YOU’RE MORE FLEXIBLE THAN YOU THINK:
RE-CONCEPTUALIZING THE TOP-DOWN CONTROL OF VISUAL-SPATIAL ATTENTIONAL CAPTURE.

by
Maria Giammarco, MSc

A thesis
presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Doctor of Philosophy
in
Psychology and Neuroscience

Guelph, ON, Canada

© Maria Giammarco, November, 2016
We have long known that unexpected salient properties of our visual environments, like the flashing lights on an ambulance, capture our spatial attention, leading to faster processing of objects appearing in the same spatial location as compared to objects appearing elsewhere. One of the more provocative discoveries of the attentional capture literature is that spatial capture can be constrained by our voluntary goals, or Attentional Control Settings (ACSs). ACSs determine which stimuli or stimulus properties capture spatial attention: When looking for a red object, the unexpected appearance of other red objects will capture attention, while, importantly, the unexpected appearance of green objects will not. Current conceptualizations of ACSs, however, constrain their role in the support of more complex behaviours: The number of items contained within ACSs is believed to be limited to one, perhaps two items, and the role of voluntary control in the implementation of these ACSs (i.e., whether we are able to directly choose what items capture attention) has been challenged by a feature priming model (i.e., prior experience responding to task targets determines what items capture attention). The purpose of this thesis was to examine whether ACSs are more flexible than outlined in current literature in terms of the number of representations they comprise and in the implementation of ACSs. Chapters 1 to 5
used modified spatial capture paradigms to demonstrate that standard measures of spatial capture were contingent on ACSs for up to 30 items stored in episodic long-term memory. Chapters 6 and 7 then contrasted a voluntary model of ACS implementation against a feature priming model. Here, manipulations of feature priming had limited influence on contingent capture; instead, behaviour adhered more closely to a voluntary implementation account, and point towards a potential role for target selection in the maintenance (rather than implementation) of ACSs. These studies re-conceptualize our understanding of ACSs with regard to both the role of the memory systems supporting the contents of ACSs, and of voluntary control in driving the implementation of ACSs, while contributing to the on-going development of more comprehensive models of this sophisticated attentional mechanism of control over capture.
ACKNOWLEDGEMENTS

Naz, thank you for being an amazing advisor, mentor, and aca-dad. Thank you for always pushing me, for being patient and supportive, for fostering a deeper passion and respect for research than I ever thought possible, and, of course, for putting up with Blaire and my antics.

Mark, thank you for all your co-advisor support, advice, and mentorship – and for being my first connection to Guelph.

Blaire, my labmate, roommate, and soulmate. I have no words for all you have done for me and everything you have got me through. I can’t wait to see all the amazing things you are going to (continue to) accomplish. You are the wind beneath my wings.

Tyler, thank you for your unwavering support, love, and understanding. You have always had the confidence in me that I so often need. Thank you for being the best partner I could ask for.

To my lab peeps Rachel and Dave, thank you for your friendships, your countless awesome discussions (research and otherwise), and unending conference fun.

A huge thank you to my thesis committee members Lana Trick and Evan Risko, my defense committee members Harvey Marmurek, Dan Meegan, and Katherine Sledge Moore, and all those who have collaborated on the research in this thesis (Emma B Guild, Adriana Paoletti, Jackson Hryciw, Kate Turner, and Krista Miller). I truly appreciate everything you have done – I would not have been able to complete this work without you all!

Finally, and absolutely most importantly, thank you Sandra, Mom, and Dad. Without you none of this would be remotely possible. Thank you for always trusting my choices without question or hesitation, and for the many ways you have supported my love of learning through the years. Thank you for being immensely proud of every accomplishment, no matter how big or small – it’s what keeps me going.

The research in this thesis was made possible by the financial support of the Natural Sciences and Engineering Research Council of Canada, a Queen Elizabeth II Graduate Scholarship in Science and Technology, and an Ontario Graduate Scholarship.
# TABLE OF CONTENTS

**ABSTRACT** .................................................................................................................................................. ii  
**Acknowledgements** ................................................................................................................................. iv  
**Table of Contents** ....................................................................................................................................... v  
**List of Tables** ............................................................................................................................................... ix  
**List of Figures** ............................................................................................................................................... x  
**Introduction** ............................................................................................................................................... 1  
  Visual Spatial Attention ................................................................................................................................. 3  
  Attentional Control Settings ......................................................................................................................... 5  
  Aim I: The Storage of Attentional Control Set Items ............................................................................... 6  
  Aim II: The Implementation of ACSs .......................................................................................................... 10  

**Chapter 1:** Selective cueing effects for episodic LTM ACS-matching cues ............................................ 15  
  Methods ....................................................................................................................................................... 18  
  Participants .................................................................................................................................................. 18  
  Apparatus, stimuli, and procedure ............................................................................................................. 18  
  Memory training task ................................................................................................................................. 18  
  Attention cueing task ............................................................................................................................... 20  
  Results ......................................................................................................................................................... 22  
  Memory training task ................................................................................................................................. 22  
  Attention cueing task ............................................................................................................................... 22  
  Discussion ................................................................................................................................................... 24  

**Chapter 2:** Selective cueing effects for episodic LTM ACS-matching cues are not impacted by poor task performance ........................................................................................................................................... 26  
  Methods ....................................................................................................................................................... 26
Participants ................................................................................................................................. 43
Apparatus, stimuli, and procedure .............................................................................................. 44
Memory training task .................................................................................................................. 44
RSVP task .................................................................................................................................... 45
Results .......................................................................................................................................... 48
Memory Training Task ................................................................................................................ 48
RSVP task ...................................................................................................................................... 48
Discussion .................................................................................................................................... 50

**Interim Summary:** Chapters 1 to 5 ......................................................................................... 51

**Chapter 6:** Inter-trial feature priming does not explain ACS contingent capture, But longer-lasting priming might .................................................................................................................. 53

Methods ....................................................................................................................................... 55
Participants ................................................................................................................................. 55
Apparatus, stimuli, and procedure .............................................................................................. 56
Memory Training Task ................................................................................................................ 56
RSVP Task ...................................................................................................................................... 56
Results .......................................................................................................................................... 57
Memory Training Task ................................................................................................................ 57
RSVP Task ...................................................................................................................................... 58
Discussion .................................................................................................................................... 62

**Chapter 7:** Longer-lasting priming does not explain ACS contingent capture either; ACSs reflect flexible control over capture ........................................................................................................... 64

Methods ....................................................................................................................................... 64
Participants ................................................................................................................................. 64
Apparatus, stimuli, and procedure .............................................................................................. 65
Memory Training Task ................................................................................................................ 65
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSVP Task</td>
<td>65</td>
</tr>
<tr>
<td>Results</td>
<td>66</td>
</tr>
<tr>
<td>Memory Training Task</td>
<td>66</td>
</tr>
<tr>
<td>RSVP Task</td>
<td>66</td>
</tr>
<tr>
<td>Discussion</td>
<td>70</td>
</tr>
<tr>
<td><strong>Interim Summary: Chapters 6 and 7</strong></td>
<td>79</td>
</tr>
<tr>
<td><strong>General Discussion</strong></td>
<td>81</td>
</tr>
<tr>
<td>Summary</td>
<td>81</td>
</tr>
<tr>
<td>Specifying the mechanisms of ACSs I: The interface between episodic long-term memory and attention</td>
<td>82</td>
</tr>
<tr>
<td>Specifying the mechanisms of ACSs II: ACSs require more than storage</td>
<td>83</td>
</tr>
<tr>
<td>Implications for the implementation of ACSs</td>
<td>85</td>
</tr>
<tr>
<td>The representation of memory items in episodic LTM ACSs: Dependent vs. independent</td>
<td>88</td>
</tr>
<tr>
<td>Specifying the timing of control</td>
<td>89</td>
</tr>
<tr>
<td>Updating our understanding of ACSs</td>
<td>91</td>
</tr>
<tr>
<td>Conclusions</td>
<td>93</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>95</td>
</tr>
</tbody>
</table>
LIST OF TABLES

*Table 1.* Error rates (%) and standard errors of the mean (SE) for Chapter 1. No effects of cueing, cue type, or their interaction on accuracy were observed. .............................................................. 24

