = -0.27$, $p = .17$, $d = -0.8$. There were no main effects of trial type or feature, $F < 2.68$. The amplitude for congruent targets with present features was significantly different relative to zero, $t(19) = -2.27$, $p = .035$, $d = -0.51$. All other amplitudes were not significant, $t < |0.54|$.

4.4.2.2. N2pc onset latencies

A two-way ANOVA was conducted using N2pc onset latencies (jackknife latency analysis, Kiesel et al., 2008), with trial type (congruent vs. incongruent) and feature (present vs. absent) as factors. There was a main effect of trial types, $F(1,19) = 154.58$, $p < .001$, $\eta_p^2 = 0.89$, and a main effect of feature, $F(1,19) = 1455.80$, $p < .001$, $\eta_p^2 = 0.99$. There was an interaction between trial type and feature, $F(1,19) = 37.05$, $p < .001$, $\eta_p^2 = 0.66$. Bonferroni-corrected comparison (corrected alpha = .05/2 = .025) were conducted, with comparisons between trial types for each feature. When the feature was present, the onset was earlier for congruent targets (M = 271.00 ms, SD = 25.96) than incongruent targets (M = 326.00 ms, SD = 2.05), $t(19) = 13.03$, $p < .001$, $d = 4.13$. When the feature was absent, the onset was earlier for congruent targets (M = 173.80 ms, SD = 4.98) than for incongruent targets (M = 191.90 ms, SD = 1.99) , $t(19) = 4.30$, $p < .001$, $d = 1.36$.

Figure 4.4

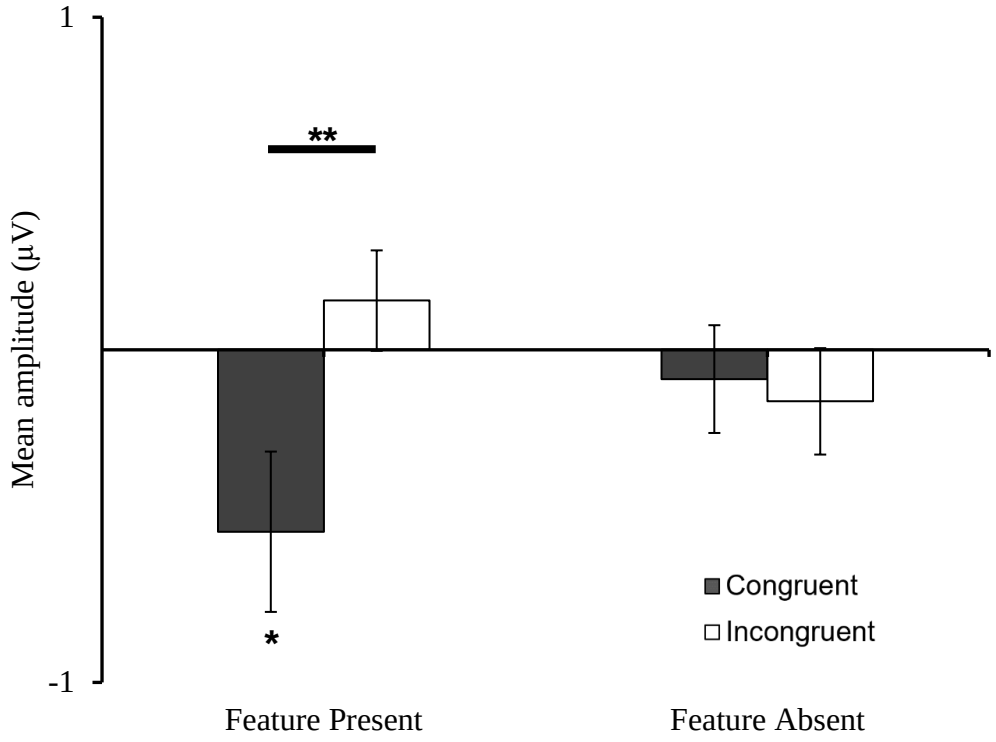
Grand averaged waveforms (Experiment 5)



Note. Parent and difference waveforms in post-imagery search. N2pc time window at 200 and 300 msec (dashed rectangle).

Figure 4.5

N2pc mean amplitudes (Experiment 5)



Note. Average of signals from 200 to 300 ms. Asterisks represent significance (*$p < .05$; **$p < .01$). Error bars represent ± 1 SE.

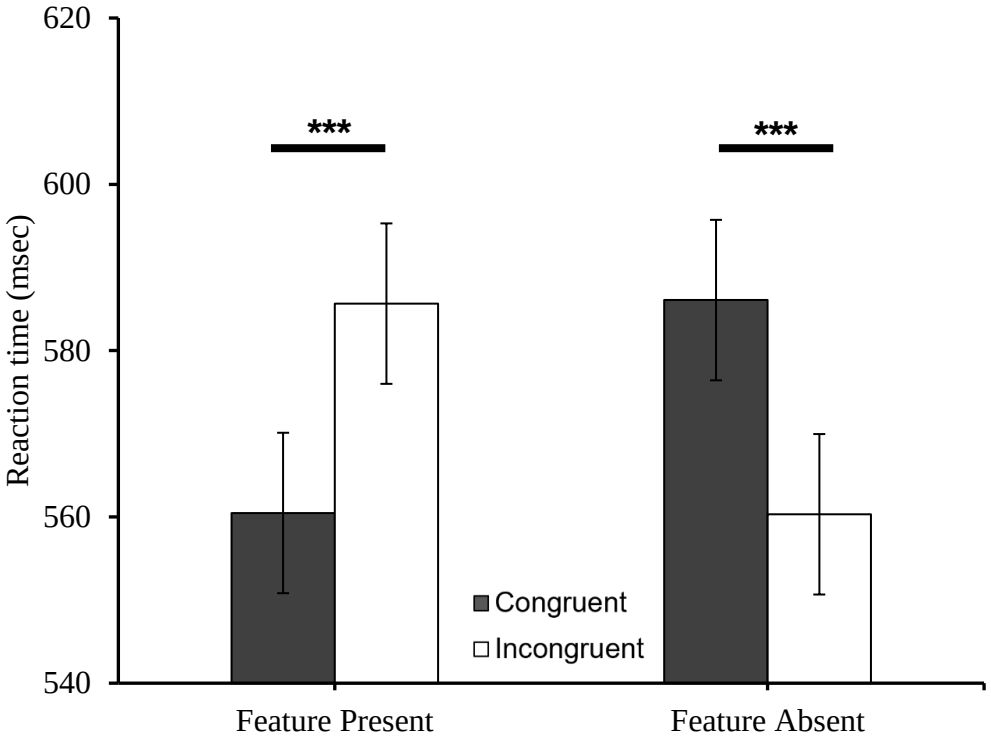
4.4.2.3. Average reaction time

Figure 4.6 shows the average reaction time. A two-way ANOVA was conducted using average reaction time, with trial type (congruent vs. incongruent) and feature (present vs. absent) as factors. There was an interaction between trial type and feature, $F(1,19) = 24.20$, $p < .001$, $\eta_p^2 = 0.56$. Bonferroni-corrected comparison (corrected alpha = .05/2 = .025) were conducted between trial types based on feature. When the feature was present, search was faster for congruent targets ($M = 560.50$ ms, $SD = 53.36$) than for incongruent targets ($M = 585.65$ ms, $SD = 38.81$), $t(19) = 3.86$, $p < .001$, $d = 0.53$. When

the feature was absent, search was faster for incongruent targets (M = 560.31 ms, SD = 52.19) than for congruent targets (M = 586.1 ms, SD = 42.21), $t(19) = 3.95$, $p < .001$, d = -0.55. There were no main effects, $F < 24.20$.

Figure 4.6

Average reaction time (Experiment 5)



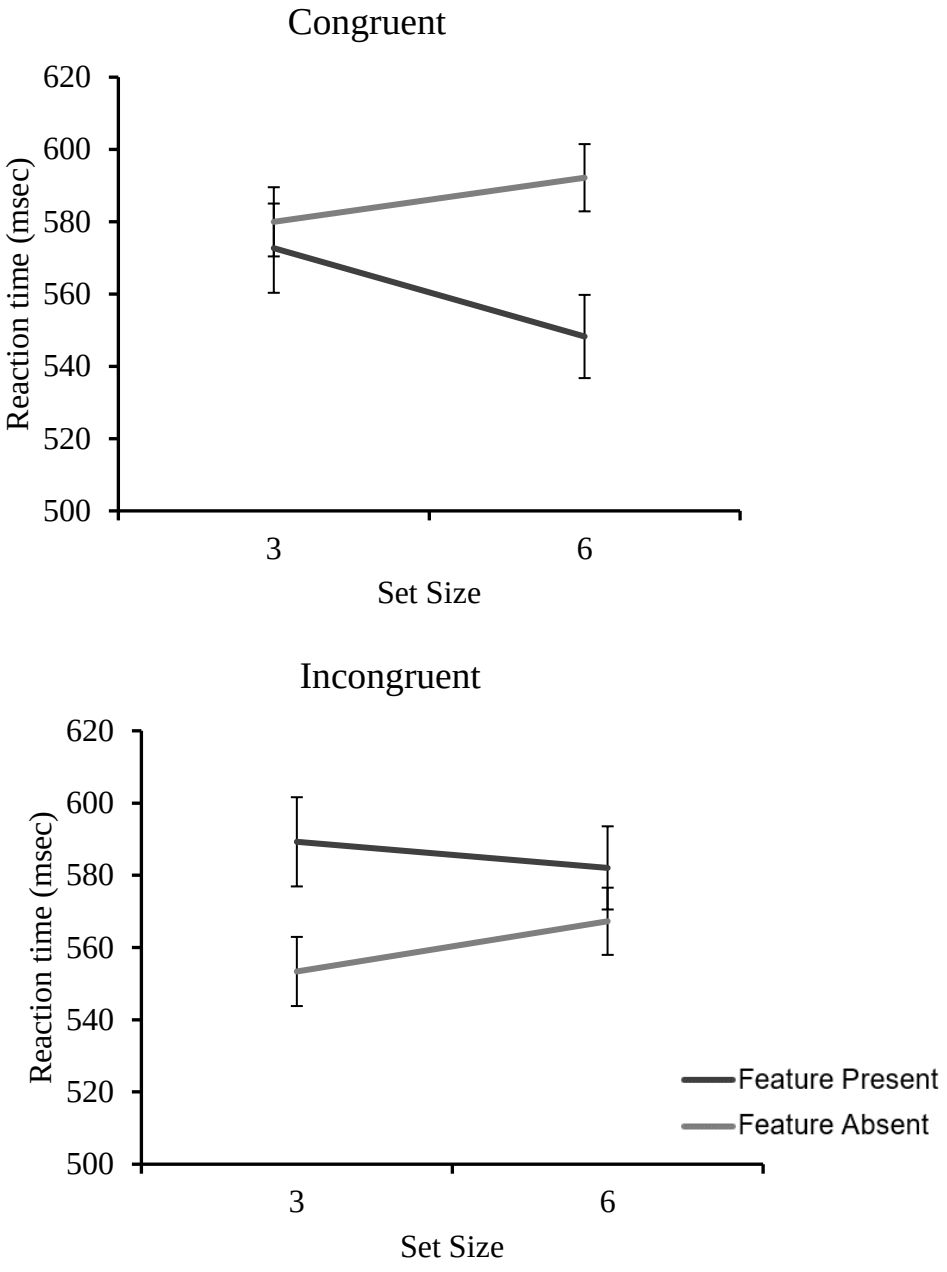
Note. Asterisks represent significance (***$p < .001$). Error bars represent ± 1 SE.

4.4.2.4. Slope reaction time

Figure 4.7 shows the slope reaction time. A two-way ANOVA was conducted using slope reaction time data, with feature and trial type as factors. There was a main effect of feature, $F(1,19) = 16.54$, $p < .001$, $\eta_p^2 = 0.47$, with a larger slope for absent feature (M = -5.28, SD = 12.31) than for present feature (M = 4.35, SD = 12.78). There was no main effect of trial type nor interaction, $t < 1.46$.

Figure 4.7

Slope reaction time (Experiment 5)



Note. Slope reaction time as a function of set size (3 or 6 search items). Error bars represent ± 1 SE.

4.4.2.5. Summary

There were three main findings. First, when the feature was present, the N2pc was larger and faster for congruent targets than incongruent targets, consistent with what was observed in Study 1. When the feature was absent, the N2pc was faster (but not different in magnitude) for congruent than incongruent targets. This finding was unexpected, as Experiment 4 revealed faster reaction for incongruent than congruent targets when the feature was absent. Second, behavioral findings regarding the average reaction time were replicated. Third, findings regarding the slope reaction time was partially replicated, with larger slopes for absent than present feature. Although there was no interaction as observed in Experiment 4, the overall findings were consistent with what have been reported in asymmetric search studies (Treisman and Souther,1985; Wolfe, 2001).

4.5. General discussion

4.5.1. Summary and interpretation of findings

I examined the question of how top-down imagery interacted with bottom-up pop-out effects. To address this question, I used asymmetric search paradigms by manipulating the presence of salient features on the target. Experiment 4 tested the behavioral effects based on overall search performance (average reaction time) and search efficiency (slope reaction time). Experiment 5 aimed to replicate these effects and also tested the neural effects based on the N2pc components. There were three major findings. First, the average reaction time was faster for congruent targets than incongruent targets when the salient feature was present (Q among Os), as predicted. Surprisingly, the reverse was found when the salient feature was absent (O among Qs).

Second, search was more efficient when the feature was present (flatter slope) than when absent (positive slope), consistent with what has been observed in asymmetric search studies (Dosher et al., 2004; Saiki et al., 2005; Shen & Reingold, 2001; Treisman & Gormican, 1988). These results were present irrespective of the trial type. Third, the N2pc effects between trial types were observed when the feature was present but not when absent, partially aligning with the effects with average reaction time. Interestingly, the N2pc onset was faster for congruent targets than incongruent targets, irrespective of the feature.

The average reaction time results supported the goal-driven account, but in a surprising way. When the salient feature was present, top-down effects aligned with pop-out effects for congruent targets but misaligned for incongruent targets. This alignment facilitated faster reaction time for congruent targets, likely because the same feature was being processed by imagery and bottom-up processing. Other imagery-search studies shared similar findings when imagery and congruent pop-out targets were matched (e.g.,(Cochrane, Townsend, Lee, et al., 2021). However, the opposite effect was observed when the salient feature was absent. According to the goal-driven account, it was predicted that imagery would still guide attention, even in the absence of pop-out. Surprisingly, the results did not support this prediction. Reaction times for congruent targets were slower when the salient feature was absent, suggesting that the absence might reduce the effectiveness of imagery in guiding attention. This could imply that imagery relies, at least in part, on the presence of distinctive visual features to effectively

direct attention, a concept that aligns with the notion that attentional guidance may vary based on the availability of visual information (Bravo & Nakayama, 1992).

The slope reaction time results supported the stimulus-driven account. Search was more efficient when the target possessed the salient feature, regardless of whether the target was congruent or incongruent. This suggests that saliency automatically captured attention, leading to flatter slopes. This finding is consistent with the stimulus-driven account, which posits that pop-out effects are driven by bottom-up processes and occur independently of top-down influences (Treisman & Gormican, 1988; Wolfe, 1994). The absence of systematic differences in search efficiency based on imagery conditions supports the idea that salient features partially operate through a mechanism that does not necessitate top-down modulation (Saiki et al., 2005; Shen & Reingold, 2001).

The N2pc findings largely supported the goal-driven account, particularly when the salient feature was present. The N2pc component was larger and faster for congruent targets than for incongruent targets, mirroring the pattern observed in the average reaction time data. These results suggest that imagery not only influenced behavioral outcomes but also has a measurable impact on the neural mechanisms underlying attention. The faster onset of the N2pc for congruent targets suggests that imagery guided attention (Cochrane et al., 2018, 2021), in line with the goal-driven account. However, when the salient feature was absent, the N2pc results were more complex. The onset of the N2pc was faster for congruent targets, but the magnitude did not differ significantly between congruent and incongruent trials. The lack of difference in magnitude could suggest that in the absence of pop-out, there is a limitation to the impact of imagery. An alternative

explanation is that there were no top-down effects in line with the stimulus-driven account. Perhaps without a salient feature, search was conducted in a serial manner (Saiki et al., 2005), evinced by the observed positive slopes. While this point cannot be ruled out, perhaps a simpler explanation is that the N2pc mean amplitude was not sensitive enough to capture the top-down effects. Previous studies showed that the N2pc was more temporally variable in search tasks without pop-out than tasks with pop-out (Dowdall et al., 2012), meaning that the expected time window did not fully capture the N2pc component (Luck, 2014).

Furthermore, a strict stimulus-driven view cannot explain the faster onset for congruent targets. One possible explanation is that attention was shifted initially to a congruent target, but subsequent attentional processes were hindered by the lack of saliency. Still, it was paradoxical that reaction time was slower for congruent targets despite faster N2pc onset. Perhaps the onset that was observed does not reflect a fast selection of the target but one of the nontargets instead, given that the salient feature was presented in the nontargets. Related search studies showed similar effects, in which nontargets that contain task-relevant features (Moon et al., 2022) or that were categorically similar to the target (Nako et al., 2014; Moon et al., 2022; Wu et al., 2013, 2015, 2018) evoked N2pc components when the target was absent. In other words, the distraction of salient features might have triggered the initial selection of attention, slowing down subsequent behavioral responses.

4.5.2. Theoretical contributions

This study revealed a relationship between top-down imagery and pop-out effects more intricate than predicted by either the stimulus-driven or goal-driven accounts. The findings suggest that imagery functions as a top-down process that modulates pop-out effects. These results oppose the strict stimulus-driven account that top-down processes do not impact bottom-up processing. Additionally, the findings warrant a mechanistic explanation for asymmetry. There are competing ideas about whether asymmetric search reflects the enhancement of target processing (A. Treisman & Gormican, 1988; Q. Wang et al., 1994) or of distractor suppression (Malinowski & Hübner, 2001; Shen & Reingold, 2001). Distractor processing is the ability to filter out irrelevant information and is separate from target enhancement (Wöstmann et al., 2022). Given that the two are separate, top-down effects on either of the two should be distinguished.

As demonstrated in Study 1 and here, the findings suggest that imagery impacts target enhancement, evinced by the N2pc effects. Several ERP studies showed that the N2pc component is better suited for target enhancement than distractor suppression (Mazza et al., 2009). Alternatively, recent studies marked suppression using the P_d component (distractor positivity), which was observed before the N2pc component (Sawaki et al., 2012; Sawaki & Luck, 2013). Does imagery also impact distractor suppression? Unfortunately, the present study was not designed to measure the P_d component, though perhaps the reversed behavioral effect observed in the feature absent trials (O among Qs) can be explained by increased suppression of nontargets. For incongruent trials, perhaps imagery of salient features suppressed attention to matching

features of the nontargets, thus allocating attention to the target faster and consequently behavior. In contrast, for congruent trials, imagery did not possess salient features, thus no suppression of nontargets could occur, resulting in slower behavior. Putting it together, perhaps imagery enhanced target processing when pop-out effect occurred and also increased distractor suppression when imagery matched salient features of nontargets. It cannot be ruled out that both occurred irrespective of the feature, which raises questions as to whether distractor suppression also was impacted in Study 1. Future studies should disentangle the effects on target enhancement and distractor suppression more clearly.

4.5.3. Limitations

Study 1 had two limitations that were addressed here. First, the results may have been confounded by intertrial priming. Previous imagery-search studies instructed participants to imagine a color that was not previously selected (e.g., imagine green if red shape was chosen; Cochrane et al., 2018). Imagery was pitted against intertrial priming, with findings showing the typical behavioral effects according to imagery, not whether the target and nontargets repeated or switched (Cochrane et al., 2018; Cochrane, Ng, Khosla, et al., 2021; Cochrane, Wang, Pratt, et al., 2021). The present study deployed the same instructions and yielded similar results, increasing confidence that the results in Study 1 were due to imagery, not priming. Second, because Study 1 used a large stimulus set with complex items, there inevitably were unsystematic differences in feature dimensions not accounted for. These differences raised concerns over whether and how strongly pop-out effects occurred. The present study systematically controlled these differences by distinguishing targets and nontargets by a single feature.

There were three limitations in this study. First, although the change to imagery instruction addressed intertrial priming within trial, the repeated presentation of two letters across trials might have led to priming between trials. This priming could have influenced the results, as participants became more adept at recognizing the letters over time. Future research should consider incorporating neutral trial types, similar to those used in Experiment 3, with other familiar letters distinguished by a single feature (e.g., T and I) to minimize potential priming effects.

Second, participants between Study 1 and Study 2 may have engaged distinct executive functions. Previous search studies distinguished the role of working memory (WM) and long-term memory (LTM) in attentional model (see Wolfe, 2021). These studies assigned different targets for each trial to test WM-related effects or the same target repeatedly each trial to test LTM-related effects (Woodman & Chun, 2006). Study 1 aligns with the former approach, while Study 2 aligns with the latter. Recent studies explored the representational aspects of imagery from working memory models (Keogh & Pearson, 2014, 2017). The Reverse Hierarchy Model states that imagery retrieves long-term memory in the MTL (see Figure 1.2; Pearson, 2019). It is unclear whether specific executive functions contributed to the observed effects in this study.

Finally, the potential influence of familiarity effects cannot be ruled out. Previous studies showed that searching for an unfamiliar target among familiar nontargets was more efficient than a familiar target among unfamiliar nontargets (Guo et al., 2021; Malinowski & Hübner, 2001; Mruczek & Sheinberg, 2005; Shen & Reingold, 2001; Q. Wang et al., 1994; Wolfe & Friedman-Hill, 1992; Zhang & Onyper, 2020). A recent

study decoded activity patterns of familiar items during an imagery task (Ragni et al., 2021). It is possible that the effects typically observed between congruent and incongruent trials were intermixed with familiarity effects of the search items. Familiarity with nontargets may have facilitated search in addition to target enhancement by imagery (Guo et al., 2021; Mruczek & Sheinberg, 2005; Zhang & Onyper, 2020). In congruent trials, perhaps familiarity of nontargets added to facilitation (Guo et al., 2021; Mruczek & Sheinberg, 2005; Zhang & Onyper, 2020). In incongruent trials, familiarity of nontargets might have suppressed the distraction caused by imagery-matching features (Shen & Reingold, 2001). More research is needed to explore these possibilities in greater detail.

4.5.4. Conclusion

In conclusion, this study tested the theoretical accounts of attention by examining the impact of imagery on asymmetric search. The interaction between imagery and pop-out was revealed to be intricate in ways that cannot fully be accounted for purely by the stimulus-driven nor goal-driven account. Overall search performance was influenced by imagery, whereas search efficiency was influenced by saliency. Although evidence suggests that selection to saliency was guided by imagery, the N2pc findings did not entirely align with the behavioral effects, warranting further research to unravel the precise ways in which top-down and bottom-up processes interact.

Chapter 5: Selection vs. Post-selection

5.1. Abstract

Selection and post-selection are understood as two separate stages in search. Selection reflects the allocation of attention to task-relevant features, marked by the N2pc component. Post-selection reflects the extraction and analysis of these features, marked by the SPCN component. Systematically driven effects on either stages may have subsequent effect on behavior, necessitating a clear distinction between what factors impact a certain stage. The majority of imagery-search studies provided behavioral findings, which alone cannot address this issue. Study 3 examined how imagery influences these stages separately by manipulating search instructions differentiated by recognition demands. Participants completed a 'locate' task in which they responded to the location of a pop-out target (low recognition demand), and a 'identify' task in which they responded to the feature of the same target (high recognition demand). N2pc and SPCN effects separately were tested by comparing imagery trial types (congruent and incongruent) for each task. ERP findings showed larger N2pc components in congruent than in incongruent trials for both tasks, suggesting enhanced selection by imagery. Surprisingly, SPCN effects were observed in the unexpected locate task but not in the expected identify task. These findings suggest that imagery may inadvertently engage post-selection without explicit requirement. Further investigation is needed to understand how imagery interacts with this stage to better understand its impact on behavior.

5.2. Introduction

Previous search studies examined attentional processes using the N2pc component as a neural marker for selection (Eimer, 1996; Luck, 2011). However, the act of selecting an item does not guarantee its immediate recognition. Rather, a subsequent post-selection process is required to analyze the features of the selected item (Huang & Pashler, 2007; Xu & Chun, 2009). Early studies have distinguished between the processes of selection and post-selection. Selection is an early, feature-driven process, influenced by the differences between targets and nontargets (Duncan & Humphreys, 1989). Post-selection is a later process, driven by the complexity of the features within the selected items (Xu & Chun, 2009).

This distinction raises an important question: If imagery induces behaviorally relevant responses during search, do these responses stem exclusively from the selection stage, or do they also involve the post-selection stage? To address this question, it is crucial to differentiate the effects of imagery on these two distinct stages. Several studies have utilized ERP components, examining the mean amplitude to compare neural effects between different conditions (Brisson & Jolicoeur, 2007; Mazza et al., 2007). Differences in the N2pc component have been interpreted as evidence of effects on the selection process, while differences in the SPCN (sustained posterior contralateral negativity) component have been associated with effects on the recognition process (Mazza et al., 2007). The SPCN emerges from the same waveform as the N2pc, occurring shortly after the N2pc (~ 400 ms after stimulus onset) and persisting for a longer duration (Berggren & Eimer, 2018; Grubert et al., 2017; Mazza et al., 2007).

The presence of these two ERP components is contingent upon the specific demands of the search task. For example, tasks requiring only spatial attention, such as those where participants allocate attention to different positions within an array, typically elicit the N2pc component (Berggren & Eimer, 2018; Grubert et al., 2017; Mazza et al., 2007). In contrast, tasks with high recognition demand, where participants must not only locate the target but also identify its features, often elicit both the N2pc and SPCN components (Eimer, 2014). Mazza et al. (2007) showed this distinction. In tasks where participants located a target with a distinctive feature (e.g., line orientation), only the N2pc was observed, whereas in tasks requiring both location and feature identification (e.g., determining left or right orientation), both the N2pc and SPCN components were present.

Given that imagery has the potential to influence selection, it stands to reason that imagery might also impact post-selection. Study 3 (Experiment 6) examined whether the effects of imagery extend beyond the initial selection stage to influence the post-selection processes involved in object recognition. Studies 1 and 2 consistently showed N2pc effects, suggesting that imagery at least impacts selection. If imagery also impacts post-selection, then tasks that have high recognition demands should be traceable by the SPCN components.

5.3. Methods

5.3.1. Participants

Twenty undergraduates participated in Experiment 6. Two additional participants were excluded in the final analyses due to excessive eye-movement artifacts. All

participants were recruited via UC Riverside's Human Subject Research Pool and were compensated with Psychology course credit. All participants self-reported normal or corrected-to-normal vision. Participants were excluded if more than 30% of their EEG data were removed and/or accuracy was less than 70%. No analyses were conducted until data collection was complete.

5.3.2. Stimuli and design

The stimuli were red and green outlined squares that contained a gap in one of the sides (left, right, top, or bottom). Basic shapes were widely used in singleton search tasks to systematically control for distinctive features between targets and nontargets (Wolfe et al., 2010). The goal was to design one search task with two separate search instructions that differ in recognition demands. This set-up required manipulating the task instructions, such that either the position of the target or the feature of the target was used as the goal. The search displays contained one target square with one color (e.g., red) and three nontarget squares with the other color (green). The squares were presented in a cross configuration surrounding a black fixation cross in the center., subtending 3.6°. Trial types were manipulated identical to Study 1: the congruent trials presented a target that matched with a previously imagined stimulus and nontargets that did not match; the incongruent trials presented a target that did not match a previously imagined stimulus and nontargets that did match. This manipulation was necessary to examine differences in the outcome measures between congruent and incongruent targets.

5.3.3. Procedures

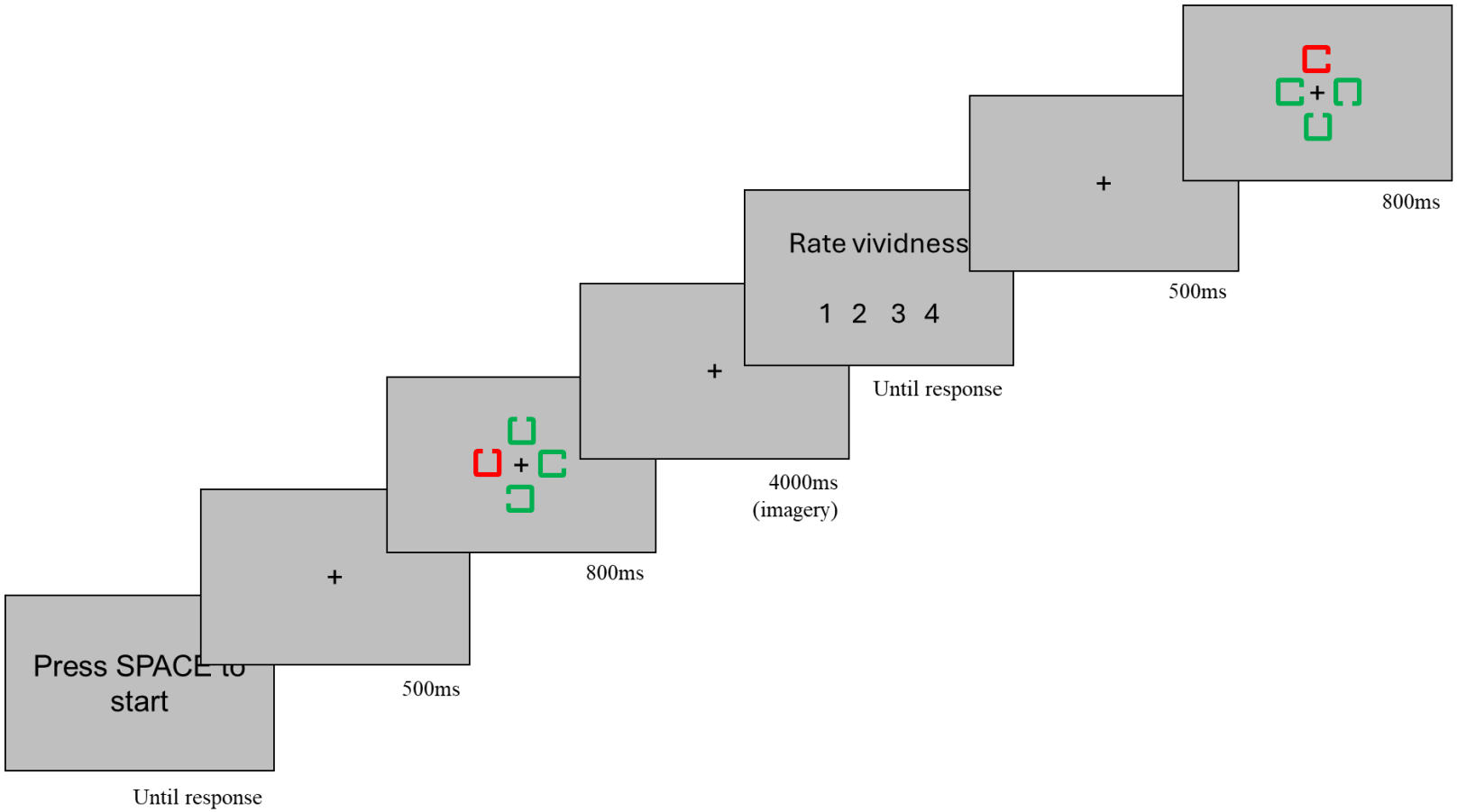
Figure 5.1 shows a sample trial. Participants completed two experiment runs, one with its own search instruction (locate and identify). Before starting each experimental run, participants completed a practice run with 10 trials to get familiar with the task. Each experimental run included 5 blocks of 32 trials (160 trials), evenly split between trial types (congruent and incongruent). The sequence of displays for each trial was identical to those in Experiments 2, 3, 4 and 5. When search items appeared in pre- and post-imagery, participants responded to the location of the pop-out target (locate task) or to the location of the missing gap on the pop-out target (identify task). The former required only locating a unique colored square (four possible locations), whereas the latter requires identifying the gap of the square (four possible gaps). For both tasks, participants gave their responses using the arrow keys corresponding to the location of the target or gap (e.g., pressing the left arrow key when the target or the gap was on the left).

5.3.4. EEG system and data preprocessing

The recording and steps for data preprocessing were identical to those in Studies 1 and 2. Two participants were excluded from final analyses due to excessive artifacts.

Figure 5.1

Sample trial



Note. Same paradigm for locate and identify task. Participants were instructed to locate the odd color (locate task) or identify the gap of the odd color (identify task).