*Table 2.* Error rates (%) and standard errors of the mean (SE) for Chapter 2. No effects of cueing, cue type, or their interaction on accuracy were observed. Error rates were significantly lower than in Chapter 1............................................................... 29

*Table 3.* Error rates (%) and standard errors of the mean (SE) for Chapter 3. No effects of cueing, SOA, or their interaction on accuracy were observed. ................................................................. 35

*Table 4.* Error rates (%) and standard errors of the mean (SE) for Chapter 4. No effects of cueing were observed. ............................................................................................................. 40

*Table 5.* Logistic regression of repetition count and distance count for lag-2 primed target set distractor trials in Chapter 6. Neither factor was a significant predictor of accuracy. ..................... 62

*Table 6.* Logistic regression of repetition count and distance count for lag-2 primed target set distractor (infrequently and infrequently primed) trials in Chapter 7. Neither factor was a significant predictor of accuracy. .............................................................................. 70

*Table 7.* Logistic regression of time for lag-2 primed target set distractor (*Top*), never-primed target set (*Middle*), and non-studied (*Bottom*) distractor trials in Chapter 6. Time was a significant predictor of accuracy for never-primed target set and non-studied distractor trials only. ........................................................................................................ 75

*Table 8.* Logistic regression of time for lag-2 primed target set vs. non-studied (*Top*), and never-primed target set vs. non-studied target set (*Bottom*) distractor trials in Chapter 6. Time by distractor type was not a significant predictor of accuracy in either test. ......................................................... 78
LIST OF FIGURES

Figure 1. Example trial sequence of the memory training task for Chapter 1. Left panel. Study phase: Participants were shown a series of 30 objects for 3000 ms each, with an inter-stimulus interval of 950 ms, and were instructed to commit these objects to memory. Right panel. Recognition phase: Participants were shown 30 trials containing two objects each and were required to indicate which was previously studied. Both phases were repeated until participants achieved 80% accuracy on two consecutive recognition phases. ................................................................. 19

Figure 2. Example trial sequence of the attention cueing task for Chapter 1. In each experiment participants were instructed to locate the studied object and ignore the concurrently presented non-studied object. The image files for both objects were rotated either clockwise or counterclockwise and participants were required to indicate the direction of rotation of the target (studied object). The response frame was preceded by a cue frame in which a single task-irrelevant object appeared in one of the placeholder boxes. The Stimulus Onset Asynchrony (SOA) was 150 ms (duration of cue frame) plus the ISI (inter-stimulus interval). ...................... 21

Figure 3. Reaction time data for Chapter 1, revealing selective cueing effects. Studied cues produced cueing effects, but not non-studied cues. Error bars in this figure, and all subsequent figures, are within-subject standard errors (Morey, 2008). ................................................................. 23

Figure 4. Example trial sequence of the attention cueing task for Chapter 2. As in Chapter 1, participants were instructed to locate the studied object and ignore the concurrently presented non-studied object. In the response frame, the placeholder boxes surrounding the objects changed colour (one always red, the other always green) and participants were required to indicate the colour of the box surrounding the studied object. As in Chapter 1, the response frame was preceded by a cue frame in which a single task-irrelevant object appeared in one of the placeholder boxes. All SOAs were 150 ms (duration of cue frame) plus the ISI (inter-stimulus interval). ........................................................................................................... 27
Figure 5. Reaction time data for Chapter 2, revealing selective cueing effects. Studied cues produced cueing effects, but not non-studied cues. ................................................................. 29

Figure 6. Example trial sequence of the attention cueing task for Chapter 3. The trial sequence was identical to that in Chapter 2, with the exception of the SOA duration (duration of the cue frame and the ISI): either 150, 250, or 400 ms. ................................................................. 33

Figure 7. Reaction time data for Chapter 3. Inhibition of return is present at long SOAs (400 ms) suggesting that cueing effects in the present experiments were caused by attentional capture rather than volitional attention. ................................................................. 35

Figure 8. Example trial sequence of the attention cueing task for Chapter 4. The trial sequence was identical to that in Chapter 2 with the exception of the cue frame: One task-irrelevant object appeared in each of the placeholder boxes, one studied and one non-studied. ......................... 38

Figure 9. Reaction time data for Chapter 4. The 150 ms SOA cueing effect for studied cue images is replicated using a modified, two-cue paradigm. ................................................................. 40

Figure 10. Trial sequence of the memory training task for Chapter 5. Left panel. Study phase: participants were shown two sets of 15 objects for 3000 ms each, with an inter-stimulus interval of 950 ms, and were instructed to commit these objects to memory. Each set was explicitly designated as separate from the other, and the presentation of each was separated by an interval of time controlled by participants. Right panel. Recollection phase: participants were shown 45 trials containing one object each and are required to indicate whether the object was from the first set of studied objects (set A), the second set (set B) or was not previously studied. Both phases were repeated until participants achieved 90% accuracy on four consecutive recognition phases. ............................................................................. 45

Figure 11. Trial sequence of the rapid serial visual presentation (RSVP) task in Chapter 5. Participants were instructed to use the probe screen to report the identity of either the target set (e.g., set A) or non-target set (e.g., set B) item (counterbalanced across participants) presented
within the RSVP stream on each trial. Two, five, or eight frames prior to the target, two
distractors appeared flanking the RSVP stream. One of the flanking objects was always a non-
studied object, while the other could be an additional non-studied object, a studied object from
the target set, or a studied object from the non-target set.  

*Figure 12.* Accuracy data for Chapter 5, revealing a contingent spatial blink (i.e., impairment of
target detection at lag 2), that occurred only for studied target distractors.

*Figure 13.* Schematic of RSVP distractor conditions and corresponding target manipulations
across Chapters 6 (*left panel*) and 7 (*right panel*). The distractor conditions in the purple boxes
represent our conditions of interest, relative to non-studied distractor conditions, for discerning
between feature priming and voluntary implementation accounts of ACSs.

*Figure 14.* Accuracy data for Chapter 6: a selective spatial blink (impairment of target detection
on lag-2 trials) for trials with a studied target set critical distractor only. This selective spatial
blink was modulated by priming: primed target set items captured attention significantly more
than never primed target set items.

*Figure 15.* The predicted likelihood of accurate vs. inaccurate performance on all lag-2 primed
target set distractor trials given repetition count i.e., the number of times a particular distractor
image has previously appeared as a target (*Left panel*) and distance count i.e., the amount of time
in trials since the last appearance of a particular distractor image as a target (*Right panel*).
Neither factor was found to significantly predict the likelihood of an inaccurate response on a
given trial.

*Figure 16.* Accuracy data for Chapter 7: a selective spatial blink (i.e., impairment of target
detection on lag-2 trials) for trials with a studied target set distractor, whether frequently or
infrequently primed.

*Figure 17.* The predicted likelihood of accurate vs. inaccurate performance on all lag-2 primed
target set distractor trials given repetition count i.e., the number of times a particular distractor
image has previously appeared as a target (Left panel), and distance count i.e., the amount of time in trials since the last appearance of a particular distractor image as a target (Right panel). Neither factor was found to significantly predict the likelihood of capture on a given trial. ....... 69

Figure 18. The predicted likelihood of accurate vs. inaccurate performance on all lag-2 primed target set distractor trials never-primed target set distractor trials, and non-studied distractor trials over time. Time was a significant predictor of accuracy for never-primed and non-studied critical distractor trials; the likelihood of making an accurate response increased over time. ............... 73
INTRODUCTION

Many of us can identify with the phenomenon of attentional capture: Physically salient properties of our visual environments, like the flashing lights on a passing ambulance, draw our attention rapidly and without intention. A large body of scientific work has been devoted to the study of attentional capture: For example, one of the classic findings in attention capture research is that when a salient stimulus unexpectedly appears in the environment, stimuli later appearing in the same location are processed significantly faster than stimuli appearing in other locations, ostensibly because the salient stimulus captured attention to its spatial location (Posner, 1980). Researchers have used this effect as the hallmark of the capture of visual spatial attention (referred to more simply as spatial capture throughout the remainder of this thesis). One of the more provocative discoveries of this capture literature is that the ability for salient stimuli in the environment to capture our attention is constrained by our internal attentional goals, more formally referred to as our Attentional Control Settings (ACSs) (Folk, Remington, & Johnston, 1992; Folk, Remington, & Wright, 1994; Folk & Remington, 1998). In this case, when looking for a red stimulus, the unexpected appearance of any red stimuli will cause spatial capture and, importantly, the appearance of green stimuli will not. Much progress has been made in the past few decades in our understanding of these ACSs, including whether or not the constrained capture effects as described truly do reflect an ability to control the initial capture of spatial attention by goal-matching items rather than, say, a difference in how quickly attention can be disengaged from goal-matching items (Awh, Belopolsky, & Theeuwes, 2012; Belopolsky, Schreij, & Theeuwes, 2009; Theeuwes, 2013). Our understanding of ACSs, though, is continuously developing, particularly when it comes to the role of ACSs in supporting more complex behaviours that are more akin to those used every day, such as looking for a number of unique and specific items rather than an individual feature or object.