5.4. Results

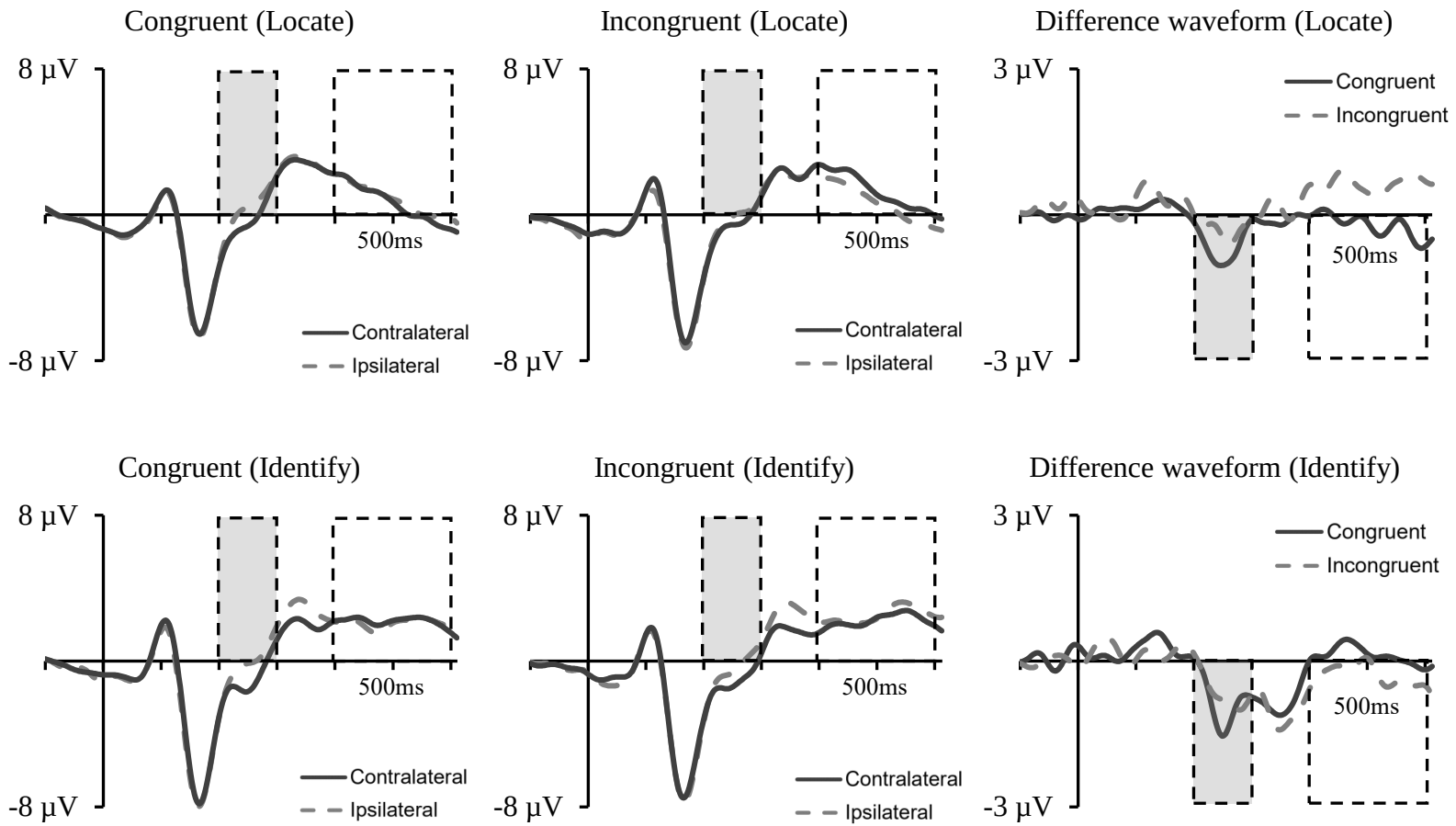
The goal was to examine differences in imagery-related effects on selection and post-selection. The N2pc and SPCN components were used to mark these stages separately.

5.4.1. N2pc mean amplitude

Figure 5.2 shows the grand-average parent (contralateral and ipsilateral) and difference (contralateral minus ipsilateral) waveforms. The N2pc time window is marked from 200 to 300 ms (dashed grey rectangle). Figure 5.3 shows the N2pc mean amplitudes. A two-way ANOVA was conducted using N2pc mean amplitude data, with trial type (congruent vs. incongruent) and task (locate vs. identify) as factors. There was a main effect of trial type, with a larger amplitude for congruent targets (M = -0.80 μV, SD = 1.24) than incongruent targets (M = -0.43 μV, SD = 1.20), $F(1,19) = 4.97$, $p = .038$, $\eta_p^2 = 0.21$. There was no main effect of task nor interaction, $F < 2.17$. For locate search, the amplitude for congruent target (M = -0.71 μV, SD = 1.09) was significantly different from zero, $t(19) = -2.91$, $p = .009$, $d = -0.65$, but not for incongruent target (M = -0.20 μV, SD = 1.29), $t(19) = -0.70$, $p = .50$, $d = -0.16$. For identify search, the amplitude was significant for congruent target (M = -0.89 μV, SD = 1.39), $t(19) = -2.85$, $p = .01$, $d = -0.64$, and for incongruent target (M = -0.65 μV, SD = 1.11), $t(19) = -2.64$, $p = .016$, $d = -0.59$.

Figure 5.2

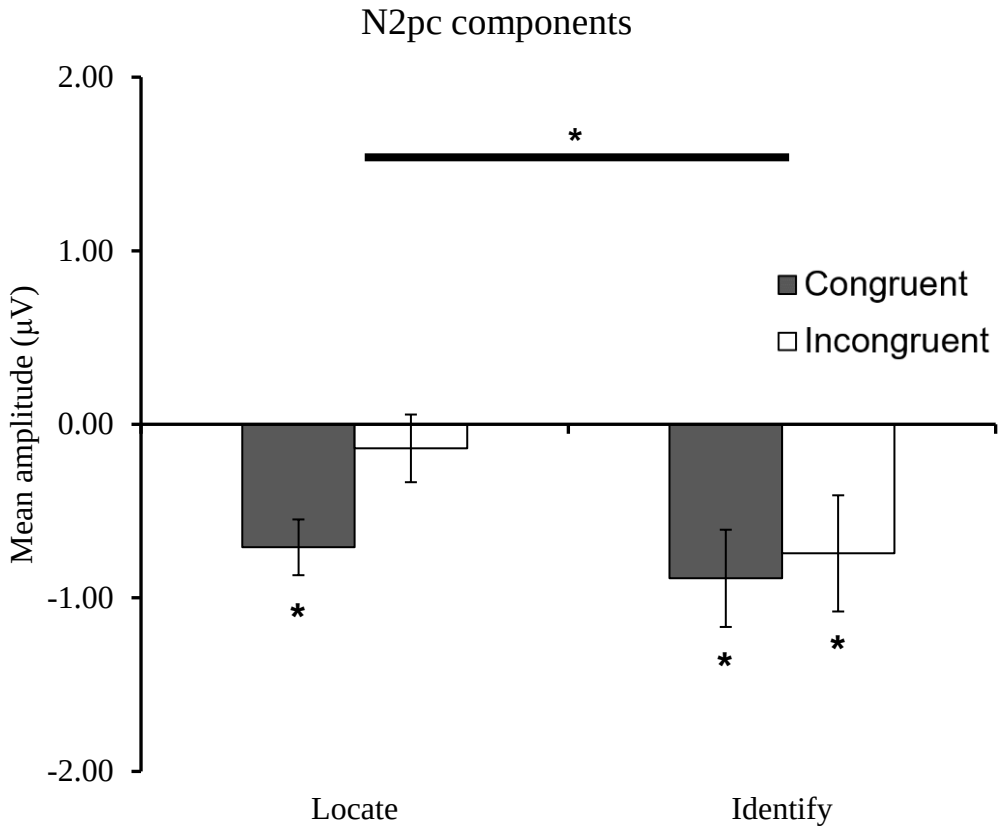
Grand averaged waveforms (Experiment 6)



Note. Parent and difference waveforms in post-imagery search. N2pc time window at 200 and 300 msec (dashed grey rectangle). SPCN time window at 400 and 600 msec (dashed blank rectangle).

Figure 5.3

N2pc mean amplitudes (Experiment 6)



Note. Average of signals from 200 ms to 300 ms. Asterisks represent significance (* $p <$ .05). Error bars represent ± 1 SE.

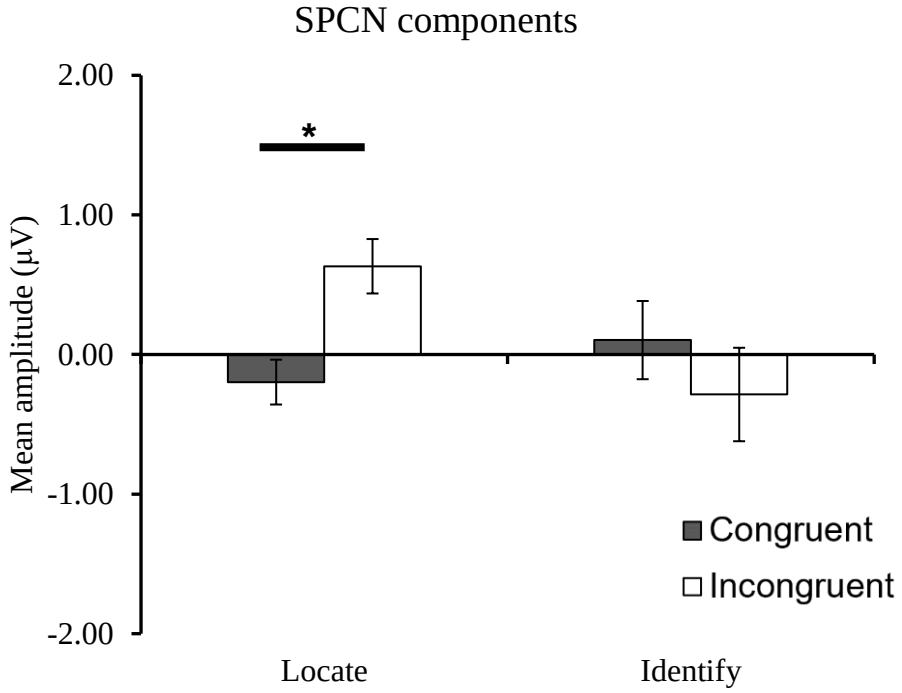
5.4.2. SPCN mean amplitude

The SPCN time window is marked from 400 to 600 ms (Figure 5.2, dashed rectangle). Figure 5.4 shows the SPCN mean amplitudes. A two-way ANOVA was conducted using the SPCN mean amplitude data, with trial type (congruent vs. incongruent) and task (locate vs. identify) as factors. There was an interaction between trial type and task, $F(1,19) = 5.44$, $p = .031$, $\eta_p^2 = 0.22$. To investigate this interaction, Bonferroni-corrected comparison (corrected alpha = .05/2 = .025) was conducted, with

comparisons between trial types for each task. For locate search, the amplitude was larger for congruent targets (M = -0.20 μV, SD = 1.05) than incongruent targets (M = 0.63 μV, SD = 1.62), $t(19) = -2.50$, $p = .017$, $d = -0.67$. For identify search, there was no significant difference in amplitude between congruent (M = 0.10 μV, SD = 1.11) and incongruent targets (M = -0.29 μV, SD = 1.06), $t(19) = 1.18$, $p = .17$, $d = 0.32$. There were no main effects of trial type nor task, $F < 1.48$. The amplitudes were not significantly different from zero, $t < 1.75$.

Figure 5.4

SPCN mean amplitudes (Experiment 6)



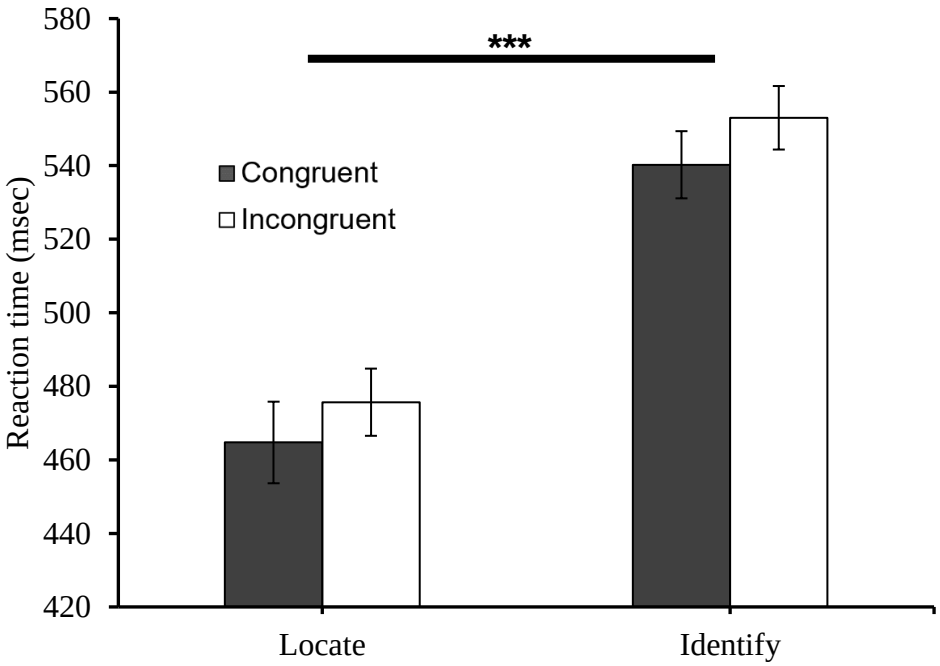
Note. Average of signals from 400 ms to 600 ms. Asterisks represent significance (* $p < .05$). Error bars represent ± 1 SE.

5.4.3. Reaction time

Figure 5.5 shows the average reaction time. A two-way ANOVA was conducted using average reaction time data, with trial type (congruent vs. incongruent) and search task (locate vs. identify) as factors. There was a main effect of trial type, with faster reaction time for congruent targets (M = 502.49 ms, SD = 45.24) than incongruent targets (M = 514.35 ms, SD = 41.97), $F(1,19) = 6.97$, $p = .016$, $\eta_p^2 = 0.27$. There was a main effect of task, with faster reaction time for locate task (M = 470.20 ms, SD = 47.47) than identify task (M = 546.64 ms, SD = 39.74), $F(1,19) = 107.24$, $p < .001$, $\eta_p^2 = 0.85$. There was no interaction between trial type and task, $F = 0.10$.

Figure 5.5

Average reaction time (Experiment 6)



Note. Asterisks represent significance (***$p < .001$). Error bars represent ± 1 SE.

5.5. General discussion

5.5.1. Summary and interpretation of findings

I examined the question of whether imagery impacts the post-selection stage of search. To address this question, I measured the N2pc and SPCN components (and behavioral measures) in two separate search tasks (locate vs. identify) with different recognition demands. Effects were assessed based on comparisons between congruent and incongruent trial types. Regarding the N2pc component, neural effects were observed irrespective of the search task, with larger N2pc for congruent than incongruent targets. These results aligned with the expectation that both tasks had high selection demands because attention needed to be allocated to the location of the target (Mazza et al., 2007). Regarding the SPCN component, neural effects were observed for the locate task but not the identify task. These results were unexpected, as the identify task required extracting features of the target whereas the locate task only required attentional allocation. Regarding the behavioral responses, reaction time was faster for congruent than incongruent targets irrespective of the search task. These results reflect typical behavioral effects observed in imagery-search studies (Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021).

Why did the task that had low recognition demands elicit SPCN effects but not the task that had high recognition demands? One possible explanation is that the task design influenced how participants engaged with the imagery. In the locate task, even though the primary demand was attentional allocation, participants might have relied more heavily on imagery to support their search, inadvertently engaging post-selection

processes that are typically associated with feature extraction (Eimer, 2014; Mazza et al., 2007). This could have led to the observed SPCN effects, reflecting a deeper level of processing than initially anticipated (Berggren & Eimer, 2018; Grubert et al., 2017). On the other hand, the identify task, which explicitly required feature extraction, might have led participants to depend more on perceptual rather than imagery-based strategies (Pearson, 2019). If imagery played a less central role in this task, it might explain the absence of SPCN effects.

Additionally, the faster reaction times for congruent targets across both tasks indicate that imagery facilitated search performance, further supporting its role as a top-down process (Cochrane, Townsend, Lee, et al., 2021; Cochrane, Wang, Pratt, et al., 2021). However, the differential neural effects suggest that the impact of imagery on the post-selection stage might be more complex and contingent on the specific demands of the task. These findings suggest that the role of imagery in post-selection processing is more flexible than previously thought depending on the nature of the task and the strategies employed by participants.

5.5.2. Theoretical contributions

The findings of this study offer insights into how imagery interacts with different stages of search. Traditionally, the N2pc and SPCN components have been employed to distinguish these stages (Eimer, 2014; Mazza et al., 2007), yet the specific role of imagery has been underexplored. By demonstrating that imagery can elicit distinct neural responses associated with both stages, this study underscores the potential of imagery as a valuable tool for investigating the temporal dynamics of attentional processing (Cochrane

et al., 2021). This contribution is significant as it opens up new avenues for exploring the interaction between imagery and attention in a more task-dependent manner.

The results support the goal-driven account by showing that imagery influences selection, as evidenced by consistent N2pc effects across both tasks. This suggests that a robust mechanism exists in which imagery impacts the selection process, reinforcing the importance of imagery in guiding attention. The unexpected SPCN effects observed in the locate task, despite its lower recognition demands, challenge existing assumptions about the relationship between task demands and post-selection processing (Mazza et al., 2007). These findings imply that imagery might engage post-selection processes even when feature extraction is not explicitly required, suggesting that imagery's role in post-selection is more flexible. The influence of imagery on attentional processes may be contingent on the nature of the task. The absence of SPCN effects in the identify task, despite its higher recognition demands, suggests that when a task permits alternative strategies, the role of imagery may be minimized. This task-dependent variability highlights the necessity for attentional models to account for the specific context in which imagery is employed.

5.5.3. Limitations

There were some limitations to consider. The length of the experimental run was larger compared to Studies 1 and 2. It is possible that participants were more fatigued in this experiment compared to others, which may have reduced compliance to imagery task (Cochrane et al., 2018; Monzel et al., 2021). Although efforts were made to balance the number of trials between the two tasks, individual differences in adherence to imagery

instructions may have introduced variability in the data. Future studies could mitigate this issue by implementing shorter experimental sessions.

Another limitation relates to the difference in recognition demands between the two tasks, which was modeled after seminal work distinguishing the N2pc and SPCN components (Mazza et al., 2007, 2009). While this distinction provided a useful framework for understanding attentional processes, it is based on the assumption that the tasks accurately captured recognition processes. The recognition demands in the current study might not fully align with the theoretical constructs of these ERP components. The locate task may have involved additional processes absent in the identify task, or the identify task could have had minimal recognition demands. Future research could refine the task designs to isolate the SPCN component in an imagery-search paradigm.