Consider the following scenario: You are late for a party and have to pull together the outfit you want to wear from your densely packed closet (i.e., your favourite green pants, that white blouse you have been meaning to wear, the matching beige cardigan, etc.). You might notice that as you open your closet door, only the particular items you need for the night “pop-out” at you from amongst the rest—even that terrible bright yellow dress doesn’t capture your
attention. The goal of finding your outfit requires an ACS for multiple items at the same time, and by necessity, adopting this ACS requires some form of memory to maintain its contents. An ACS of multiple items, however, seems to be at odds with the leading conceptualization of ACSs, which is that we are only able to specify ACSs by at most one or two items maintained in visual working memory, our capacity limited memory system responsible for holding visual information in an online, accessible state (Carlisle & Woodman, 2013; Kumar, Soto, & Humphreys, 2009; Olivers, Meijer, & Theeuwes, 2006; Olivers, Peters, Houtkamp, & Roelfsema, 2011; Soto, Heinke, Humphreys, & Blanco, 2005; Soto et al., 2008; van Moorselaar, Theeuwes, & Olivers, 2014). The first major aim of this thesis will therefore be to evaluate the possibility that ACSs can be specified by multiple items stored in episodic long-term memory—the seemingly capacity unlimited memory system responsible for storing our experiences of past events—thus allowing for increased flexibility in the number of representations that guide capture.

Returning to our scenario, we can highlight another limitation in our understanding of ACSs. Once on your way to the party, you need to be on the lookout for rapidly changing landmarks to avoid getting lost: When you see the river, take a right; then quickly after keep an eye out for the hockey arena on your left; once you pass the arena, make a right on Cardigan street. In the described scenario, you are able to effectively control which landmarks guide your attention (e.g., I’m now looking for the hockey arena instead of the river). Such an ability, however, is counter to a particular modern conceptualization of ACSs espousing our inability to effectively choose the types of things that end up capturing our attention; instead, past experience with stimuli determines whether they will capture our attention in the future (Awh et al., 2012; Belopolsky et al., 2009). The second major aim of this thesis will therefore be to address debate over how it is that ACSs are implemented in the first place: Are we truly able to choose what captures our attention, or are ACSs better explained by a model that relies solely on the role of past experiences of attending to relevant items in guiding capture (i.e., these landmarks only capture your attention if you have successfully located them before)?

Together the studies in this thesis will investigate both the number of representations by which ACSs are specified, and the implementation of these ACSs. Doing so provides advancement to our current conceptualization of ACSs: These ACSs involve more complex
behaviours than evidenced by the current literature and indeed operate as a mechanism of behaviourally relevant, voluntary control over spatial capture.  

VISUAL SPATIAL ATTENTION

Before reviewing the current state of ACS research, it is worth taking a step back to evaluate the role that attention, broadly speaking, plays in human visual information processing. At any given moment in time, our visual environments provide us with an overwhelming amount of sensory information such that we are unable to fully make sense of (i.e., perceive) it all at once. This limitation is a consequence of the structure of the visual system: Perception is a hierarchical process beginning with low-level feature-based representations (e.g., orientations, colours, shapes, etc.) that are later combined into higher-level representations of objects (e.g., faces, buildings, animals, etc.). This integration necessitates competition amongst low-level representations for representation at the higher-level, such as when two objects overlap in space. In order to resolve this competition, the visual system employs a number of attentional mechanisms that bias the outcome of the competition in, hopefully, behaviourally relevant ways, including object-based, feature-based, and spatial-based mechanisms (Cavanagh, 2011; Desimone & Duncan, 1995; Petersen & Posner, 2012; Posner & Cohen, 1984; Yantis, 2005). Our everyday experience of being able to focus on some parts of the visual environment, while ignoring others, is a consequence of selection: attentional mechanisms bias the processing of particular properties or stimuli over others, and in the case of the work in this thesis, the prioritization of stimuli in particular spatial locations over stimuli in other locations.

In investigations of the biases that guide perceptual competition, researchers classically distinguish between voluntary, goal-directed attentional mechanisms, and stimulus-driven attentional mechanisms. Voluntary mechanisms, such as looking for traffic before crossing the street, or searching for a friend in a crowd, are viewed as relatively active and volitional in nature, and guide selection on a slower timecourse as compared to stimulus-driven mechanisms.

---

1 Sections of this Introduction and the General Discussion have been adapted from one published paper (Giammarco, Paoletti, Guild, & Al-Aidroos, 2016) and one draft of a paper in preparation for submission (Giammarco, Hryciw, & Al-Aidroos, in prep). Sections of the Methods are direct excerpts from these papers; Chapters 1 to 5 (Giammarco et al., 2016) and Chapters 6 and 7 (Giammarco et al., in prep).
In contrast, stimulus-driven mechanisms operate on the basis of the saliency of stimulus properties (Carrasco, 2011; Chun, Golomb, & Turk-Browne, 2011; Corbetta & Shulman, 2002; Kastner & Pinsk, 2004; Ruz & Lupiáñez, 2002). Saliency is defined by the uniqueness of a stimulus’ visual attributes relative to stimuli in the same environment (e.g., ambulance lights tend to capture our attention because they are brighter than other stimuli on the road; Itti & Koch, 2000), and the prioritized selection of salient stimuli tends to confer behavioural advantages (i.e., get out of the way an ambulance is coming!; Corbetta & Shulman, 2002; Kastner & Pinsk, 2004). This form of selection is viewed as relatively rapid, passive, and transitory in nature, occurring in response to unexpected events and without the need or use for voluntary control. While attentional capture is typically characterized as the result of stimulus-driven selection, contingent capture (i.e., capture constrained by ACSs), invokes a role for voluntary mechanisms in controlling the types of things that capture attention (Al-Aidroos, Harrison, & Pratt, 2010; Bacon & Egeth, 1994; Folk et al., 1992, 1994; Folk & Remington, 1998; Folk, Leber, & Egeth, 2002; Ruz & Lupiáñez, 2002; Wyble, Folk, & Potter, 2013). Thus ACSs cannot be explained as either a stimulus-driven or voluntary attentional mechanism alone; instead, ACSs involve the use of goal-based factors to constrain the impact of stimulus-based factors (i.e., saliency) to then determine selection. The notion that voluntary control over selection can guide stimulus-driven capture is supported by recent evidence suggesting voluntary, goal-driven and stimulus-driven mechanisms rarely operate in isolation: What is considered salient in the first place is often determined by our voluntary goals (Corbetta & Shulman, 2002; Ruz & Lupiáñez, 2002). In regards to spatial attention in particular, ACSs involve the use of voluntary goals to guide stimulus-driven selection in such ways that the spatial locations containing stimuli that match our goals (i.e., the clothes you plan on wearing to the party) end up being selected for, despite the physical saliency of stimuli in other locations that do not match these goals (i.e., your bright yellow dress).
ATTENTIONAL CONTROL SETTINGS

To appreciate the goals of this thesis in further developing our understanding of ACSs, it is helpful to examine the current state of the field with regard to the evidence supporting the existence of ACSs and what we currently know about the nature of ACSs. ACSs determine the objects or properties that guide capture, in turn allowing for enhanced processing of goal-matching information (Folk et al., 1992, 1994; Folk & Remington, 1998; Folk, Leber, & Egeth, 2008). Evidence for the existence of ACSs has often come from the attention cueing paradigm, a commonly used measure of stimulus-driven spatial capture (i.e., the selection of particular spatial locations over others) (Posner, 1980). In this task participants are required to report the appearance of a pre-defined target stimulus that appears in one of several spatial locations. Shortly before the appearance of the target (i.e., about 150 ms) a non-predictive cue stimulus briefly appears in one of the possible target locations; participants are told to ignore the cue as it is not relevant in locating the subsequent target. Despite this, the task irrelevant cue still impacts performance. In a typical cueing task, target detection tends to be faster on trials in which the cue stimulus and target were in the same spatial location than differing spatial locations; this difference in reaction time is referred to as a cueing effect. This suggests that the cue stimulus induced a shift in spatial selection towards its location, resulting in faster processing of cued targets (i.e., targets in the same location; Folk et al., 1992; Folk et al., 1994; Folk & Remington, 1998).