5.5.4. Conclusion

In conclusion, this study showed that imagery influenced both selection and post-selection processes in visual search, with congruent imagery enhancing the initial selection of targets and potentially influencing the sustained processing of selected items, depending on the task demands. The task-specific effects observed in the SPCN component suggest that task demands modulated the extent to which imagery continued to influence post-selection processes. These findings contribute to our understanding of how top-down imagery interacts with attentional mechanisms in visual search and underscore the importance of considering both selection and post-selection stages when investigating the effects of imagery on attention.

Chapter 6: Individual differences in vividness

6.1. Abstract

Individuals naturally vary in how vividly they experience mental imagery, yet the impact of these differences on attentional processing remains unclear. While previous studies have explored the neural correlates of subjective imagery using data-driven techniques, the role of vividness in attentional processing has not been adequately addressed. Study 4 investigates whether vividness modulated neural activity patterns during visual search and behavioral performance. Using the previously collected dataset, two analyses were conducted: multivariate pattern analysis (MVPA) on EEG data to examine the temporal dynamics of neural signals, and reaction time data to predict behaviorally relevant outcomes. The results revealed that neural signals during both the selection and post-selection stages of search were distinguishable based on vividness ratings, and that behavioral effects were more pronounced when imagery was rated as more vivid. These findings suggest that an individual's capacity for vivid imagery may set an upper limit on their performance in attention-demanding tasks.

6.2. Introduction

People's ability to experience imagery varies naturally, with some individuals vividly visualizing images while others may see less distinct images or nothing at all (Marks, 1973; Zeman et al., 2015). This variability has fueled the imagery debate regarding the representational aspects—whether imagery is depictive or propositional (Pearson & Keogh, 2015). Previous studies used subjective measures like the Vividness of Visual Imagery Questionnaire (VVIQ) to quantify individual differences in vividness

(Marks, 1973; M. Runge et al., 2015; M. S. Runge et al., 2017). FMRI studies showed correlations between VVIQ scores and neural activity in the visual cortex, with more vivid imagers showing stronger activation patterns in areas like V1 (Cui et al., 2007; Bergmann et al., 2016). These findings were further supported by MVPA techniques, which demonstrated significant overlap in neural activity between imagery and perception when vividness scores were high (Stokes et al., 2009; Thirion et al., 2006).

Previous studies provided insight into the neural correlates of vivid imagery, but they leave open the question of whether vividness affects other cognitive processes. For example, the sensory strength of imagery, as measured by the binocular rivalry task, has been shown to correlate with vividness ratings (Keogh & Pearson, 2017, 2018, 2021). In these studies, participants who imagined a color more vividly were more likely to perceive that color in subsequent perceptual tasks, suggesting a bias in perceptual awareness driven by imagery. Moreover, this sensory strength has been linked to performance in tasks involving visual working memory, suggesting that vividness may influence broader cognitive functions (Jacobs et al., 2018; Keogh et al., 2021). These findings suggest that vividness could play a modulating role in other processes, warranting further investigation.

The relationship between vividness and attentional processing remains underexplored. Most models of attention assume a uniform ability to generate visual representations, potentially overlooking the impact of vividness on attentional mechanisms (Wolfe, 2021). To address this gap, Study 4 examined whether trial-by-trial variations in vividness modulated the top-down effects of imagery on visual search.

Previous research using fMRI has employed MVPA to analyze the neural correlates of imagery, showing that vividness can influence neural activity patterns (Dijkstra et al., 2017). However, given that visual search involves temporally distinct stages (Eimer, 2014), EEG may be better suited for investigating how these stages are influenced by imagery. Recent adaptations of MVPA techniques for EEG data allowed for the detection of condition-specific neural activities across timepoints (Bae & Luck, 2018, 2019). Furthermore, previous imagery-search studies showed that the typical congruency effect (faster reaction time for congruent than incongruent targets) were more pronounced when vividness was high (Cochrane, Townsend, Lee, et al., 2021). Study 4 extends this line of research by examining these effects with a more varied set of stimuli and a larger dataset, aiming to provide a more comprehensive understanding of how vividness modulates attentional processing during search tasks.

6.3. Methods

6.3.1. Participants

Behavioral and EEG data from previous EEG experiments (excluding the active control group) were analyzed collectively. Based on the EEG data, I excluded those who had less than 25 trials per bin (highly vivid vs. less vivid). A total of 17 participants were excluded, leaving 83 participants in the final analyses.

6.3.2. MVPA technique with EEG data

The MVPA technique was described in Chapter 2. For each participant, support vector machines (SVMs) were trained to classify segments based on median-split vividness ratings (highly vivid vs. less vivid). The analysis was conducted separately for

each trial type. A classification accuracy measured the proportion of segments in which the SVMs correctly predicted the vividness category. For each trial type, a final classification accuracy was collapsed across participants.

6.4. Results

The goal was to test how well vividness predicted EEG and behavioral data. There were two predictions. First, from the EEG data, neural activity patterns during search could be distinguished based on vividness, reflected by significant classification accuracy. Second, from the behavioral data, reaction time would be different across individual vividness ratings.

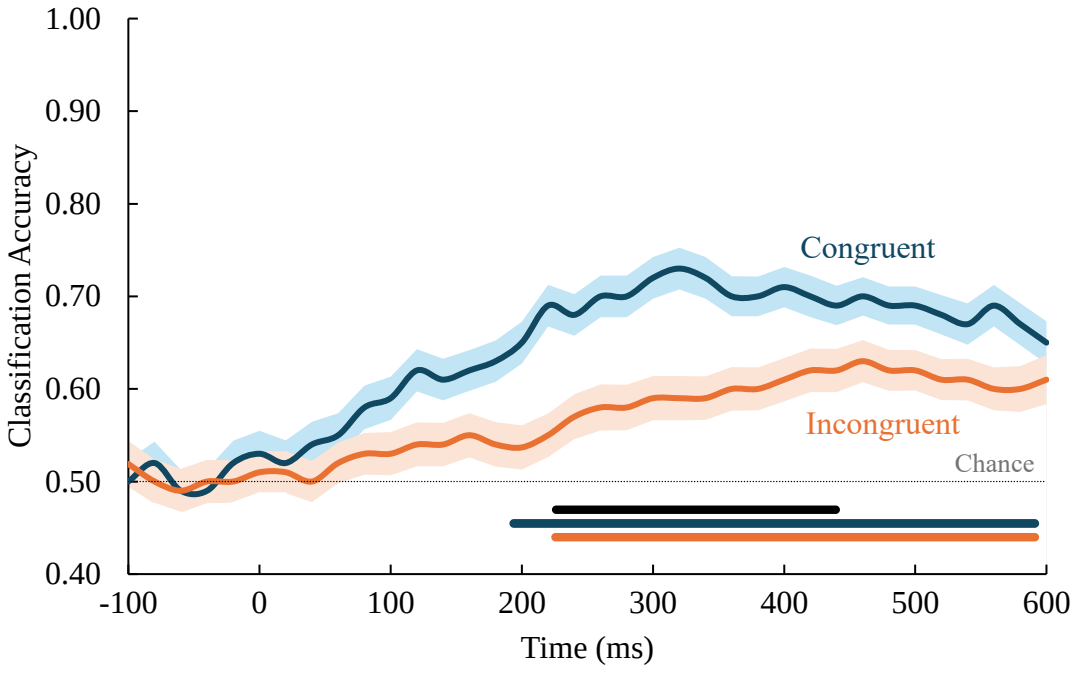
6.4.1. Classification accuracy

Figure 6.1 shows the classification accuracy. If vividness modulated activity patterns, then the classification accuracy should track differences in these patterns based on the vividness rating. First, each signal starting from 150 to 600 ms after stimulus onset was analyzed using a one-sample *t*-test compared to chance (Bae & Luck, 2018). The time window of analyses was chosen to include the approximate window in which selection (~200 to 300 ms) and post-selection (~400 to 600 ms) occurred. As there were only two vividness categories (highly vivid vs. less vivid), the chance performance was at 50%. Second, a cluster-level *t* mass was computed by summing the *t* scores from contiguous significant time points with correction (Bae & Luck, 2019). Given that several analyses were conducted (signals at each individual timepoint), Type I error needed to be controlled. To test the significance of the observed cluster-level *t* mass, permuted cluster-level *t* mass was generated from 10,000 iterations. The observed mass was considered

significant if it was larger than at least 95% of the permuted mass ($p < 10^{-4}$). For congruent trials, there was a significant cluster ($p < .0001$, one-tailed) between 198 and 600 ms (critical t mass = 95.22). For incongruent trials, there was a significant cluster ($p < .0001$, one-tailed) between 220 and 600 ms (critical t mass = 42.99). Finally, to test whether the SVM predicted vividness better for one trial type, signals from congruent and incongruent trials were compared at each time point, and cluster-level t mass was computed and compared to permuted mass like before. There was a significant cluster between 220 and 440 ms (critical t mass = 39.95), with higher accuracy for congruent trials (= 70%) than incongruent trials (= 59%).

Figure 6.1

Classification accuracy of vividness



Note. Sample-to-sample comparison of accuracy to chance (=0.50; Congruent, blue; Incongruent, orange) and between trial types (grey). Error bars (blue and orange shades) represent ± 1 SE.

6.4.2. Reaction time based on vividness

Figure 6.2 shows the average reaction time by vividness rating. To test whether vividness predicted behavioral effects, a linear mixed-effects (LME) analysis was conducted using average reaction time. Vividness ratings (1, 2, 3, and 4) and trial types (congruent and incongruent) were used as fixed effects and participants as the random effects intercepts. Likelihood ratio tests were used to compare the full model with vividness included and the reduced model without vividness included (e.g., Kay et al., 2022). There was a significant effect of vividness, χ^2 (3) = 20.30, $p < .001$, with faster reaction time following more vivid imagery, relative to less vivid imagery, for both trial types. Table 6.1 shows the fixed effects estimates.

Table 6.1

Fixed effects for LME of vividness data

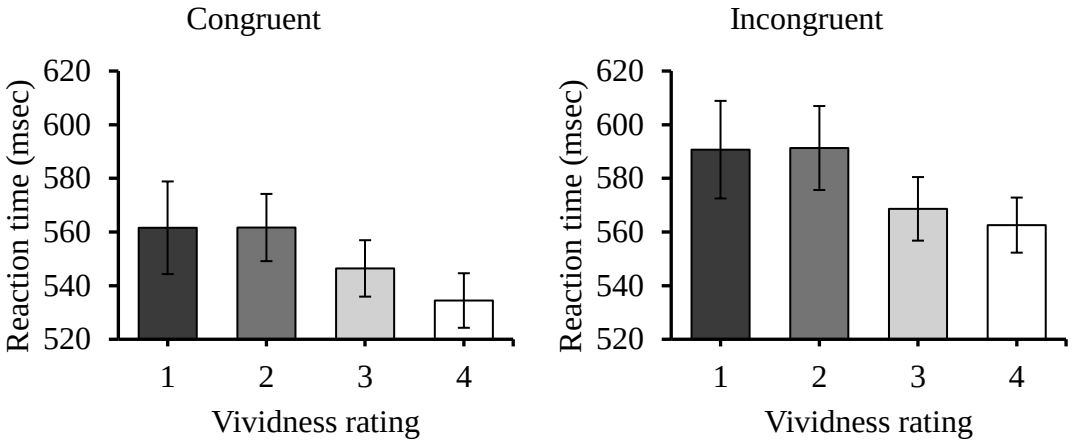
Fixed Effects Estimates

Term	Estimate	SE	df	t
Intercept	562.609	10.366	83.959	54.277
Vividness (1)	7.794	4.563	465.255	1.708
Vividness (2)	10.315	3.614	461.993	2.855
Vividness (3)	-4.170	3.480	462.064	-1.198
TrialType (1)	-12.766	2.070	460.367	-6.168

Note. Reaction time ~ Vividness Rating + Trial type + (1 | Subject).

Figure 6.2

Average reaction time by vividness rating



Note. Split between congruent and incongruent trials. Error bars represent ± 1 SE.

6.5. General discussion

6.5.1. Summary and interpretation of findings

I examined the question of whether individual differences in vividness ratings influenced neural activity patterns and behavioral responses. To address the question, two exploratory analyses were conducted. To examine the temporal dynamics of neural signals, an MVPA technique with EEG was used based on median split vividness ratings. For both trial types, signals were decoded during the selection stage (~200 to 300 ms). This finding was unsurprising, as imagery was predicted to impact signals associated with selection. Related studies that showed between-condition N2pc effects used similar MVPA techniques (Bayet et al., 2018; Moon, Zhao, et al., 2022). For example, Bayet et al. (2018) decoded signals during the selection stage between trials with a specific target (e.g., a specific human face) and trials with a category target (a category of human faces).

Moon et al. (2022) decoded signals in the same time window based on the probable target location (e.g., left or right of cued location). In my own unpublished work, I also decoded signals in the same time window based on target present trials and target absent trials with distractors. Collectively, these results showed that between-condition N2pc effects were aligned with high classification accuracy during the time that selection was expected. Here, the findings suggest that vividness modulated these selection-driven signals.

Signals also were decoded during the post-selection stage. There are two potential explanations. First, it is possible that modulating effects originating during the selection stage persisted beyond the typical N2pc window, extending into the post-selection phase (~400 to 600 ms). This explanation aligns with the previous results showing differences in N2pc onset between trial types. Perhaps vividness modulated selection-driven signals with faster selection for highly vivid trials compared to less vivid trials. However, classification accuracy only reflects the distinguishability of neural signals based on vividness, not the direction of effects between vividness categories. I discuss later behavioral findings to support faster modulating effects of vividness. Second, the findings suggest that vividness modulated signals related to post-selection processes. The inclusion of trials from Experiment 6, which demonstrated significant SPCN effects, may have contributed to the results. Alternatively, the results suggest that vivid imagery enhanced not only the selection of relevant stimuli but also the maintenance and processing of these stimuli during the post-selection stage. Interestingly, the previously

mentioned search studies that deployed MVPA techniques also revealed similar patterns of results, with tasks that did not have recognition demands (e.g., Bayet et al., 2018).

The accuracy was higher for congruent trials than incongruent trials during and after the N2pc time window. This difference is unsurprising because the N2pc component was larger for congruent targets than incongruent targets. The higher accuracy suggests that vivid imagery enhanced target selection, leading to more pronounced and distinguishable neural signals compared to less vivid imagery. Additionally, the difference in accuracy was absent during the post-selection stage (after 400 ms), suggesting that vividness modulated signals to the same extent for both congruent and incongruent trials after the initial selection process. These findings reflect a temporal specificity in how vividness modulates attentional processing. The fact that the difference in accuracy was concentrated around the N2pc window suggests that vividness exerts its strongest influence in early stages of attentional processing. If true, then vivid imagery is particularly effective at enhancing target selection.

To examine whether vividness predicted behavioral responses, an LME model was applied to reaction time split across individual vividness ratings. The findings revealed faster reaction times following more vivid imagery compared to less vivid imagery for both trial types. For congruent trials, these results align with previous imagery-search studies that demonstrated faster reaction times for vividly imagined targets (Cochrane, Townsend, Lee, et al., 2021). Related studies showed that highly vivid imagery biased upcoming perceptual events more strongly than less vivid imagery (Keogh & Pearson, 2017, 2018, 2021; Pearson et al., 2011). An unexpected result was in

incongruent trials, which showed a similar pattern of faster reaction times. It was hypothesized that more vivid imagery would lead to greater distraction from matching nontargets, resulting in slower reaction times for incongruent trials (Cochrane, Ng, Khosla, et al., 2021; Cochrane, Wang, Pratt, et al., 2021). In contrast, the data showed faster reaction times for more vivid incongruent trials. Perhaps vivid imagery may enhance overall perceptual processing speed in the presence of distractors. It is possible that vivid imagery facilitated a more efficient processing of incongruent targets by reducing the impact of irrelevant distractors that matched imagery, thereby speeding up the overall reaction time. Further research is needed to explore these results.

6.5.2. Theoretical contributions

An important contribution of this study is its examination of how vividness influences attentional processing. The findings suggest that vividness might impose an upper limit on attentional capacity, such that imagery guides attention to the extent of its vividness. Previous research has demonstrated that individuals with high vividness scores exhibit stronger neural activation patterns in the visual cortex (Cui et al., 2007; Bergmann et al., 2016) and greater neural overlap between imagery and perception (Albers et al., 2013; Stokes et al., 2009; Thirion et al., 2006). This study supports the notion that vivid imagery engages perception networks more robustly than less vivid imagery, potentially enhancing the precision of visual representations. Our results align with this interpretation by showing that vivid imagery influences both neural activity and behavioral performance in visual search tasks. This suggests that the effectiveness of imagery in guiding attention may be constrained by its vividness, reinforcing the idea that

attentional processes are intricately linked to the perceptual quality of mental images, as described by the depictive theory (Kosslyn, 1975, 1981, 2005).

While previous studies have explored the neural correlates of vivid imagery (Dijkstra et al., 2019), the present study advances this by examining the temporal dynamics of vivid imagery during search tasks. Although individual differences in vividness have been investigated (M. Runge et al., 2015; M. S. Runge et al., 2017), there has been limited research on how vividness specifically relates to attentional processing. The current findings highlight the need to update existing attentional models, which often assume a uniform ability to generate visual representations. This study underscores the importance of incorporating individual differences in vividness into attentional frameworks, emphasizing that attentional mechanisms are not only influenced by imagery but also by its qualitative aspects (Wolfe, 2021). By doing so, future research can more accurately reflect the complex interplay between imagery and attention, leading to a deeper understanding of how vividness impacts cognitive processes.