Although cueing effects in this task are triggered by an external stimulus (i.e., the cue), these cueing effects have been found to be contingent on an observer’s internal goals or their ACSs (Al-Aidroos, Harrison, & Pratt, 2010; Bacon & Egeth, 1994; Folk et al., 1992, 1994; Folk & Remington, 1998; Folk, Leber, & Egeth, 2002; Gibson & Amelio, 2000; Ruz & Lupiáñez, 2002; Wyble et al., 2013). Specifically, salient but irrelevant cue stimuli in a cueing task only produce cueing effects (i.e., capture attention) when they contain the target-defining feature of the task. For example, when participants are instructed to locate green targets, cueing effects are selectively present on trials in which the cue stimulus is also green, and absent on trials in which the cue stimulus is another colour (Folk & Remington, 1998). This contingent capture has been replicated numerous times using feature-based targets such as colours, shapes, and onsets (Al-Aidroos et al., 2010; Ansorge, Horstmann, & Carbone, 2005; Atchley, Kramer, & Hillstrom,
Selective cueing effects are often taken to support the *contingent capture* model (Al-Aidroos et al., 2010; Ansorge et al., 2005; Ansorge, Kiss, Worschech, & Eimer, 2011; Atchley et al., 2000; Chen & Mordkoff, 2007; Eimer, Kiss, Press, & Sauter, 2009; Eimer & Kiss, 2008; Folk et al., 1992, 1994; Folk & Remington, 1998; Folk et al., 2002; Folk et al., 2008; Goodhew et al., 2014; Kiss, Grubert, & Eimer, 2013; Kiss, Jolicœur, Dell’Acqua, & Eimer, 2008; Wyble et al., 2013) in which top-down control (via ACSs) selectively guides the types of stimuli that bias spatial selection so that only ACS-matching items actually capture attention and not non-matching stimuli. In the current literature, the contingent capture model is contrasted with a *rapid disengagement* model (Awh et al., 2012; Belopolsky et al., 2009; Theeuwes, 2013). According to the rapid disengagement model, ACSs do not prevent non-matching stimuli from capturing attention, rather they speed the subsequent disengagement of attention from non-matching stimuli, and thus capture by non-matching stimuli is resolved before cueing effects can be measured. Although settling the debate over the timing of control by ACSs is valuable for our larger, theoretical understanding of the timecourse of top-down control, the effects of control can still be observed and evaluated in the interim. While the debate over the timing of control is ongoing, I will use the contingent capture terminology for the present thesis and provide a more thorough discussion of the debate in the General Discussion.

**AIM I: THE STORAGE OF ATTENTIONAL CONTROL SET ITEMS**

In order for an ACS to exist, a memory system is required for maintaining the target items that are a part of someone’s selection goals. A multitude of research suggests that this maintenance is accomplished using visual working memory (Carlisle & Woodman, 2013; Kumar et al., 2009; Olivers et al., 2006; Olivers et al., 2011; Soto et al., 2005; Soto et al., 2008; van Moorselaar et al., 2014). Visual working memory is a capacity limited memory system responsible for the online maintenance, as well as manipulation and processing of, visual stimuli,
with the purpose of supporting current behaviours, such as remembering what cars are behind you after checking your rear-view mirror (Luck & Hollingworth, 2008; Oberauer, 2002). While visual working memory is severely capacity limited and the representations stored within are fairly transitory (working memory is only capable of maintaining about four objects worth of visual information for short periods of time; Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Wilken & Ma, 2004), it is a likely candidate memory system for encoding and storing the contents of ACSs (Adamo, Pun, Pratt, & Ferber, 2008; Moore & Weissman, 2010; Olivers & Eimer, 2011; Olivers et al., 2011; Soto et al., 2008). As it is, the majority of work investigating ACSs has used simple, single-item targets, such as colours or shapes, the most complex (Adamo et al., 2008; Moore & Weissman, 2010, 2014) of which could easily be stored within visual working memory. In fact, in other tasks measuring attentional capture, namely the additional singleton visual search paradigm, the contents of visual working memory have been shown to directly impact the attentional effects of distractors. In this task, participants are given both a memory item for end-of-trial recall, and a search target for detection in a stimulus array; in the search array, a singleton distractor (i.e., a unique stimulus that pops-out against a homogenous search array) can be present or absent, and attentional capture by singleton distractors is measured as slowed target detection on distractor present trials (i.e., distractor costs). Importantly, singleton distractors that resemble the item stored in visual working memory (e.g., share the same colour, or shape) produce larger distractor costs than distractors that do not (Olivers & Eimer, 2011; Olivers et al., 2011; Olivers et al., 2006; Soto et al., 2008). As well, the electrophysiological signature of visual working memory (the contralateral delay activity) can be observed after participants are initially told the identity of their search target, suggesting that the attributes of their search target or search goal are at least sometimes maintained in working memory (Carlisle & Woodman, 2011).

Furthermore, evidence suggests that there are multiple states of representation in visual working memory: active and accessory (Olivers et al., 2011). A single item can reside in the active state at any given moment, and it is only this active item that acts to guide attentional capture while the remaining items in the accessory state do not (Olivers et al., 2011; Peters, Goebel, & Roelfsema, 2009; van Moorselaar et al., 2014). In the aforementioned visual search task, when the visual search target is variable (i.e., changes from trial-to-trial), distractor costs
are not modulated by memory-matching distractors; however, when the search target is held constant, the memory-matching distractor does capture attention, producing larger distractor costs. This dissociation has been taken as evidence for multiple functional states in visual working memory: When search targets are continuously updated, they remain in an active state, and since no other stimuli can simultaneously exist in this state, things like the to-be-remembered item of every trial likely exist in an accessory state, thus explaining why, when the distractor matches this memory item, it does not induce greater distractor costs. When search targets remain constant, though, they can be offloaded into an accessory state while the memory item can move into the active state, and therefore guide capture (Olivers et al., 2011). Thus, holding an item in the active state of visual working memory has been likened to implementing an ACS for that item.

While the active vs. accessory model of visual working memory and its role in supporting ACSs is under debate (Hollingworth & Beck, 2016), the evidence in support of visual working memory as responsible for the storage of ACS contents is compelling. However, this model paints a fairly constrained picture of ACSs: Only several items — potentially even only a single active item — at a time may be able to control stimulus-driven attention. As in the aforementioned scenario, however, a task such as putting together an outfit comprising a number of different items questions whether we may in fact be able to adopt more complex ACSs, with complexity being defined here as quantity. Our visual working memory storage capacity is severely limited, and the representations contained within are short-lived, active only when they are in our current state of focus (Oberauer, 2002); thus more complex ACSs, as in the described scenario, cannot be adequately maintained in visual working memory. The representations supporting our memory for the items in this example are more akin to the types of representations stored in episodic LTM: our essentially capacity-unlimited memory for past experiences that are tied to particular spatial/temporal contexts and long-lasting in nature, available for recall even when not held online in visual working memory. Might episodic LTM therefore also be a feasible memory system for the storage of ACS items? Indeed, support for the notion that LTM can guide selection comes from work using the previously described additional singleton paradigm: If the search target is repeated across numerous trials, the contralateral delay activity present when that target is first introduced (i.e., on the first repetition of that particular
target) disappears with target repetitions, yet search times become increasingly efficient (i.e., the target is successfully used to guide search), suggesting ACS contents can be transferred out of working memory and into LTM, although the particular type of LTM is left unspecified (Carlisle, Arita, Pardo, & Woodman, 2011; Woodman, Carlisle, & Reinhart, 2013). Of note, Woodman et al. (2013) focuses on the memory systems used for the storage of visual search target templates and guidance of selection in task performance, and it is unclear from their work if maintaining a search target in LTM leads to the contingent capture effects associated with ACSs.