As a methodological contribution, this study addresses the question of whether EEG signals can be decoded based on subjective experiences such as vividness ratings. While previous research has successfully decoded EEG signals based on experimental conditions, such as the specificity of the target (Bayet et al., 2018) or the location of the target (Moon et al., 2022), the application of decoding techniques based on individual subjective experiences had not been thoroughly explored in search studies. This approach extends the utility of EEG decoding beyond objective experimental manipulations to

include subjective states, providing a novel methodological framework for investigating the relationship between subjective experiences and neural activity.

6.5.3. Limitations

One limitation of this study is the variability in experimental paradigms across studies, which complicates direct comparisons and interpretations of the results. Study 1 utilized real-world items, which were visually more complex compared to Study 2's letters (Q and O) and Study 3's colored shapes (red and green squares). This difference in stimulus complexity could have affected attentional demands with increased task difficulty. Additionally, the number of search items varied between studies: Study 1 included three complex items, Study 2 had either three or six letters, and Study 3 involved four shapes. The response instructions also differed; Studies 1 and 3 required responses to four separate positions, while Study 2 instructed responses to left or right positions. In Study 2, trials without pop-out effects (e.g., O among Qs) were part of the congruent trial type, and participants were instructed to imagine the opposite stimulus. The inclusion of trials without pop-out effects and the instruction to imagine the opposite stimulus in Study 2 could have altered participants' engagement with the imagery task.

Another limitation is that the MVPA techniques used in this study cannot establish causal relationships. Although the findings suggest that vividness modulated EEG signals during attentional processing, it cannot be concluded that these signals were directly driven by vividness. While MVPA can identify patterns (Bae & Luck, 2018), it cannot confirm whether these patterns are the result of vividness alone. Inevitably, the interpretation remains correlational. Additionally, distinguishing between imagery-driven

signals and attention-driven signals is inherently challenging because both occurred during the search task. The current analysis did not differentiate whether these signals are primarily driven by imagery or by the attentional processes. If the signals are largely attention-driven, then the effects attributed to vividness might instead reflect changes in attentional focus. Clarifying this distinction would require more targeted experimental designs that isolate the specific contributions of imagery and attention.

6.5.4. Conclusion

In conclusion, this study enhanced current understanding of how vividness influences attentional processing by demonstrating that vivid imagery can modulate both neural activity and behavioral responses. These findings align with previous research showing stronger neural activation with higher vividness (Cui et al., 2007; Bergmann et al., 2016). The results contribute to a more nuanced understanding of how vivid imagery can pose upper limits to attentional mechanisms.

Chapter 7: Conclusion

7.1. Overview

Mental imagery is not merely a passive ability but an important strategy for tasks that demand attention. Decades of imagery studies examined the nature of imagery, revealing both conceptual and neuroanatomical similarities to perception (Pearson, 2019). Similarly, decades of attention studies examined the functional dynamics of attentional processing during visual search tasks (Wolfe, 1994; Duncan & Humphreys, 1989). More recently, studies have started to examine the impact of imagery on attentional processing, linking these two fields of research (Cochrane et al., 2021; Moriya, 2018). Anecdotally, many people would report using imagery as a strategy when preparing for a search task by imagining the target beforehand. Empirical studies also showed that imagery can influence behavior, such as improving reaction times (Cochrane et al., 2021; Moriya, 2018). Given that imagery has clear effects on behavior, one may assume that it impacts processes occurring before the overt behavior, particularly attentional mechanisms. However, the findings to date have been underwhelming, leaving room for further exploration into the interaction between imagery and attention.

The four studies in this dissertation provide evidence that imagery can impact attentional processing. I implemented a visual search paradigm to examine both attentional processing and subsequent behavioral responses following imagery instructions. The findings consistently revealed that imagery elicited early effects on attentional processing, during the time when selection was expected to occur. This dissertation pushes forward the field of attention research by explicitly investigating the

strategy of mental imagery, as opposed to assuming its use implicitly. First, I review the main findings of each study and assess whether they addressed the research questions posed. Next, I discuss the overall contributions of these findings. Then, I explore the broader implications of this work. Finally, I address the general limitations of this work that warrant further research.

7.2. Main findings

As a general hypothesis, I tested whether imagery elicited top-down effects on attentional processing. This hypothesis was addressed across three experimental studies (Studies 1, 2, and 3) and one analytical study (Study 4). Each study addressed a specific research question stated in each subsection. Table 7.1 summarizes the main findings of each experiment.

7.2.1. Did imagery elicit top-down effects?

Study 1 tested the hypothesis that imagery guides attention to imagery-matching items. To address this hypothesis, both the N2pc component and reaction time were utilized as relevant measures. Previous studies demonstrating behavioral effects of imagery alone were insufficient to determine whether imagery specifically impacts attentional selection (Cochrane et al., 2018; Cochrane, Ng, et al., 2021; Cochrane, Townsend, et al., 2021; Cochrane, Wang, et al., 2021; Monzel et al., 2021; Moriya, 2018). The N2pc component provided a means to capture top-down effects (Eimer, 1996; Luck, 2011). Across the three experiments, the study replicated previously observed behavioral effects of imagery (Cochrane et al., 2018; Cochrane, Townsend, et al., 2021; Moriya, 2018). Experiment 1 demonstrated that the N2pc component was larger and faster for

congruent compared to incongruent trials, aligning with the hypothesis that imagery impacts selection. Experiment 2 replicated these N2pc effects and minimized these effects when removing imagery instructions, though behavioral effects persisted. This suggests the presence of other influences, such as intertrial priming (Maljkovic & Nakayama, 1994, 1996). Experiment 3 again observed N2pc effects between congruent and incongruent trials and additionally showed that N2pc onset was faster for congruent than neutral trials and slower for incongruent than neutral trials. These results collectively support the goal-driven account of attention, demonstrating that imagery functions as a top-down process influencing target selection.

7.2.2. Did imagery modulate pop-out effects?

Study 2 tested the hypothesis that imagery would modulate pop-out effects. To test this hypothesis, I used asymmetric search paradigms by manipulating the presence of salient features on the target. The target and nontargets differed by one salient feature (slanted line on Q), presented only either on the target (Q among Os) or on the nontargets (O among Qs). This manipulation allowed for comparing top-down effects when pop-out effects were present (Q among Os) and absent (O among Qs). According to the goal-driven account, imagery should modulate pop-out effects based on behavioral and neural differences between congruent and incongruent trial types, similar to what was observed in Study 1. According to the stimulus-driven account, pop-out effects should not be influenced by the trial types (Treisman & Gormican, 1988; Wolfe, 1994).

Table 7.1

Summary of findings

Study	Exp	Conditions	DV	Main findings
1	1	Trial types: Congruent, Incongruent	N2pc, RT	• N2pc was larger and faster in congruent than incongruent. • RT was faster in congruent than incongruent
	2	Trial types: Congruent, Incongruent Group: Imagery, Active Control	N2pc, RT	• For Imagery group, N2pc was larger in congruent than incongruent • For both groups, RT was faster in congruent than incongruent
	3	Trial types: Congruent, Incongruent, Neutral	N2pc, RT	• N2pc was faster in congruent than neutral, and slower in incongruent than neutral • RT was faster in congruent than incongruent, slower in incongruent than neutral
2	4	Trial types: Congruent, Incongruent Salient feature: Present, Absent Set size: 3 items, 6 items	RT, Slope RT	• For present feature, RT was faster in congruent than incongruent (opposite for absent feature) • For both trial types, slope RT was steeper when feature was absent than present
	5	Trial types: Congruent, Incongruent Salient feature: Present, Absent Set size: 3 items, 6 items	N2pc, RT, Slope RT	• For present feature, N2pc was larger in congruent than incongruent (faster for both features) • Behavioral effects from Exp 4 were replicated
3	6	Trial types: Congruent, Incongruent Tasks: Locate, Identify	N2pc, SPCN, RT	• For both tasks, N2pc was larger for congruent than incongruent • For locate task, SPCN was larger for congruent than incongruent • RT was faster for congruent than incongruent
4	Analytical	Trial types: Congruent, Incongruent Vividness: Median split (MVPA), individual ratings (LME)	Classification ACC, RT	• Signals were decoded during N2pc and SPCN window • Between trial types, signals were decoded during N2pc window • For both trial types, higher vividness predicted faster RT

Note. DV, Dependent Variable; RT, reaction time.

Experiment 4 tested the hypothesis using average reaction time and slope reaction time, as traditionally used in asymmetric search (Dosher et al., 2004; Saiki et al., 2005; Shen & Reingold, 2001; Treisman & Gormican, 1988). Experiment 5 replicated the behavioral findings and additionally tested the N2pc effects as done in Study 1. Three major findings emerged. First, average reaction time was faster for congruent targets than incongruent targets when the salient feature was present (Q among Os), as predicted by the goal-driven account. However, the reverse pattern was observed when the salient feature was absent (O among Qs). Reaction times for congruent targets were slower when the salient feature was absent, suggesting that the absence of the feature might reduce the effectiveness of imagery in guiding attention. Second, search efficiency was higher (i.e., flatter slope) when the salient feature was present, regardless of trial type, consistent with previous asymmetric search studies (Dosher et al., 2004; Saiki et al., 2005; Shen & Reingold, 2001; Treisman & Gormican, 1988). These results align with the stimulus-driven account, suggesting that efficiency is driven by saliency and not top-down processes (Wolfe, 1994).

Third, the N2pc effects were observed between trial types only when the salient feature was present, partially aligning with the average reaction time results. The N2pc onset was faster for congruent targets compared to incongruent targets, regardless of the feature's presence. The N2pc component was larger and faster for congruent targets when the salient feature was present, supporting the goal-driven account of imagery influencing attention (Cochrane et al., 2018; 2021). When the feature was absent, the N2pc onset was faster for congruent targets, but the magnitude did not differ. This lack of difference could

indicate that imagery's influence on attention is limited without the presence of a salient feature. Alternatively, the absence might result in a serial search process (Saiki et al., 2005), which could be less affected by top-down imagery. Putting it together, the interaction between imagery and pop-out was more nuanced than either the stimulus-driven or goal-driven accounts alone can fully explain. The top-down effects from Study 1 were replicated, suggesting that pop-out effects *can* be modulated by imagery in some ways.

7.2.3. Did imagery impact post-selection?

Study 3 tested the hypothesis that imagery influences the selection and potentially the post-selection stages of search. Previous imagery-search studies that demonstrated behavioral effects typically involved tasks requiring both selection and post-selection processes (e.g., Cochrane et al., 2021; Pearson et al., 2015), raising the question of how these effects arise from attentional processing. To test the hypothesis, I measured the N2pc and SPCN components in two distinct search tasks: 'locate' and 'identify', which varied in their recognition demands. The 'locate' task required participants to respond to the location of the target, while the 'identify' search task required participants to respond to a specific feature of the target. Although both tasks involved selection, only the identify task necessitated recognition. The N2pc and SPCN components served as neural markers to differentiate between the effects on selection and post-selection stages of attention (Eimer, 2014; Mazza et al., 2007).

N2pc effects were observed regardless of the search task, with a larger N2pc for congruent targets compared to incongruent targets. These results were expected, as both

tasks had high selection demands, requiring attention to be allocated to the target's location (Mazza et al., 2007). SPCN effects were observed in the locate task but not in the identify task. This finding was unexpected, as the identify task, which required extracting the target's features, was presumed to engage post-selection processes more intensively than the locate task, which primarily involved attentional allocation (Mazza et al., 2007). Putting it together, the findings suggest that imagery influenced both selection and post-selection, but the extent of this influence depends on the task demands.

7.2.4. Did vividness modulate top-down effects?

Study 4 tested the hypothesis that vividness modulated the top-down effects of imagery. Several studies used subjective measures of vividness (M. Runge et al., 2015; M. S. Runge et al., 2017) to examine neural correlates of vivid imagery (Cui et al., 2007; Dijkstra et al., 2017). Much less is known about whether vividness of imagery imposes limitation on attentional processing. Current models of attention often assume a uniform ability to generate visual representations (Wolfe, 2021).

To test this hypothesis, two exploratory analyses were conducted. First, to examine the temporal dynamics of neural signals, an MVPA technique with EEG (Bae & Luck, 2018, 2019) was used, based on a median split of vividness ratings. If imagery-related activities differed based on vividness, then MVPA techniques could potentially detect these differences during search. Signals were decoded during the selection stage (~200 to 300 ms) for both trial types, aligning with predictions that imagery would impact signals associated with selection. This finding is consistent with related studies that observed between-condition N2pc effects using similar MVPA techniques (Bayet et

al., 2018; Moon et al., 2022). Signals were also decoded during the post-selection stage (~400 to 600 ms). This could indicate that modulating effects originating during the selection stage persisted into the post-selection phase, aligning with previous results showing differences in N2pc onset between trial types. Alternatively, the findings might suggest that vividness directly modulates signals related to post-selection processes, as observed in Experiment 6. When comparing classification accuracy between trial types, accuracy was higher for congruent trials than for incongruent trials around the N2pc time window, aligned with the N2pc effects observed in previous experiments.

Second, to examine whether vividness predicted behavioral responses, an LME model was applied to reaction times split across individual vividness ratings. Previous imagery-search studies found larger congruent-incongruent differences when vividness ratings were high (Cochrane, Ng, Khosla, et al., 2021). In this study, the results revealed faster reaction times following more vivid imagery compared to less vivid imagery for both trial types. For congruent trials, the results aligns with previous findings, where vividly imagined targets led to faster reaction times (Cochrane, Townsend, Lee, et al., 2021). Unexpectedly, the same pattern of results were observed in the incongruent trials, contrary to previous studies that showed slower reaction times (Cochrane, Ng, Khosla, et al., 2021; Cochrane, Townsend, Lee, et al., 2021). It is possible that vivid imagery facilitated more efficient processing of incongruent targets by reducing the impact of irrelevant distractors that matched the imagery, consequently speeding up reaction time.

Putting it together, the findings showed that vividness modulated top-down effects in both neural and behavioral domains. Perhaps vivid imagery increased the selection of

target features, ultimately impacting behavior. The unexpected facilitation observed in incongruent trials suggests that vivid imagery may also reduce the impact of matching features in nontargets.

7.3. Contributions

7.3.1. Informing the debate between stimulus-driven and goal-driven account

The current work provided evidence in favor of the goal-driven account and against a strict stimulus-driven account. The debate concerned whether attentional selection could be influenced by processes besides bottom-up processing (Eimer & Kiss, 2010). According to the goal-driven account, internally generated representations called templates (relation to imagery discussed in a later section) are activated to guide attention to matching features (Olivers et al., 2011). According to the stimulus-driven account, if salient features are present, attention would be prioritized by saliency (i.e., pop-out effect) over the guidance of templates (Theeuwes, 2010). To address this debate, the N2pc component previously was measured in cue-search tasks with the following predictions. If the stimulus-driven account is correct, then the N2pc should mark the selection of salient features regardless if they matched the template (i.e., the target) or were task-irrelevant (nontarget). If the goal-driven account is correct, then the N2pc should mark the selection of salient features only when matched but not when irrelevant. Previous studies showed results as predicted by the goal-driven account (Eimer & Kiss, 2008, 2010).

The findings here aligned with the N2pc results of the goal-driven account. N2pc effects were observed consistently with respect to imagery, revealing two points. First,

imagery functions as a top-down process that influences selection (Study 1). Collectively, the findings resembled those that supported the goal-driven account. Second, imagery modulates bottom-up processing of pop-out effects (Study 2). When there were pop-out effects (Q among O), N2pc effects occurred with respect to matching features of the imagery. When there were no pop-out effects (O among Qs), N2pc effects remained present following imagery. These findings do not oppose the notion that bottom-up processing was at work. Rather, this process was under the influence of top-down processes like imagery, contrary to the stimulus-driven account.

A counterargument from the stimulus-driven side was that the N2pc effects reflect the late stages of processing (Hickey et al., 2006). Study 3 tested this claim by measuring the SPCN component to mark post-selection. If imagery instead influenced subsequent stages of selection, then SPCN effects should have been revealed following imagery. However, only one of two tasks with identical imagery instructions showed this prediction. Moreover, the N2pc effects were replicated in both tasks and observed earlier than the SPCN, revealing a temporal distinction between two components. This temporal ordering challenges the stimulus-driven account, demonstrating that top-down processes like imagery can influence attentional selection rather than merely later stages of processing.

7.3.2. Informing the imagery debate

The findings support the depictive theory that imagery can activate visual representations (Kosslyn, 1975, 1981, 2005; Kosslyn et al., 2006). This theory proposed that imagery engages in perceptual networks to activate visual representations. If true,

then representations of imagery should impact behavior similarly to those of perception (Cochrane & Milliken, 2019, 2020; Wantz et al., 2015). Previous imagery-search studies showed comparable behavioral search outcomes following either an imagery cue or perceptual cue (Cochrane et al., 2021). Study 1 extended these findings with N2pc effects, suggesting that visual representations were used to guide attention to matching items. Study 2 shared similar results while eliminating intertrial priming effects by instructing imagery of the opposite letter.