Direct evidence that the contents of ACSs can be specified based on LTM representations comes from two recent studies; however, these studies examined semantic LTM rather than episodic LTM. Goodhew et al. (2014) used a cueing paradigm to demonstrate that when participants look for a red-coloured target, the word “red” presented as a cue will capture attention. Similarly, Wyble et al. (2013) demonstrated contingent attentional capture when targets were defined by conceptual category membership; for example, sports equipment, or kitchen furniture. Here they used a rapid serial visual presentation (RSVP) task to measure capture, in which a stream of central sequential stimuli was presented and participants were required to identify a categorically defined target within the stream amidst the appearance of flanking distractors at randomized intervals preceding the target. When distractors were from the target category, they produced an attentional blink effect (i.e., a reduced likelihood of detecting the target), but not when they did not match the target category, suggesting attentional capture by distractors was contingent on the observers’ ACSs. Given that the conceptual representations used to define category membership do not resemble the type of visual-object representations that can be stored in visual working memory, it is likely that ACSs were defined based on semantic representations in LTM. While the knowledge that the contents of ACSs can be stored in semantic memory adds considerable value to the function of attentional capture, allowing us to orient towards categories of relevant items (e.g., all food stimuli when hungry) rather than single visual objects or features, this form of memory is constrained. In particular, the acquisition of semantic knowledge often occurs on a relatively slow timescale, and semantic memories tend to be domain-general and relatively immune to changes or adaptations over time (Holdstock, Mayes, Isaac, Gong, & Roberts, 2002). On the other hand, episodic LTM contains a vast amount
of contextually rich information about our lives and experiences that may be directly relevant to our behaviour at any given moment (e.g., figuring out where you left your keys yesterday as you are rushing to get to work on time), while allowing for rapid relational associations between large amounts of even arbitrary information (Henke, 2010) (e.g., semantically unrelated experiences of putting down your keys while checking a phone notification can be associated with the same time and place and thus bound in a single episodic memory). Given both the importance of episodic long-term memories for guiding behaviour, and emerging evidence suggesting LTM in general, and episodic memory more specifically, is a potential candidate memory system for supporting ACSs, the first major focus of this thesis will be to evaluate whether ACSs can in fact be defined by representations stored in episodic LTM. Determining whether this is the case is a relevant and important question for our understanding of ACSs more generally, and would allow for a greater degree of flexibility in the types of representations that are utilized for the top-down control of spatial capture. To preview the results, Chapters 1 to 5 provide multiple pieces of evidence across multiple paradigms in favour of the existence and use of episodic LTM ACSs.

AIM II: THE IMPLEMENTATION OF ACSs

Another hot topic in the attentional capture literature is how ACSs are implemented in the first place. A commonly held idea is that ACSs are established through the use of top-down control; a volitional goal (e.g., look for green) is implemented, and this goal will subsequently bias perceptual representations towards goal-matching features or stimuli (Al-Aidroos et al., 2010; Bacon & Egeth, 1994; Folk et al., 1992, 1994; Folk & Remington, 1998; Folk, Leber, & Egeth, 2002; Gibson & Amelio, 2000; Ruz & Lupiáñez, 2002; Wyble et al., 2013). The notion that we are essentially able to decide what information our ACSs are composed of, though, has recently been challenged by the idea that feature priming of target stimuli is instead critical in producing the selective capture results used to support the contingent capture hypothesis. Rather than choosing, in a sense, what will capture attention via the implementation of an internal selection goal from the onset of a task (e.g., “your target is green, look for green…greeeeeeeen”), the feature priming camp argues that capture is instead modulated by recent experiences of
selecting and responding to targets (Awh et al., 2012; Belopolsky et al., 2009). According to the priming literature, selecting a particular feature or object automatically influences subsequent processing of that feature or object through bottom-up priming (Hillstrom, 2000; Kristjansson, Wang, & Nakayama, 2002; Maljkovic & Nakayama, 1994, 1996; Maljkovic & Nakayama, 2010; Moore & Weissman, 2010; Müller, Heller, & Ziegler, 1995; Olivers & Humphreys, 2003; Pinto, Olivers, & Theeuwes, 2005). Thus, target-defining features are incidentally primed through the active selection of a target, and it is this priming that subsequently biases perceptual selection towards certain stimuli (i.e., those that match or contain the target-defining feature that has been primed) rather than others (i.e., those that do not match or contain the target-defining feature and therefore have not been primed).

Perhaps contingent capture is simply a consequence of selecting primed features or objects? In fact, a feature priming account can explain ACS-selective capture effects in standard capture tasks, such as the spatial cueing paradigm. To re-cap, in the cueing task, although the cue is irrelevant, it can contain properties that capture attention. In support of a top-down, contingent capture model, standard cueing effects have been shown to be selectively present on trials in which the cue matches participants’ ACS (i.e., contains the target-defining feature), and absent when the cue does not match participants’ ACS. However, in traditional contingent capture cueing tasks, participants are given a single, unchanging target item or feature, such that participants continuously select the same target throughout an entire experimental session. Following active selection of a consistent target – a requirement for successful task performance – this target (i.e., its critical, target-defining feature) is ostensibly primed. This priming is passive in nature, that is, it occurs unintentionally and without control, but acts automatically to influence selection in future trials: Target-matching stimuli, including task-irrelevant cue stimuli, may capture attention because they have been primed, not because they are part of a voluntary ACS.

Given past work evaluating the role of priming on selection, this potential model seems to be justified. Specifically, Maljkovic and Nakayama (1994, 1996, 2000) found priming effects in a speeded visual search task for “pop-out” targets (i.e., heterogeneous targets against a homogeneous array of distractors): Search times were significantly faster when the target-defining feature or target location was repeated from one trial to the next, rather than when
switched, even when participants knew the target feature in advance. Similar effects of the impact of inter-trial feature priming on behaviour have also been reported in the additional singleton visual search paradigm: Singleton distractor costs to speeded search performance, the measure of capture, were modulated by the relationship between the target on trial \( n-1 \) and distractor on trial \( n \). Distractor costs were attenuated when the target-defining feature was repeated in the distractor, while enhanced when the target-defining feature switched, suggesting feature priming was critical in modulating the strength of capture by singleton distractors (Pinto et al., 2005). Notably, these inter-trial priming effects are relatively transient in nature, exerting maximum influence within several trials.

In order to directly test the possibility that inter-trial feature priming can better account for selective capture than the voluntary implementation of an ACS, Belopolsky et al. (2009) developed a modified version of the spatial cueing task used to measure contingent capture. Critically, the target-defining feature changed on a trial-by-trial basis. This manipulation allowed for a clear, falsifiable prediction: If capture is reliant on the top-down establishment of ACSs, then capture should be selective to cues possessing the target-defining feature for the current trial, and be unrelated to the target-defining feature on the preceding trial (i.e., in particular if the target-defining feature switched). If contingent capture is instead the result of inter-trial feature priming, the opposite should be true. Capture should be dependent on the target-defining feature from the previous trial rather than the current trial. The evidence supported the latter prediction: When target identity randomly switched between a colour singleton and an abrupt onset stimulus (i.e., the appearance of an additional stimulus in one of several potential target locations), both target matching and non-matching cues captured attention. Notably, when the target on a given trial was unknown in advance, the target on trial \( n-1 \) significantly impacted whether or not the cue on trial \( n \) would capture attention (Experiments 1 & 2; although see Folk & Remington, 1998: here similar results were obtained, but taken to conclude that inter-trial priming only plays a role when effective top-down control is lacking). Together with the known effects of feature-priming in visual search, these data provide converging evidence that inter-trial feature priming can indeed better account for the contingent capture results often taken to support the voluntary implementation of ACSs.
Despite these findings, the depth of work exploring the role of a feature priming account of selective capture is relatively lacking, with only a handful of studies assessing the issue (Belopolsky et al., 2009; Folk & Remington, 1998; Lien, Ruthruff, & Johnston, 2010; Lien, Ruthruff, & Naylor, 2014; Pinto et al., 2005; Theeuwes, 2013). As well, the evidence in favour of the feature priming account is contentious and inconsistent (Eimer & Kiss, 2008; Irons & Remington, 2013; Lamy & Kristjánsson, 2013). One important factor to note across these priming studies is the use of single target ACSs (e.g., on a given trial, look for one specific shape, colour, or onset). This presents an important confound within these studies: When testing inter-trial feature priming by changing the target-defining feature on a trial-by-trial basis, every time that the primed ACS changes from one trial to the next, participants are also asked to adopt a new ACS. It is possible that one trial is not enough time to switch and effectively implement a new ACS, and, thus, inefficient top-down switching may be the reason why capture effects are non-selective, rather than ACSs being implemented through priming (Lien et al., 2010; Lien et al., 2014). Thus in the current state of the ACS field, the role of inter-trial priming is confounded with the potential impact of ACS-switching costs. In order to isolate the role of inter-trial feature priming, ACS-switching must be controlled for, yet such control is not possible with single feature or single item ACSs.

The second major aim of this thesis, therefore, addresses the confound present in the feature-primed ACS literature by taking advantage of the findings from Chapters 1 to 5 (i.e., evidence supporting the existence and use of episodic LTM ACSs) and providing participants with multiple-item target ACSs, a form of target set that allows for the features of target items to change on a trial-by-trial basis without invoking the need to change their ACS. Across the studies in Chapters 6 and 7, these episodic LTM ACSs are used to directly test a feature priming account of selective capture in an RSVP capture paradigm; while the potential costs of switching ACSs is controlled for using episodic LTM ACSs, the same target stimuli are repeatedly selected over the course of an experiment, and thus the accumulation of priming instances over time may still bias selection such that contingent capture is driven by feature priming. To preview the results, both the presence and frequency of target priming had limited contributions to the implementation of ACSs, elucidating the important debate over the implementation of ACSs, and providing evidence for the role of flexible top-down control over the implementation of ACSs.
THE CURRENT THESIS

The consideration of more complex behaviours, such as looking for many different unique items at the same time, and being able to effectively choose what items will guide attentional capture in the first place, is fairly restricted under our current view of ACSs. The major aims of this thesis—to determine whether ACSs can be specified by episodic LTM representations and to address the debate over the role of top-down control vs. feature priming in the implementation of these ACSs—work to re-conceptualize our understanding of ACSs as a more flexible form of control over selection than demonstrated by the literature thus far. Across all studies, participants first memorized a set of 16 to 30 unique, naturalistic images of everyday objects, a feat that is likely accomplished using episodic long-term memory given that 16 objects far exceeds the capacity of visual working memory (Brady, Konkle, Alvarez, & Oliva, 2008). These items were then designated as the targets in subsequent tasks designed to measure attentional capture. To determine whether individuals can implement ACSs based on representations stored in episodic LTM, the studies in Chapters 1 to 4 used modified cueing paradigms to test for contingent capture. Specifically, the presence of two well-established markers of capture that are measurable in a standard, single cue paradigm—cueing effects (Chapters 1 and 2) and inhibition of return (Chapter 3)—were found to be modulated by episodic LTM ACS-matching cues: The presence of inhibition of return in Chapter 3 confirms that the cueing effects in Chapters 1 and 2 are unlikely to be driven by voluntary attention. In Chapter 4, a two-cue paradigm was used to provide converging evidence in support of the findings from Chapters 1 to 3. Chapter 5 introduces a modified version of the RSVP paradigm used to measure capture in order to extend the generalizability of the findings from previous chapters, and, critically, to demonstrate that episodic LTM ACSs can be supported by recollection, rather than familiarity memory, highlighting the detail-rich, context-specific nature of the representations that ACSs can be based on.

The role of flexibility in the implementation of ACSs is then addressed in Chapters 6 and 7. Specifically, the role of top-down control vs. feature priming in this implementation is directly tested. The findings from Chapters 1 to 5, that ACSs can be specified by multiple unique items,
allow for control over a major confound of previous work in the top-down vs. priming debate (i.e., changes in feature priming have conventionally also required changes in participants’ ACS), since target and distractor features in the current paradigm can change on a trial-by-trial basis, while participants’ ACS remains the same. Using a modified RSVP task to measure capture, these studies reveal that both the presence of priming (Chapter 6) and the frequency of priming (Chapter 7) are insufficient for the implementation of ACSs. Instead, the data adhere more strongly with a flexible control account of ACSs in which ACSs are implemented efficiently through the use of top-down task goals, without invoking the need for prior experience selecting target items.

Together these studies advance our understanding of ACSs, highlighting the flexibility in both the representations by which ACSs are specified (i.e., capacity unlimited, context-rich episodic LTM representations) and in the implementation of these ACSs in the first place (i.e., the ability to simply choose what captures our attention). The use of ACSs to guide capture reflects a highly sophisticated attentional mechanism, and developing our understanding of the nature of these ACSs allows us to build more comprehensive models of how we are able to flexibly guide our limited processing resources in relevant and important ways.

CHAPTER 1: SELECTIVE CUEING EFFECTS FOR EPISODIC LTM
ACS-MATCHING CUES

A significant amount of research has explored the interactions between episodic LTM and attentional processes (Chun & Jiang, 2003; Chun & Turk-Browne, 2007; Rosen, Stern, Michalka, Devaney, & Somers, 2015; Shiffrin & Schneider, 1977; Summerfield, Lepsien, Gitelman, Mesulam, & Nobre, 2006). For example, individuals are faster at navigating through familiar visual environments (Chun & Jiang, 2003; Rosen et al., 2015; Summerfield et al., 2006), visual search is more efficient when guided by memory-based cues rather than low-level visuo-spatial cues (Chun & Turk-Browne, 2007; Olson, Chun, & Allison, 2001; Peterson & Kramer, 2001; Summerfield et al., 2006), and activity has been observed in brain regions associated with
LTM, such as the hippocampus and other medial temporal areas, during attention tasks (Chun & Phelps, 1999; Chun & Turk-Browne, 2007; Manns & Squire, 2001), while activity in brain regions associated with attention, such as the frontal-parietal network, are active during episodic memory tasks (Chun & Turk-Browne, 2007; Ciaramelli, Grady, & Moscovitch, 2008; Summerfield et al., 2006). It thus seems intuitive to think that LTM should guide attentional capture. However, the focus of the aforementioned work has been solely on the guidance of voluntary or goal-directed attention, whereas the focus of the current work is instead on the top-down control of attentional capture.

There have, however, been numerous demonstrations that episodic memory plays a role in guiding what ACS an observer will adopt, and when. For example, when participants are performing a visual search task in which their target is a circle among diamond distractors, it is equally effective to adopt a broad ACS for shape singletons (i.e., a unique shape that pops-out against a homogenous shape array), or a more narrow feature-based ACS for circles (i.e., the particular target-defining feature). The choice to use a broad versus narrow ACS is significantly influenced by participants’ activity in a previous task: If they previously completed a task that required the use of a feature-specific target like a circle (Leber & Egeth, 2006), then they are more likely to adopt a feature-based ACS in the subsequent visual search task, as evidenced by greater distractor costs on trials with circle distractors, and this effect is fairly durable, lasting for more than a week (Leber, Kawahara, & Gabari, 2009). Interestingly, LTM ceases to guide attention in this way when either the contextual details required to support episodic representations are disturbed, such as changes to the environment of previously studied search array images (Cosman & Vecera, 2013a), or when individuals have sustained damage to the medial temporal lobe, the neural substrate of episodic memory (Cosman & Vecera, 2013b). Thus, there is strong evidence to suggest that LTM guides what ACS an observer will adopt and when. It remains an open question, though, whether this memory system is also capable of representing the contents of ACSs.