If the propositional theory was true, then visual representations should have been absent (Z. W. Pylyshyn, 1973, 1973, 1981). The findings here suggest otherwise. First, N2pc effects were used to support the goal-driven account (Eimer & Kiss, 2008), which modeled visual representations as the basis for enhancing selection (Desimone & Duncan, 1995; Duncan & Humphreys, 1989). Studies 1, 2, and 3 support the goal-driven account, thus no N2pc effects should have occurred in the absence of visual representations (e.g., Experiment 2). Second, if imagery did not activate visual representations, then measures of imagery pertaining to its visual qualities (e.g., vividness) should not predict effects of imagery. In contrast to the propositional theory, previous fMRI studies decoded activity patterns in the visual cortex based on vividness (Dijkstra et al., 2017). Similarly, Study 4 decoded activity patterns during attentional processing based on median split vividness, and behavioral outcomes differed based on individual vividness ratings. Collectively, the results were in line with the depictive theory, suggesting that visual representations were activated in search tasks and functionally relevant to attentional processing.

7.3.3. Building on the Reverse Hierarchy Model

The current work builds on the Reverse Hierarchy Model (Pearson, 2019). The model described imagery as a top-down process, though not specific for attention. Briefly, imagery triggers activity patterns in frontal regions (Ishai et al., 2000; Ragni et al., 2021) that pass through the medial temporal lobe for memory retrieval (Bird et al., 2010; Ranganath & D'Esposito, 2005). These activities cascade across the visual region to activate visual representations (Pearson & Kosslyn, 2015), in line with the depictive theory. The depictive theory describes how imagery is represented, and this model described how its representation come about. The model describes how top-down imagery functions in a given imagery task, whereas my work describes how top-down imagery functions following an imagery task. The N2pc effects revealed that these representations enhanced selection of matching items (Study 1) and modulated bottom-up processing (Study 2). Separately, the SPCN effects revealed that these representations may impact post-selection depending on task demands (Study 3). All in all, the findings build upon the Reverse Hierarchy Model by informing the functional role of imagery.

7.3.4. An upper limit to attentional processing

Study 4 revealed differences in the EEG signals between highly vivid and less vivid trials, suggesting a potential upper limit to attentional processing influenced by the vividness of imagery. Specifically, the extent to which internally represented targets enhance target selection appears to depend on the subjective experience of these representations. This finding is crucial for several reasons. First, it aligns with observations from fMRI studies, where vivid imagery elicited activity patterns similar to

those observed in perception (Cui et al., 2007; Bergmann et al., 2016). The analogy proposed by the depictive theory—where imagery is likened to viewing a picture (see Figure 1.1)—implies that the visual quality of imagery can influence the effectiveness of attentional processes. Consequently, vivid mental images could potentially set an upper limit on the effectiveness of attentional processes by enhancing the precision of attentional focus. Second, an upper limit raises the question of whether there is an inherent limitation to one's vividness that impacts attentional skills, or if vividness can be trained to improve these skills. Identifying this upper limit highlights potential constraints of attentional resources and their modulation by imagery (Cochrane et al., 2021). Third, current theoretical models of attention do not account for the subjective qualities of internal representations (Wolfe, 2019). By demonstrating that vividness can modulate both activity patterns and behavioral outcomes, this research prompts a reevaluation of attentional models to incorporate these subjective factors.

7.3.5. Methodological advancements

There were three methodological advancements. First, although EEG has been widely used in search studies to mark specific attentional stages (Eimer, 2014), this technique rarely has been used in imagery-search studies (Cochrane et al., 2021). The significance of this advancement lies in the ability of EEG to provide real-time insights into how imagery influences attentional processes. Traditionally, EEG studies have focused on specific attentional stages or components, such as the N2pc and SPCN, without integrating the effects of mental imagery. By applying EEG to imagery-search paradigms, these studies bridge a gap in the literature, allowing for a more direct way to

examine the effects of imagery on attentional mechanisms in a temporally resolved manner (Mazza et al., 2007; Eimer, 2014).

Second, Study 4 introduced a new application of EEG decoding techniques. This was significant in understanding the temporal dynamics of imagery in tasks that demand attention, an understudied area. Findings showed that meaningful differences based on vividness could be extracted from scalp EEG data. This method was significant because it allowed for the examination of how individual differences in imagery vividness influence neural activity. Prior research primarily utilized EEG decoding to differentiate between experimental conditions, such as target specificity or location (Bayet et al., 2018; Moon et al., 2022). By incorporating subjective ratings of imagery vividness, this approach extended the utility of EEG decoding, revealing meaningful differences in neural responses related to individual differences in imagery quality.

Third, the search paradigm across studies showed robust N2pc effects, a critical component in understanding top-down attentional processes. By using a search task to measure N2pc effects, the studies established a reliable method for assessing how imagery impacts visual search performance and attentional mechanisms. This approach also allowed for the evaluation of different strategies, such as comparing the effectiveness of imagery versus other cognitive strategies in guiding attention (Keogh et al., 2021; Jacobs et al., 2018). Furthermore, the ability to establish a task-specific measure for top-down effects provided insights into how variations in imagery strategies might influence search efficiency, especially in individuals with extremely poor imagery capabilities (e.g., aphantasia, Keogh & Pearson, 2014).

7.4. Implications outside current work

7.4.1. Are imagery and templates one and the same? The next debate to come

Models of the goal-driven account state that templates of the targets are formed by activating selective visual representations (Desimone & Duncan, 1995). According to the depictive theory, imagery also activates visual representations (Pearson et al., 2015). Given that both templates and imagery rely on visual representations, the question arises: are templates and imagery fundamentally the same, or do they operate differently? The findings here and from other imagery-search studies (Cochrane et al., 2018; Moriya, 2018; 2023) were similar to those from studies under the template framework (Conci et al., 2019; Müller et al., 2003; Töllner et al., 2010; Vickery et al., 2005), suggesting templates and imagery share a functional role in guiding attention. Similarly, Keogh & Pearson (2021) implemented the binocular rivalry task to compare the sensory strength of imagery (perceptual bias to previous mental image, Pearson et al., 2008) with the sensory strength of templates (perceptual bias to previously attended feature). Participants were primed to report a particular color following either imagery of a certain color or attention to the same color. Findings revealed similar sensory strengths between the two conditions. Moreover, individuals self-described as aphantasic (inability to imagine; more on this in a later section) displayed no sensory strength in either condition. In summary, the 'potency' of imagery and templates on perceptual awareness was similar, suggesting that imagery engages with underlying mechanisms that prepare templates. Aphantasia may suggest an impaired mechanism shared between imagery and templates. In line with this, Monzel et al. (2021) implemented an imagery-search paradigm, sampling from both

aphantasic and non-aphantasic populations. Reaction time was slower for the aphantasic group than a control group, as expected if imagery and templates share mechanisms.

However, some factors must be considered before concluding that templates and imagery are similar. First, the previously mentioned studies compared intact imagery with no imagery (Keogh & Pearson, 2021; Monzel et al., 2021). However, as Study 4 showed, variation in imagery skills, as opposed to present or absent skills, implicate meaningful differences in search outcomes that need to be accounted for in models under the template framework (Wolfe, 2019; Eimer, 2014). Second, there are nuances to how templates and imagery are evaluated. In cue-search tasks that involve templates, observers are typically informed about the specific features of the target in advance (Müller et al., 2003). In this sense, templates are 'fixed', as the target is predefined. In imagery-search tasks like here, observers often are unaware what to search for until the stimulus are presented (Cochrane et al., 2021). Trial types were evenly split and randomized across trials, so observers cannot form expectations of the upcoming target. In this sense, imagery is 'variable' because it relies on the observer's ability to flexibly generate a mental image. These differences challenge the notion that templates and imagery are fundamentally the same because they operate under different conditions.

To address this issue, Cabbai et al. (2023) implemented a search paradigm in which participants were cued the upcoming target with diagnostic features (e.g., lemon is associated with the color yellow). This cue served to generate templates. Before the search array was presented, a separate cue display was presented, in which one of the locations was presented with the diagnostic feature (yellow). Search was faster when the

target appeared in the same location as the diagnostic feature than a different location – a search phenomenon known as the contingent capture effect (Folk et al., 1992). If imagery and templates are similar, then explicit instructions to imagine the target object should impact the contingent capture effect. That is, trials with more vivid imagery should result in stronger effects compared to less vivid imagery. Moreover, individuals with aphantasia should exhibit less impact on contingent capture effect as imagery is unavailable. However, findings revealed no impact of imagery instructions either in the control or aphantasic groups, suggesting that the preparation of templates does not operate under the same mechanisms of imagery.

Putting it together, the evidence for the similarity between imagery and templates remains mixed (Cabbai et al., 2023; Keogh & Pearson, 2021; Monzel et al., 2021). My findings suggest similarities in their function, but the underlying mechanisms shared by the two processes remain unknown. This conflict in recent studies hints at an emerging debate that could potentially bridge the imagery and attention fields.

7.4.2. Target enhancement vs. distractor suppression

Study 2 revealed an interaction between top-down effects of imagery and pop-out effects. A peculiar finding was that when pop-out effects were absent (O among Qs), reaction time was faster for incongruent targets (imagined Q, then searched for O) than for congruent targets (imagined O, then searched for O). This reverse effect contrasted the traditional effect of imagery, with faster reaction time for congruent than incongruent targets (Cochrane et al., 2021a). These results raise questions about what exactly the top-down effects of imagery are. In a congruent trial, the effect was suggested to be that

imagery guided attention to target-matching features. In an incongruent trial, imagery still would guide attention but to matching features on the nontargets. That is, imagery influenced target enhancement, evinced by the N2pc effects (Mazza, 2009). However, this explanation falls short with the O among Qs case, suggesting that imagery elicits other types of effects. One possibility is that, in addition to target enhancement, imagery influences distractor suppression. That is, once attention was allocated to matching features on the nontargets, it was deployed elsewhere more effectively than to nonmatching features (Sawaki et al., 2012). Recent ERP studies used the P_d component (distractor positivity), which was observed before the N2pc component (Sawaki et al., 2012). Future studies should consider optimizing the search paradigm to test effects on both target enhancement and distractor suppression.

Additionally, Study 4 showed that in incongruent trials, reaction time was faster in trials with higher vivid ratings compared to less vivid ratings. This finding contradicts the prediction that vivid imagery would lead to more distraction from matching nontargets (Cochrane et al., 2021). If vivid imagery increases distractor suppression, then it would align with the observation that more vivid imagery resulted in more pronounced behavioral effects in the incongruent trials. This interpretation further underscores the importance of distinguishing between target enhancement and distractor suppression in attentional models, suggesting that vivid imagery can modulate both processes.

7.4.3. Causal role of imagery: a case for studying aphantasia

A causal claim could not be made with the current set-up. However, given that the used paradigm successfully elicited N2pc effects, perhaps the same tasks can be deployed

to assess those who cannot use imagery, a condition known as aphantasia. In 2015, the term aphantasia was coined to describe the inability to use imagery (Zeman et al., 2015). While there have been references to this condition dating as far back as Galton's work (Blomkvist, 2022; Galton, Francis, 1883), aphantasia has only recently been empirically examined. Seminal research has shown that aphantasia is not simply a psychological condition where individuals are unaware of their own imagery (De Vito & Bartolomeo, 2016). Recent findings have demonstrated that imagery-related effects, or the absence thereof, can be quantified through sensory strength (such as perceptual dominance of color following imagery, Keogh & Pearson, 2018) and pupil dilation (in response to bright versus dark mental images, L. Kay et al., 2022). Importantly, these measures are not subject to volitional control, unlike self-reported vividness ratings. Thus, the evidence suggests that aphantasia may have neurological underpinnings (Milton et al., 2021), raising questions about which cognitive processes are affected when reliant on imagery. There has been limited progress in investigating whether aphantasics can effectively use templates or imagery to guide their attention. Recent findings suggest that this might not be the case (Monzel et al., 2021), as evidenced by the absence of differences in search behavior between congruent and incongruent trials. Given the existence of certain behavioral deficits associated with aphantasia, this condition could serve as a 'knock-out' model of imagery (Muraki et al., 2023) for assessing the causal role of imagery.

Sampling from the aphantasic population could also help resolve several limitations in the current work. First, there were concerns that intertrial priming contributed to the effects in Study 1. If that were the case, then an aphantasic group

should exhibit similar results to the active control group. Second, it was unclear whether the mental task in the active control group successfully prohibited the use of imagery. An aphantasic group would resolve this concern, as they naturally cannot use imagery. Additionally, it would help identify any other strategies deployed aside from imagery. Third, if pop-out effects were present in Study 1, then an aphantasic group should exhibit results according to the stimulus-driven account, assuming that no other strategies were deployed. Lastly, if vividness was responsible for the effects observed in Study 4, these effects should be absent in individuals with aphantasia. All in all, aphantasia provides a promising avenue for assessing the causal role of imagery in attentional processing, strengthening our understanding of imagery as a top-down process.

7.4.4. Is imagery the best strategy?

During the imagery debate between the depictive and propositional theories, a previous study showed that those who sided with the depictive theory described their imagery as visual and those who sided with the propositional described their imagery as symbolic (Reisberg et al., 2003). Given the naturalistic differences in imagery skills, an important question is whether imagery is the best strategy for everyday search. In a related field about visual working memory (WM), a previous study showed that imagery strength correlated with WM performance in tasks that required memorizing visual information but not in tasks that required memorizing verbal memory (Keogh & Pearson, 2014). In other words, a poor imager does not benefit from using imagery in certain tasks. Another study showed that those with aphantasia performed similarly to those without in WM tasks (Keogh et al., 2021), further suggesting that imagery is not the best strategy for

certain individuals. These findings suggest that the utility of imagery as an explicit strategy may vary depending on the task and individual abilities. While vivid imagery can enhance performance in tasks heavily reliant on visual detail, it may not be the most effective approach for those with lower imagery vividness. This underscores the importance of considering individual differences in cognitive strategies and adapting them to the demands of real-world tasks.

Given the differences in how individuals experience imagery, it is plausible that some may rely on alternative strategies to guide attention in search tasks. For example, in some of my lab's work, we implemented search tasks in which the target was a category of items (e.g., animate-inanimate items, Moon, He, et al., 2022); alphanumeric items, Wu et al., 2013). Since the potential targets varied significantly in visual details, participants likely relied on non-visual information, such as the abstract concept of an animal or letter. Importantly, N2pc effects were observed in these category search tasks, suggesting that participants were able to direct attention effectively without or in addition to a visual strategy. Putting it together, future studies should consider not just the individual differences in imagery skills but also the differences in strategies deployed. The paradigm here enforced imagery as the default strategy, though it cannot be ruled out that imagery was in fact the best strategy. Differences in strategies may be particularly important for individuals with limited imagery capabilities, further emphasizing the need to consider a broader range of attentional strategies beyond mental imagery.

7.5. Unresolved issues to address in future studies

7.5.1. Active control group

The control condition in Experiment 2 did not successfully eliminate the behavioral effects observed in the imagery group. One possibility is that the active control task used, while designed to prevent the use of imagery, may not have been fully effective in this regard. Participants in the control group may have still engaged in some form of mental visualization or utilized other cognitive strategies that influenced their performance, thereby confounding the comparison with the imagery group. Additionally, intertrial priming might have contributed to the observed effects. One constraint to the paradigm was that the task should not lead to additional EEG artifacts. For example, previous studies used verbal articulation to minimize the use of imagery (Cochrane et al., 2018), but such movement would have introduced several artifacts here and cost excessive EEG trials. To address this issue in future studies, it will be crucial to develop a more robust control task that prevents the use of imagery. As previously suggested, recruiting individuals with aphantasia, who naturally lack the ability to form mental images, could serve as a more effective control group, as they would not be able to rely on imagery regardless of the task setup. By refining the control condition, future studies can more accurately isolate the role of imagery in attentional processes and ensure that the observed effects are not confounded by alternative strategies or uncontrolled cognitive processes.

7.5.2. Unclear effects on post-selection

Study 3 showed unexpected results, with SPCN effects in the locate task but not the identify task. This study was the only one that did not have additional experiments to replicate the effects. To speculate with the limited dataset, the findings go against the notion that the top-down effects originate at the post-selection stage, going against the stimulus-driven account (Theeuwes, 2008). The absence of SPCN effects in the identify task indicates that post-selection processes are not necessary for top-down effects to occur. Instead, the results suggest that top-down influences on attention may occur earlier, during the selection stage, aligning with the robust behavioral effects observed both here and in other imagery-search studies (Cochrane et al., 2021).

The previous points raise an important question: When does post-selection become sensitive to top-down influences? Additionally, under what conditions does imagery specifically facilitate object recognition? These questions remain unanswered, as the results suggest that post-selection may not always be influenced by imagery. Theories positing that top-down effects primarily occur at the post-selection stage may need revision if post-selection is not consistently influenced by imagery (Theeuwes, 2008). Furthermore, understanding how and when imagery impacts post-selection could aid in identifying specific conditions or individuals who benefit most from imagery strategies. If imagery affects not only the selection but also the post-selection stages, then experimental tasks must be optimized to capture effects at later stages of search. It is also possible that some conditions benefit from imagery during post-selection but not during selection. These findings would present intriguing implications for the stimulus-driven

account, which asserts that top-down effects are exclusively targeted at the post-selection stage (Theeuwes, 2008). Future research must investigate these dynamics to better understand the nuanced role of imagery in visual tasks.