Stimulus-driven capture occurs on a fairly rapid timescale in comparison to other mechanisms of attention (i.e., voluntary attention): It onsets faster and terminates as quickly as 200-300 ms after attention capturing stimuli onset (Lupianez, 2010; Posner, Rafal, Choate, & Vaughan, 2007). As well, it is rather effortless in nature, occurring involuntarily. Episodic LTM
retrieval, on the other hand, has traditionally been characterized as a relatively slow and effortful cognitive process (Gabrieli, Fleischman, Keane, Reminger, & Morrell, 1995; Gazzaniga, Heatherton, Halerpn, & Heine, 2010). Thus, it might be surprising to think of episodic LTM as capable of guiding fast, reflexive attentional capture. This traditional view of episodic LTM, though, has recently been replaced by a two-stage model of recollection, in which an initial rapid, unconscious, and obligatory stage is followed by a slower, explicit, and effortful stage (with this second stage akin to traditional models of recollection) (Moscovitch, 2008). In line with the traditional model, this initial rapid and unconscious (Degonda et al., 2005; Guild, Cripps, Anderson, & Al-Aidroos, 2014; Moscovitch, 2008) retrieval stage is supported by activity in the neural substrate of episodic retrieval: the hippocampus (Hannula & Ranganath, 2009; Hannula, Ryan, Tranel, & Cohen, 2007; Hannula, Tranel, & Cohen, 2006). The existence of an initial rapid stage in episodic retrieval is further supported by work documenting deficits in implicit context-specific learning, despite typical perceptual learning (Chun & Phelps, 1999), in individuals with hippocampal damage. As well, hippocampal activity is predictive of implicit episodic retrieval, even in the absence of explicit retrieval (Hannula & Ranganath, 2009), and viewing times are biased towards faces embedded in scenes that they were previously associated with, despite the concurrent presence of equally familiar faces, as early as 1000 ms prior to explicit retrieval (Hannula et al., 2007). Rapid, unconscious episodic retrieval has also been shown to guide performance in speeded visual search tasks (Guild et al., 2014). Thus a multitude of evidence in the episodic retrieval literature supports the notion that retrieval can in fact occur on a rapid timescale, perhaps a timescale rapid enough to support the top-down control of stimulus-driven attention.

The purpose of Chapter 1 was to determine whether or not this is the case: whether ACSs can be specified based on episodic LTM representations in a traditional attentional capture task i.e., a modified cueing task. When an observer is looking for any one of a large number of specific visual objects, will the sudden appearance of one of those objects capture visual spatial attention, but not the sudden appearance of other stimuli? In the experiment presented in this Chapter, participants memorized a set of 30 complex, naturalistic images of everyday objects, a feat likely accomplished using episodic long-term memory given the capacity limitations of visual working memory. These studied images were then designated as targets in an attention
cueing task. In the cueing task, two potential targets appeared simultaneously on each trial, one of which was from the studied list. Participants were required to locate the studied object; this behavioural goal ought to guide them to establish an ACS specific to the studied objects. On each trial, an irrelevant cue appeared beforehand either in the same location as the target (a cued trial) or in the other location (an uncued trial). This cue was either a different studied image than the target or a non-studied image. If episodic LTM representations can be used to specify ACSs, then a contingent cueing effect should be observed: Shorter response latencies for cued vs. uncued trials, only when cues are studied images.

METHODS

PARTICIPANTS

Twenty-eight undergraduate students (mean age of 17.96 years) at the University of Guelph received partial course credit as compensation for their participation. Participants with error rates at 25% or higher (11 of 28 participants) were excluded from the analyses, resulting in a sample of 17 participants. I selected a high threshold for inclusion post-hoc, after observing participants’ difficulty with the task (see error rates below); subsequent experiments designed to be easier use a lower threshold (i.e., 15%). Of note, including all participants did not alter the pattern of statistically significant effects reported below, and the high rate of errors is an issue addressed in Chapter 2. All participants reported having normal colour vision and normal, or corrected-to-normal, visual acuity.

APPARATUS, STIMULI, AND PROCEDURE

The experiment was conducted on a 17 inch iMac monitor with a 1680 × 1050 resolution LCD display at a screen refresh rate of 60 Hz, and responses were made on a standard keyboard. Object images were selected from Brady et al. (2008). Target images for the training and cueing tasks included images from many categories, including animals, food, toys, vehicles, clothing, appliances, and tools.

MEMORY TRAINING TASK.

The experimental session began with a memory task during which participants studied 30 images of objects that were randomly selected from a pool of 130 images (see Figure 1). Each
image was presented one at a time for 3000 ms, with an inter-stimulus interval of 950 ms. Participants were instructed to memorize the images. Images subtended approximately 3° of visual angle in width and height, and were presented in the centre of the display on a white background. After viewing all images, participants’ memory was assessed using a two-item forced-choice recognition test. Specifically, each studied image was re-presented, now paired with a non-studied image randomly selected from a pool of 2100 images. Participants were asked to determine which image in the pair was the previously studied item. Image pairs were presented until participants made their response, with an interstimulus interval of 500 ms. Participants repeated the memory training phase until achieving 80% accuracy or greater on the recognition test on two consecutive blocks.

Figure 1. Example trial sequence of the memory training task for Chapter 1. **Left panel.** Study phase: Participants were shown a series of 30 objects for 3000 ms each, with an inter-stimulus interval of 950 ms, and were instructed to commit these objects to memory. **Right panel.** Recognition phase: Participants were shown 30 trials containing two objects each and were required to indicate which was previously studied. Both phases were repeated until participants achieved 80% accuracy on two consecutive recognition phases.
ATTENTION CUEING TASK.

Following memory training, participants completed an attention cueing task (see Figure 2). I used this task to evaluate whether participants could adopt an ACS for the specific targets studied during training. Each trial of the task began with a fixation frame, a black central fixation point, radius 0.07°, and two placeholder boxes, black square outlines 7.5 × 7.5°, one located 5° to the left of fixation and the other 5° to the right. The fixation frame was presented for 500 ms, after which a 5 × 5° cue image was added to the display inside one of the two placeholder boxes. The cue was either one of the studied images, or was a non-studied image selected randomly from the same 130-image pool as the studied images. Accordingly, the cue either matched the participants’ ACSs (studied cue), or not (non-studied cue). Stimuli for both cue types were selected from the same pool of images so that, across participants, each image was equally likely to be a studied or non-studied cue; thus, any differences in cueing effects between cue types can only be attributed to participants’ internal experiences and attentional goals, and not low-level features of the images or semantic properties of the objects. Participants were told that the cue was irrelevant to their task, and should be ignored.

Task: Find the studied item

![Diagram of the task](image-url)
Figure 2. Example trial sequence of the attention cueing task for Chapter 1. In each experiment participants were instructed to locate the studied object and ignore the concurrently presented non-studied object. The image files for both objects were rotated either clockwise or counter-clockwise and participants were required to indicate the direction of rotation of the target (studied object). The response frame was preceded by a cue frame in which a single task-irrelevant object appeared in one of the placeholder boxes. The Stimulus Onset Asynchrony (SOA) was 150 ms (duration of cue frame) plus the ISI (inter-stimulus interval).

The cue was visible for 100 ms, followed by the fixation frame for 50 ms, after which two $5 \times 5^\circ$ target images were added to the display, one centred in each placeholder box. One target was a studied image and the other a non-studied image, and each was randomly rotated either $25^\circ$ degrees clockwise or counter-clockwise (i.e., for studied images, rotation was relative to the original presentation during training). Additionally, on every trial, the image selected for the cue was always different than the two images presented in the target frame to avoid priming effects. Participants were required to locate the studied image and report the direction of its rotation by pressing the up or down cursor on the keyboard, while ignoring the non-studied image. I had participants report a feature of the target (orientation) that always had a different value in target images than in cues (i.e., cue images were always in their original orientation, whereas targets were always rotated) to limit the potential for cue-driven response biases. The non-studied image in the target frame was randomly selected without replacement from the pool of 2,100 images. After 500 ms all stimuli were removed from the display. Participants were encouraged to be as fast and accurate as possible in their responses; if a response was not made within 2,500 ms of target onset, then the trial timed out and the next one began. For incorrect or response-absent trials, a 500 Hz error tone was presented for 50 ms.

The cueing task consisted of four blocks of 80 trials each, with trials presented in a random order. All levels of cue type (studied cue vs. non-studied cue), cue location (left vs. right placeholder), and target location (left vs. right placeholder) were fully crossed within participants, and each unique combination of conditions occurred with equal frequency. Notably, because the cue appeared in either spatial location with equal likelihood, the location of the cue was not predictive of the location of the target image, and therefore could not be used to improve task performance. For the analyses that follow, cue location and target location were combined to create a single factor based on whether the cue and target had the same location (cued) or not.
(uncued), resulting in a fully crossed 2 (Cue Type: studied cue vs. non-studied cue) × 2 (Cueing: cued vs. uncued) within-subjects design.