7.5.3. Limitation of the MVPA technique

The MVPA technique in Study 4 revealed differences in neural signals based on vividness, highlighting variations in participants' subjective experiences with imagery from a large scale dataset. This method allowed for tracking variations on a trial-by-trial basis (Dijkstra et al., 2017), rather than relying solely on averaged vividness scores (Cui et al., 2008). This approach provided a finer-grained analysis of how vividness influences neural activity. However, the MVPA technique has notable limitations that future studies should address. The differences in neural signals based on vividness do not inherently mean that these differences were directly driven by vividness. The MVPA results demonstrate associations rather than causal relationships. Differences in neural signals could be influenced by other factors related to task performance which may not be directly related to vividness.

Because trials between the vividness categories needed to be approximately split, participants who reported consistent ratings were excluded from the analyses. This exclusion criterion inadvertently removed participants with poor imagery (those reporting low vividness on most trials) and those with exceptionally vivid imagery (those reporting high vividness). Aphantasia, characterized by low vividness scores, and hyperphantasia, characterized by extremely vivid imagery, represent these extremes (Milton et al., 2021; Zeman et al., 2015). The exclusion of these participants could skew results by not fully

representing the range of imagery experiences, potentially overlooking significant variations in how vividness affects attentional processing. Additionally, because data was analyzed collectively across studies, participants imagined different stimuli, ranging from simple features in Studies 2 and 3 to complex, whole items in Study 1. It is uncertain whether subjective experiences of imagery were consistent across these different stimulus sets, which could introduce variability in the results. Addressing these limitations is crucial for a more accurate understanding of the role of vividness in attentional processing. Future research should aim to include a broader range of imagery experiences and carefully control stimulus variability to better elucidate the modulating role of vividness.

7.5.4. Did participants actually follow the imagery instruction?

While the observed differences between trial types across all experiments bolster confidence that participants engaged in imagery, this question cannot be answered with complete certainty. Comparing trial types is a common approach to assess the effects of imagery, as demonstrated in similar studies (Cochrane et al., 2018; Monzel et al., 2021). However, adherence to imagery instructions can be challenging to verify, as participants' internal experiences were not directly observable. Self-reported data on imagery vividness may not always accurately reflect whether participants engaged in the intended mental processes (Dijkstra et al., 2017). The variability in imagery skills may partially reflect how well participants followed imagery instructions. As previously discussed, individuals with different levels of imagery vividness, such as those with aphantasia or hyperphantasia (Milton et al., 2021; Zeman et al., 2015), or just those with poor or good

imagery without these diagnoses (Keogh & Pearson, 2014), might engage with the task differently, even when instructed to use imagery.

Lastly, the effectiveness of imagery could have varied. Interestingly, previous studies proposed pupil dilation as a physiological marker for imagery (Binda et al., 2013; L. Kay et al., 2022). For example, imagery of bright and dark stimuli was followed by dilation and contraction, respectively (Kay et al., 2022), similar to what the pupils do when directly perceiving the same stimuli (Mathot, 2018). In conclusion, while the differences observed between trial types suggest that imagery was likely used, confirming adherence to imagery instructions remains a challenge. Addressing these issues in future research will help improve the accuracy of imagery studies and enhance our understanding of the role of mental imagery in attentional processing.

7.6. Conclusion

In conclusion, this dissertation has examined the complex interplay between mental imagery and visual attention. The impact of imagery on attentional process aligns with the theoretical framework of the goal-driven account. The integration of EEG data with behavioral measures has provided a deeper understanding of the attentional mechanisms underlying top-down effects. Collectively, the presented studies contribute to a more nuanced view of attentional mechanisms, emphasizing the dynamic interaction between top-down and bottom-up processes. This work contributes to updating existing theoretical models, paving the way for future research that can further refine current understanding of these processes.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. *Current Biology*, *23*(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

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

Bae, G.-Y., & Luck, S. J. (2019). Decoding motion direction using the topography of sustained ERPs and alpha oscillations. *NeuroImage*, *184*, 242–255. https://doi.org/10.1016/j.neuroimage.2018.09.029

Bannert, M. M., & Bartels, A. (2018). Human V4 Activity Patterns Predict Behavioral Performance in Imagery of Object Color. *Journal of Neuroscience*, *38*(15), 3657–3668. https://doi.org/10.1523/JNEUROSCI.2307-17.2018

Bartolomeo, P. (2002). The Relationship Between Visual Perception and Visual Mental Imagery: A Reappraisal of the Neuropsychological Evidence. *Cortex*, *38*(3), 357–378. https://doi.org/10.1016/S0010-9452(08)70665-8

Bayet, L., Zinszer, B., Pruitt, Z., Aslin, R. N., & Wu, R. (2018). Dynamics of neural representations when searching for exemplars and categories of human and non-human faces. *Scientific Reports*, *8*(1), 13277. https://doi.org/10.1038/s41598-018-31526-y

Belopolsky, A. V., Zwaan, L., Theeuwes, J., & Kramer, A. F. (2007). The size of an attentional window modulates attentional capture by color singletons. *Psychonomic Bulletin & Review*, *14*(5), 934–938. https://doi.org/10.3758/BF03194124

Berggren, N., & Eimer, M. (2018). Object-based target templates guide attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1368–1382. https://doi.org/10.1037/xhp0000541

Bergmann, J., Genç, E., Kohler, A., Singer, W., & Pearson, J. (2016). Smaller Primary Visual Cortex Is Associated with Stronger, but Less Precise Mental Imagery. *Cerebral Cortex*, *26*(9), 3838–3850. https://doi.org/10.1093/cercor/bhv186

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Pupil constrictions to photographs of the sun. *Journal of Vision*, *13*(6), 8–8. https://doi.org/10.1167/13.6.8

Bird, C. M., Capponi, C., King, J. A., Doeller, C. F., & Burgess, N. (2010). Establishing the Boundaries: The Hippocampal Contribution to Imagining Scenes. *Journal of Neuroscience*, *30*(35), 11688–11695. https://doi.org/10.1523/JNEUROSCI.0723-10.2010

Blomkvist, A. (2022). Aphantasia: In search of a theory. *Mind & Language*, mila.12432. https://doi.org/10.1111/mila.12432

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology*, *55*(6), e13049. https://doi.org/10.1111/psyp.13049

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics*, *51*(5), 465–472. https://doi.org/10.3758/BF03211642

Brisson, B., & Jolicœur, P. (2007). The N2pc component and stimulus duration. *NeuroReport*, *18*(11), 1163. https://doi.org/10.1097/WNR.0b013e3281e72d1b

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. *PLOS ONE*, *9*(9), e106953. https://doi.org/10.1371/journal.pone.0106953

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, *13*(6), 407–420. https://doi.org/10.1038/nrn3241

Cabbai, G., Brown, C. R. H., Dance, C., Simner, J., & Forster, S. (2023). Mental imagery and visual attentional templates: A dissociation. *Cortex*, *169*, 259–278. https://doi.org/10.1016/j.cortex.2023.09.014

Carnevale, N. T., & Hines, M. L. (2006). *The NEURON Book*. Cambridge University Press.

Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, *2*(7), 671–676. https://doi.org/10.1038/10230

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of Neurons in Inferior Temporal Cortex During Memory-Guided Visual Search. *Journal of Neurophysiology*, *80*(6), 2918–2940. https://doi.org/10.1152/jn.1998.80.6.2918

Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012). Imagery and Perception Share Cortical Representations of Content and Location. *Cerebral Cortex*, *22*(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Cochrane, B. A., & Milliken, B. (2019). Imagined event files: An interplay between imagined and perceived objects. *Psychonomic Bulletin & Review*, *26*(2), 538–544. https://doi.org/10.3758/s13423-019-01572-2

Cochrane, B. A., Ng, V., Khosla, A., & Milliken, B. (2021). Looking into the mind's eye: Directed and evaluated imagery vividness modulates imagery-perception congruency effects. *Psychonomic Bulletin & Review*, *28*(3), 862–869. https://doi.org/10.3758/s13423-020-01868-8

Cochrane, B. A., Ng, V., & Milliken, B. (2021). Top-down imagery overrides the influence of selection history effects. *Consciousness and Cognition*, *93*, 103153. https://doi.org/10.1016/j.concog.2021.103153

Cochrane, B. A., Nwabuike, A. A., Thomson, D. R., & Milliken, B. (2018). An imagery-induced reversal of intertrial priming in visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 572–587. https://doi.org/10.1037/xlm0000470

Cochrane, B. A., Townsend, B., Lee, R. C., Legere, J. K., Milliken, B., & Shedden, J. M. (2021). Visual imagery influences attentional guidance during visual search: Behavioral and electrophysiological evidence. *Attention, Perception, & Psychophysics*, *83*(1), 58–66. https://doi.org/10.3758/s13414-020-02208-7

Cochrane, B. A., Wang, C., Pratt, J., Milliken, B., & Sun, H. (2021). Comparing imagery and perception: Using eye movements to dissociate mechanisms in search. *Attention, Perception, & Psychophysics*, *83*(7), 2879–2890. https://doi.org/10.3758/s13414-021-02336-8

Conci, M., Deichsel, C., Müller, H. J., & Töllner, T. (2019). Feature guidance by negative attentional templates depends on search difficulty. *Visual Cognition*, *27*(3–4), 317–326. https://doi.org/10.1080/13506285.2019.1581316

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, *47*(4), 474–478. https://doi.org/10.1016/j.visres.2006.11.013

De Vito, S., & Bartolomeo, P. (2016). Refusing to imagine? On the possibility of psychogenic aphantasia. A commentary on Zeman et al. (2015). *Cortex*, *74*, 334–335. https://doi.org/10.1016/j.cortex.2015.06.013

Debanne, D. (2004). Information processing in the axon. *Nature Reviews Neuroscience*, *5*(4), 304–316. https://doi.org/10.1038/nrn1397

Dell'Acqua, R., Sessa, P., Jolicœur, P., & Robitaille, N. (2006). Spatial attention freezes during the attention blink. *Psychophysiology*, *43*(4), 394–400. https://doi.org/10.1111/j.1469-8986.2006.00411.x

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

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

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2017). Vividness of Visual Imagery Depends on the Neural Overlap with Perception in Visual Areas. *The Journal of Neuroscience*, *37*(5), 1367–1373. https://doi.org/10.1523/JNEUROSCI.3022-16.2016

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2019). Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences*, *23*(5), 423–434. https://doi.org/10.1016/j.tics.2019.02.004

Dosher, B. A., Han, S., & Lu, Z.-L. (2004). Parallel Processing in Visual Search Asymmetry. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 3–27. https://doi.org/10.1037/0096-1523.30.1.3

Dowdall, J. R., Luczak, A., & Tata, M. S. (2012). Temporal variability of the N2pc during efficient and inefficient visual search. *Neuropsychologia*, *50*(10), 2442–2453. https://doi.org/10.1016/j.neuropsychologia.2012.06.015

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Eimer, M. (1996). *The N2pc component as an indicator of attentional selectivity*. 10.

Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526–535. https://doi.org/10.1016/j.tics.2014.05.005

Eimer, M., & Kiss, M. (2008). Involuntary Attentional Capture is Determined by Task Set: Evidence from Event-related Brain Potentials. *Journal of Cognitive Neuroscience*, *20*(8), 1423–1433. https://doi.org/10.1162/jocn.2008.20099

Eimer, M., & Kiss, M. (2010). The top-down control of visual selection and how it is linked to the N2pc component. *Acta Psychologica*, *135*(2), 100–102. https://doi.org/10.1016/j.actpsy.2010.04.010

Ferguson, L., Ahmed, C., Dang, C., & Wu, R. (2021). Neural target selection as a marker of real-world familiarity during search for perceptually distinct objects. *European Journal of Neuroscience*, *53*(5), 1517–1532. https://doi.org/10.1111/ejn.15051

Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 847–858. https://doi.org/10.1037/0096-1523.24.3.847

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030–1044. https://doi.org/10.1037/0096-1523.18.4.1030

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 317–329. https://doi.org/10.1037/0096-1523.20.2.317

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, *65*(7), 999–1010. https://doi.org/10.3758/BF03194829

Galton, Francis. (1883). *Inquiries into human faculty and its development*. Macmillan.

Ghorashi, S., Enns, J. T., Klein, R. M., & Di Lollo, V. (2010). Spatial selection and target identification are separable processes in visual search. *Journal of Vision*, *10*(3), 7. https://doi.org/10.1167/10.3.7

Giesbrecht, B., Weissman, D. H., Woldorff, M. G., & Mangun, G. R. (2006). Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*, *1080*(1), 63–72. https://doi.org/10.1016/j.brainres.2005.09.068

Grubert, A., Fahrenfort, J., Olivers, C. N. L., & Eimer, M. (2017). Rapid top-down control over template-guided attention shifts to multiple objects. *NeuroImage*, *146*, 843–858. https://doi.org/10.1016/j.neuroimage.2016.08.039

Guo, Z., Niu, M., & Wang, Q. (2021). Target familiarity and visual working memory do not influence familiarity effect in visual search. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-86669-2

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, *104*(5), 1726–1731. https://doi.org/10.1073/pnas.0610561104

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., & Cichy, R. M. (2018). The representational dynamics of task and object processing in humans. *eLife*, *7*, e32816. https://doi.org/10.7554/eLife.32816

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological Evidence of the Capture of Visual Attention. *Journal of Cognitive Neuroscience*, *18*(4), 604–613. Journal of Cognitive Neuroscience. https://doi.org/10.1162/jocn.2006.18.4.604

Horton, J. C., & Hoyt, W. F. (1991). The Representation of the Visual Field in Human Striate Cortex: A Revision of the Classic Holmes Map. *Archives of Ophthalmology*, *109*(6), 816–824. https://doi.org/10.1001/archopht.1991.01080060080030

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, *114*(3), 599–631. https://doi.org/10.1037/0033-295X.114.3.599

Huffmeijer, R., Bakermans-Kranenburg, M. J., Alink, L. R. A., & van IJzendoorn, M. H. (2014). Reliability of event-related potentials: The influence of number of trials and electrodes. *Physiology & Behavior*, *130*, 13–22. https://doi.org/10.1016/j.physbeh.2014.03.008

Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed Neural Systems for the Generation of Visual Images. *Neuron*, *3*(28), 979–990.

Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex*, *105*, 61–73. https://doi.org/10.1016/j.cortex.2017.10.014

Jolicœur, P., Brisson, B., & Robitaille, N. (2008). Dissociation of the N2pc and sustained posterior contralateral negativity in a choice response task. *Brain Research*, *1215*, 160–172. https://doi.org/10.1016/j.brainres.2008.03.059

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*(4), 346–354. https://doi.org/10.3758/BF03208805

Kappenman, E. S., Farrens, J. L., Zhang, W., Stewart, A. X., & Luck, S. J. (2021). ERP CORE: An open resource for human event-related potential research. *NeuroImage*, *225*, 117465. https://doi.org/10.1016/j.neuroimage.2020.117465

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature*, *452*(7185), 352–355. https://doi.org/10.1038/nature06713

Kay, L., Keogh, R., Andrillon, T., & Pearson, J. (2022). The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength. *eLife*, *11*, e72484. https://doi.org/10.7554/eLife.72484

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, *14*(12), 7–7. https://doi.org/10.1167/14.12.7

Keogh, R., & Pearson, J. (2017). The perceptual and phenomenal capacity of mental imagery. *Cognition*, *162*, 124–132. https://doi.org/10.1016/j.cognition.2017.02.004

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, *105*, 53–60. https://doi.org/10.1016/j.cortex.2017.10.012

Keogh, R., & Pearson, J. (2021). Attention driven phantom vision: Measuring the sensory strength of attentional templates and their relation to visual mental imagery and aphantasia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1817), 20190688. https://doi.org/10.1098/rstb.2019.0688

Keogh, R., Wicken, M., & Pearson, J. (2021). Visual working memory in aphantasia: Retained accuracy and capacity with a different strategy. *Cortex*, *143*, 237–253. https://doi.org/10.1016/j.cortex.2021.07.012

Kiesel, A., Miller, J., Jolicœur, P., & Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, *45*(2), 250–274. https://doi.org/10.1111/j.1469-8986.2007.00618.x

Kiss, M., Grubert, A., Petersen, A., & Eimer, M. (2012). Attentional Capture by Salient Distractors during Visual Search Is Determined by Temporal Task Demands. *Journal of Cognitive Neuroscience*, *24*(3), 749–759. https://doi.org/10.1162/jocn_a_00127

Kiss, M., Jolicœur, P., Dell'Acqua, R., & Eimer, M. (2008). Attentional capture by visual singletons is mediated by top-down task set: New evidence from the N2pc component. *Psychophysiology*, *45*(6), 1013–1024. https://doi.org/10.1111/j.1469-8986.2008.00700.x

Kiss, M., Van Velzen, J., & Eimer, M. (2008). The N2pc component and its links to attention shifts and spatially selective visual processing. *Psychophysiology*, *45*(2), 240–249. https://doi.org/10.1111/j.1469-8986.2007.00611.x

Koenig-Robert, R., & Pearson, J. (2019). Decoding the contents and strength of imagery before volitional engagement. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39813-y

Kok, P., Brouwer, G. J., Gerven, M. A. J. van, & Lange, F. P. de. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *Journal of Neuroscience*, *33*(41), 16275–16284. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Kosslyn, S. M. (1975). Information representation in visual images. *Cognitive Psychology*, *7*(3), 341–370. https://doi.org/10.1016/0010-0285(75)90015-8

Kosslyn, S. M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review*, *88*(1), 46–66. https://doi.org/10.1037/0033-295X.88.1.46

Kosslyn, S. M. (2005). Mental images and the Brain. *Cognitive Neuropsychology*, *22*(3–4), 333–347. https://doi.org/10.1080/02643290442000130

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, *129*(5), 723–746. https://doi.org/10.1037/0033-2909.129.5.723

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The Case for Mental Imagery*. Oxford University Press.

Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *NeuroImage*, *59*(4), 4064–4073. https://doi.org/10.1016/j.neuroimage.2011.10.055

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00213

Luck, S. J. (2011). *Electrophysiological Correlates of the Focusing of Attention within Complex Visual Scenes: N2pc and Related ERP Components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0161

Luck, S. J. (2012). Event-related potentials. In *APA handbook of research methods in psychology, Vol 1: Foundations, planning, measures, and psychometrics* (pp. 523–546). American Psychological Association. https://doi.org/10.1037/13619-028

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1000–1014. https://doi.org/10.1037/0096-1523.20.5.1000

Luck, S. J., & Kappenman, E. S. (2013). *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press.

Malinowski, P., & Hübner, R. (2001). The effect of familiarity on visual-search performance: Evidence for learned basic features. *Perception & Psychophysics*, *63*(3), 458–463. https://doi.org/10.3758/BF03194412

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, *22*(6), 657–672. https://doi.org/10.3758/BF03209251

Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: II. The role of position. *Perception & Psychophysics*, *58*(7), 977–991. https://doi.org/10.3758/BF03206826

Marks, D. F. (1973). Visual Imagery Differences in the Recall of Pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Mathot, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, *1*(1), Article 1. https://doi.org/10.5334/joc.18

Mazza, V., Turatto, M., & Caramazza, A. (2009). Attention selection, distractor suppression and N2pc. *Cortex*, *45*(7), 879–890. https://doi.org/10.1016/j.cortex.2008.10.009

Mazza, V., Turatto, M., Umiltà, C., & Eimer, M. (2007). Attentional selection and identification of visual objects are reflected by distinct electrophysiological responses. *Experimental Brain Research*, *181*(3), 531–536. https://doi.org/10.1007/s00221-007-1002-4

Mechelli, A. (2004). Where Bottom-up Meets Top-down: Neuronal Interactions during Perception and Imagery. *Cerebral Cortex*, *14*(11), 1256–1265. https://doi.org/10.1093/cercor/bhh087

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, *24*(Volume 24, 2001), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Milton, F., Fulford, J., Dance, C., Gaddum, J., Heuerman-Williamson, B., Jones, K., Knight, K. F., MacKisack, M., Winlove, C., & Zeman, A. (2021). Behavioral and Neural Signatures of Visual Imagery Vividness Extremes: Aphantasia versus Hyperphantasia. *Cerebral Cortex Communications*, *2*(2), tgab035. https://doi.org/10.1093/texcom/tgab035

Monzel, M., Keidel, K., & Reuter, M. (2021). Imagine, and you will find – Lack of attentional guidance through visual imagery in aphantasics. *Attention, Perception, & Psychophysics*, *83*(6), 2486–2497. https://doi.org/10.3758/s13414-021-02307-z

Moon, A., He, C., Ditta, A. S., Cheung, O. S., & Wu, R. (2022). Rapid category selectivity for animals versus man-made objects: An N2pc study. *International Journal of Psychophysiology*, *171*, 20–28. https://doi.org/10.1016/j.ijpsycho.2021.11.004

Moon, A., Zhao, J., Peters, M. A. K., & Wu, R. (2022). Interaction of prior category knowledge and novel statistical patterns during visual search for real-world objects. *Cognitive*

Research: Principles and Implications, *7*(1), 21. https://doi.org/10.1186/s41235-022-00356-y

Moriya, J. (2018). Visual mental imagery influences attentional guidance in a visual-search task. *Attention, Perception, & Psychophysics*, *80*(5), 1127–1142. https://doi.org/10.3758/s13414-018-1520-0

Moriya, J. (2023). Visual mental imagery of atypical color objects attracts attention to an imagery-matching object. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-023-02804-3

Mruczek, R. E. B., & Sheinberg, D. L. (2005). Distractor familiarity leads to more efficient visual search for complex stimuli. *Perception & Psychophysics*, *67*(6), 1016–1031. https://doi.org/10.3758/BF03193628

Müller, H., Reimann, B., & Krummenacher, J. (2003). Visual Search for Singleton Feature Targets Across Dimensions: Stimulus- and Expectancy-Driven Effects in Dimensional Weighting. *Journal of Experimental Psychology. Human Perception and Performance*, *29*, 1021–1035. https://doi.org/10.1037/0096-1523.29.5.1021

Muraki, E. J., Speed, L. J., & Pexman, P. M. (2023). Insights into embodied cognition and mental imagery from aphantasia. *Nature Reviews Psychology*, *2*(10), Article 10. https://doi.org/10.1038/s44159-023-00221-9

Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014). Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans? *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1283–1288. https://doi.org/10.1037/a0036885

Naselaris, T., Kay, K. N., Nishimoto, S., & Gallant, J. L. (2011). Encoding and decoding in fMRI. *NeuroImage*, *56*(2), 400–410. https://doi.org/10.1016/j.neuroimage.2010.07.073

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage*, *105*. https://doi.org/10.1016/j.neuroimage.2014.10.018

Niedermeyer, E. (1996). Dipole Theory and Electroencephalography. *Clinical Electroencephalography*, *27*(3), 121–131. https://doi.org/10.1177/155005949602700305

Nunez, P. L., & Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press.

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, S1364661311000854. https://doi.org/10.1016/j.tics.2011.05.004

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, *20*(10), 624–634. https://doi.org/10.1038/s41583-019-0202-9

Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The Functional Impact of Mental Imagery on Conscious Perception. *Current Biology*, *18*(13), 982–986. https://doi.org/10.1016/j.cub.2008.05.048

Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *Proceedings of the National Academy of Sciences*, *112*(33), 10089–10092. https://doi.org/10.1073/pnas.1504933112

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, *19*(10), 590–602. https://doi.org/10.1016/j.tics.2015.08.003

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the Mind's Eye: The Metacognition of Visual Imagery. *Psychological Science*, *22*(12), 1535–1542. https://doi.org/10.1177/0956797611417134

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Pylyshyn, Z. (2003). Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Sciences*, *7*(3), 113–118. https://doi.org/10.1016/S1364-6613(03)00003-2

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*(1), 1–24. https://doi.org/10.1037/h0034650

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*(1), 16–45. https://doi.org/10.1037/0033-295X.88.1.16

Ragni, F., Lingnau, A., & Turella, L. (2021). Decoding category and familiarity information during visual imagery. *NeuroImage*, *241*, 118428. https://doi.org/10.1016/j.neuroimage.2021.118428

Ramgir, A., & Lamy, D. (2022). Does feature intertrial priming guide attention? The jury is still out. *Psychonomic Bulletin & Review*, *29*(2), 369–393. https://doi.org/10.3758/s13423-021-01997-8

Ranganath, C., & D'Esposito, M. (2005). Directing the mind's eye: Prefrontal, inferior and medial temporal mechanisms for visual working memory. *Current Opinion in Neurobiology*, *15*(2), 175–182. https://doi.org/10.1016/j.conb.2005.03.017

Rashal, E., Santandrea, E., Ben-Hamed, S., Macaluso, E., Chelazzi, L., & Boehler, C. N. (2023). Effects of top-down and bottom-up attention on post-selection posterior contralateral

negativity. *Attention, Perception, & Psychophysics*, *85*(3), 705–717. https://doi.org/10.3758/s13414-022-02636-7

Reisberg, D., Pearson, D. G., & Kosslyn, S. M. (2003). Intuitions and introspections about imagery: The role of imagery experience in shaping an investigator's theoretical views. *Applied Cognitive Psychology*, *17*(2), 147–160. https://doi.org/10.1002/acp.858

Rietdijk, W. J. R., Franken, I. H. A., & Thurik, A. R. (2014). Internal Consistency of Event-Related Potentials Associated with Cognitive Control: N2/P3 and ERN/Pe. *PLOS ONE*, *9*(7), e102672. https://doi.org/10.1371/journal.pone.0102672

Runge, M., Bakhilau, V., Omer, F., & D'Angiulli, A. (2015). Trial-by-Trial Vividness Self-Reports Versus VVIQ: A Meta-Analytic Comparison of Behavioral, Cognitive and Neurological Correlations. *Imagination, Cognition and Personality*, *35*(2), 137–165. https://doi.org/10.1177/0276236615587490

Runge, M. S., Cheung, M. W.-L., & D'Angiulli, A. (2017). Meta-analytic comparison of trial- versus questionnaire-based vividness reportability across behavioral, cognitive and neural measurements of imagery. *Neuroscience of Consciousness*, *2017*(1). https://doi.org/10.1093/nc/nix006

Saiki, J., Koike, T., Takahashi, K., & Inoue, T. (2005). Visual search asymmetry with uncertain targets. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1274–1287. https://doi.org/10.1037/0096-1523.31.6.1274

Sawaki, R., Geng, J. J., & Luck, S. J. (2012). A Common Neural Mechanism for Preventing and Terminating the Allocation of Attention. *Journal of Neuroscience*, *32*(31), 10725–10736. https://doi.org/10.1523/JNEUROSCI.1864-12.2012

Sawaki, R., & Luck, S. J. (2013). Active suppression after involuntary capture of attention. *Psychonomic Bulletin & Review*, *20*(2), 296–301. https://doi.org/10.3758/s13423-012-0353-4

Serences, J. T., Yantis, S., Culberson, A., & Awh, E. (2004). Preparatory Activity in Visual Cortex Indexes Distractor Suppression During Covert Spatial Orienting. *Journal of Neurophysiology*, *92*(6), 3538–3545. https://doi.org/10.1152/jn.00435.2004

Shen, J., & Reingold, E. M. (2001). Visual search asymmetry: The influence of stimulus familiarity and low-level features. *Perception & Psychophysics*, *63*(3), 464–475. https://doi.org/10.3758/BF03194413

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas. *Cerebral Cortex*, *15*(10), 1570–1583. https://doi.org/10.1093/cercor/bhi035

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-Down Activation of Shape-Specific Population Codes in Visual Cortex during Mental Imagery. *Journal of Neuroscience*, *29*(5), 1565–1572. https://doi.org/10.1523/JNEUROSCI.4657-08.2009

Stuart, G., Spruston, N., & Häusser, M. (2007). *Dendrites*. Oxford University Press.

Südhof, T. C., & Malenka, R. C. (2008). Understanding Synapses: Past, Present, and Future. *Neuron*, *60*(3), 469–476. https://doi.org/10.1016/j.neuron.2008.10.011

Teixidó, M., François, C., Bosch, L., & Männel, C. (2018). Chapter 5. The role of prosody in early speech segmentation and word-referent mapping. In E. by P. Prieto & N. Esteve-Gibert (Eds.), *The Development of Prosody in First Language Acquisition* (pp. 79–100). https://doi.org/10.1075/tilar.23.05tei

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. https://doi.org/10.3758/BF03211619

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Theeuwes, J. (1994). Endogenous and Exogenous Control of Visual Selection. *Perception*, *23*(4), 429–440. https://doi.org/10.1068/p230429

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

Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, *1*(1), 29. https://doi.org/10.5334/joc.13

Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1342–1353. https://doi.org/10.1037/0096-1523.24.5.1342

Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition*, *14*(4–8), 466–489. https://doi.org/10.1080/13506280500195110

Thigpen, N. N., Kappenman, E. S., & Keil, A. (2017). Assessing the internal consistency of the event-related potential: An example analysis. *Psychophysiology*, *54*(1), 123–138. https://doi.org/10.1111/psyp.12629

Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J.-B., Lebihan, D., & Dehaene, S. (2006). Inverse retinotopy: Inferring the visual content of images from brain activation patterns. *NeuroImage*, *33*(4), 1104–1116. https://doi.org/10.1016/j.neuroimage.2006.06.062

Töllner, T., Zehetleitner, M., Gramann, K., & Müller, H. J. (2010). Top-down weighting of visual dimensions: Behavioral and electrophysiological evidence. *Vision Research*, *50*(14), 1372–1381. https://doi.org/10.1016/j.visres.2009.11.009

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*(1), 15–48. https://doi.org/10.1037/0033-295X.95.1.15

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282–299. https://doi.org/10.1037/h0027242

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, *114*(3), 285–310. https://doi.org/10.1037/0096-3445.114.3.285

Vickery, T. J., King, L.-W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1), 8. https://doi.org/10.1167/5.1.8

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

Wang, B., & Theeuwes, J. (2020). Salience determines attentional orienting in visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(10), 1051–1057. https://doi.org/10.1037/xhp0000796

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, *56*(5), 495–500. https://doi.org/10.3758/BF03206946

Wantz, A. L., Borst, G., Mast, F. W., & Lobmaier, J. S. (2015). Colors in mind: A novel paradigm to investigate pure color imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1152–1161. https://doi.org/10.1037/xlm0000079

Winlove, C. I. P., Milton, F., Ranson, J., Fulford, J., MacKisack, M., Macpherson, F., & Zeman, A. (2018). The neural correlates of visual imagery: A co-ordinate-based meta-analysis. *Cortex*, *105*, 4–25. https://doi.org/10.1016/j.cortex.2017.12.014

Wolfe, J. M. (1998). What Can 1 Million Trials Tell Us About Visual Search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics*, *63*(3), 381–389. https://doi.org/10.3758/BF03194406

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 483–502. https://doi.org/10.1037/0096-1523.29.2.483

Wolfe, J. M., & Friedman-Hill, S. R. (1992). On the Role of Symmetry in Visual Search. *Psychological Science*, *3*(3), 194–198. https://doi.org/10.1111/j.1467-9280.1992.tb00026.x

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), Article 3. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, *44*(12), 1411–1426. https://doi.org/10.1016/j.visres.2003.11.024

Wolfe, J. M., Palmer, E. M., & Horowitz, T. S. (2010). Reaction time distributions constrain models of visual search. *Vision Research*, *50*(14), 1304–1311. https://doi.org/10.1016/j.visres.2009.11.002

Woodman, G. F., Arita, J. T., & Luck, S. J. (2009). A cuing study of the N2pc component: An index of attentional deployment to objects rather than spatial locations. *Brain Research*, *1297*, 101–111. https://doi.org/10.1016/j.brainres.2009.08.011

Woodman, G. F., & Chun, M. M. (2006). The role of working memory and long-term memory in visual search. *Visual Cognition*, *14*(4–8), 808–830. https://doi.org/10.1080/13506280500197397

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, *400*(6747), 867–869. https://doi.org/10.1038/23698

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 121–138. https://doi.org/10.1037/0096-1523.29.1.121

Wöstmann, M., Störmer, V. S., Obleser, J., Addleman, D. A., Andersen, S. K., Gaspelin, N., Geng, J. J., Luck, S. J., Noonan, M. P., Slagter, H. A., & Theeuwes, J. (2022). Ten simple rules to study distractor suppression. *Progress in Neurobiology*, *213*, 102269. https://doi.org/10.1016/j.pneurobio.2022.102269

Wu, R., McGee, B., Echiverri, C., & Zinszer, B. D. (2018). Prior knowledge of category size impacts visual search. *Psychophysiology*, *55*(8), e13075. https://doi.org/10.1111/psyp.13075

Wu, R., Nako, R., Band, J., Pizzuto, J., Ghoreishi, Y., Scerif, G., & Aslin, R. (2015). Rapid Attentional Selection of Non-native Stimuli despite Perceptual Narrowing. *Journal of Cognitive Neuroscience*, *27*(11), 2299–2307. https://doi.org/10.1162/jocn_a_00857

Wu, R., Pruitt, Z., Zinszer, B. D., & Cheung, O. S. (2017). Increased experience amplifies the activation of task-irrelevant category representations. *Attention, Perception, & Psychophysics*, *79*(2), 522–532. https://doi.org/10.3758/s13414-016-1254-9

Wu, R., Scerif, G., Aslin, R. N., Smith, T. J., Nako, R., & Eimer, M. (2013). Searching for Something Familiar or Novel: Top–Down Attentional Selection of Specific Items or Object Categories. *Journal of Cognitive Neuroscience*, *25*(5), 719–729. https://doi.org/10.1162/jocn_a_00352

Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*(4), 167–174. https://doi.org/10.1016/j.tics.2009.01.008

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery – Congenital aphantasia. *Cortex*, *73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

Zeman, A., Milton, F., Della Sala, S., Dewar, M., Frayling, T., Gaddum, J., Hattersley, A., Heuerman-Williamson, B., Jones, K., MacKisack, M., & Winlove, C. (2020). Phantasia–The psychological significance of lifelong visual imagery vividness extremes. *Cortex*, *130*, 426–440. https://doi.org/10.1016/j.cortex.2020.04.003

Zhang, Y. (Raven), & Onyper, S. (2020). Visual search asymmetry depends on target-distractor feature similarity: Is the asymmetry simply a result of distractor rejection speed? *Attention, Perception, & Psychophysics*, *82*(1), 80–97. https://doi.org/10.3758/s13414-019-01818-0