RESULTS

Recursive trimming (van Selst & Jolicoeur, 1994) was used on the reaction time (RT) data to eliminate outliers. To do this, the mean and standard deviation for each condition was calculated and trials for which the participant’s RTs were greater than three standard deviations from the condition mean were eliminated. The mean and standard deviation were then re-calculated with the outliers removed and this process was repeated until no outlier trials were detected. Trimming resulted in the removal of 1.75% of trials on average. Error trials (18.44% of all trials) were also excluded from RT analyses.

MEMORY TRAINING TASK
Participants took an average of 2.18 blocks to complete memory training: Out of the 17 participants, only one required a third block of training, while the remainder required two (the minimum). The average accuracy (proportion of correct target recognition trials) across participants and blocks was 99.24% suggesting that participants were able to effectively commit the required images to memory: This high level of memory performance is common for such tasks (Guild et al., 2014; Wolfe, 2012)

ATTENTION CUEING TASK

Reaction Times
RTs for the cueing task are plotted in Figure 3. As can be seen, studied cues produced a spatial cueing effect: faster RTs to targets appearing at the cued location than the uncued location. Non-studied cues, however, did not. To assess the statistical significance of these differences I subjected RTs to a 2 (Cue Type: studied cue vs. non-studied cue) × 2 (Cueing: cued vs. uncued) within-subjects ANOVA. This analysis revealed a main effect of cue type, $F (1,16) =$
4.763, \( p = .044, \eta^2 = 0.229 \), and no main effect of cueing, \( F(1,16) = 3.019, p = .101, \eta^2 = 0.159 \). Importantly, this analysis also revealed a significant two-way interaction between cue type and cueing, \( F(1,16) = 10.265, p = .006, \eta^2 = 0.391 \), which is the signature of a contingent cueing effect. Planned post-hoc t-tests assessing the cueing effect for each cue type separately to better understand this two-way interaction were conducted. These analyses revealed a significant cueing effect for studied cues, \( t(16) = 2.904, p = .010, d = 0.704 \), but not for non-studied cues, \( t(16) = 1.103, p = .286, d = 0.267 \). Altogether, the results indicate that attentional capture was contingent on whether images were previously studied or not, suggesting that ACSs can be specified based on representations in LTM.

![Figure 3](image.png)

*Figure 3.* Reaction time data for Chapter 1, revealing selective cueing effects. Studied cues produced cueing effects, but not non-studied cues. Error bars in this figure, and all subsequent figures, are within-subject standard errors (Morey, 2008).
Accuracy

To assess speed-accuracy trade-offs, task accuracy on non-catch trials was subjected to the same 2 (Cue Type: studied vs. non-studied) × 2 (Cueing: cued vs. uncued) ANOVA. This analysis revealed no main effects or interaction, all F-values < 1.62, suggesting that the differences in RT were not driven by speed-accuracy trade-offs. Error rates are noted in Table 1 below.

<table>
<thead>
<tr>
<th>Cue Type</th>
<th>Cued</th>
<th>Mean</th>
<th>SE</th>
<th>Uncued</th>
<th>Mean</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studied</td>
<td>18.21</td>
<td>1.10</td>
<td></td>
<td>19.88</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td>Non-Studied</td>
<td>18.09</td>
<td>1.00</td>
<td></td>
<td>17.58</td>
<td>1.50</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1.* Error rates (%) and standard errors of the mean (SE) for Chapter 1. No effects of cueing, cue type, or their interaction on accuracy were observed.

Discussion

In the present study I observed a selective cueing effect. Cues only captured visual spatial attention when they matched the studied targets in our task, and thus matched participants’ ACSs. Whereas previous contingent capture studies have implemented cueing tasks where targets and ACSs are defined by a single feature, feature dimension, or semantic category, here participants were required to search for any one of a large set of objects. In doing so, the potential memory systems that participants could use to define the content of their ACSs were constrained. Working memory is unlikely to contribute much to ACSs in the present task because the number of target objects far exceeds the information capacity of visual working memory. Semantic memory is also unlikely to play a role for two reasons. First, target items were selected from multiple semantic categories. Second, because studied and non-studied cues were selected from the same pool of images, there was inevitably semantic overlap between the two sets of cues; yet, only studied cues captured attention. Remembering a set of 30 images is,
however, conducive to the use of episodic LTM. Accordingly, the selective cueing effects in the present experiment suggest that participants can adopt ACSs that are defined by representations stored in episodic LTM.

Beyond the selective cueing effect, a main effect of cue type was observed. Overall, targets were reported more slowly following studied cues than non-studied cues. This finding suggests that studied cues may have an additional non-spatial orienting effect on behaviour. One speculative interpretation is that this slowing is a consequence of the reflexive memory retrieval of task-relevant items. Importantly, because the effect is not driven by spatial orienting it is distinct from the effects of visual spatial attentional capture.

One limitation of Chapter 1 is that error rates were generally high in the attention cueing task, forcing us to eliminate 11 out of 28 participants (39% of participants). Given that participants performed near ceiling on the memory training task (average of 98.68% before elimination of participants), it is unlikely that low accuracy in the cuing task was due to a failure to retrieve the target images from LTM. Instead, high error rates may reflect the difficulty of the speeded orientation judgements made in the cueing task. In Chapter 2, the cueing-task judgement was changed to an easier colour discrimination task to assess the extent to which high error rates contributed to the selective cueing effects in Chapter 1.
CHAPTER 2: SELECTIVE CUEING EFFECTS FOR EPISODIC LTM
ACS-MATCHING CUES ARE NOT IMPACTED BY POOR TASK PERFORMANCE

In Chapter 2 I used a modified version of the attention cueing task from Chapter 1 to evaluate whether the observed selective cueing effects were influenced by low task accuracy. To reduce errors, the response required of participants in the cueing task was modified. Rather than judging the direction of rotation of the images, the target-frame images remained in their original orientation, and when they appeared the placeholder boxes changed colour: one to red, the other to green. Participants were again required to locate the target, but now reported the colour of the surrounding placeholder box. To ensure participants were actively searching both placeholders, catch trials (13% of trials) were included for which both target-frame images were non-studied, and participants were required to withhold responding. To preview the results, the error rates were successfully reduced relative to Chapter 1 and the selective cueing effect was replicated, supporting the conclusion that ACSs can be specified based on LTM representations.

METHODS

PARTICIPANTS

Twenty-one undergraduate students (mean age of 19.62) at the University of Guelph received partial course credit as compensation for their participation. Participants with error rates at 15% or higher (one of 21 participants) were excluded from the analyses, resulting in a sample of 20 participants. Of note, including all participants did not alter the pattern of statistically significant effects reported below. All participants reported having normal colour vision and normal, or corrected-to-normal, visual acuity.

APPARATUS, STIMULI, AND PROCEDURE.

The apparatuses, image databases, and training task were identical to Chapter 1. The cueing task was very similar, but differed in one respect (displayed in Figure 4). On each trial, target images were presented in their original orientation; however, at the same time that targets...
appeared one of the placeholder boxes was randomly selected to change colour from black to green, and the other from black to red. Participants were instructed to report the colour of the placeholder surrounding the target using a key press. On 13% of trials (48 of 368), both images were non-studied and from the pool of 2,100 images (catch trials). If neither box contained a target, participants were instructed to withhold responding and wait for the trial to time out.

**Figure 4.** Example trial sequence of the attention cueing task for Chapter 2. As in Chapter 1, participants were instructed to locate the studied object and ignore the concurrently presented non-studied object. In the response frame, the placeholder boxes surrounding the objects changed colour (one always red, the other always green) and participants were required to indicate the colour of the box surrounding the studied object. As in Chapter 1, the response frame was preceded by a cue frame in which a single task-irrelevant object appeared in one of the placeholder boxes. All SOAs were 150 ms (duration of cue frame) plus the ISI (inter-stimulus interval).
RESULTS

The same method of recursive trimming from Chapter 1 was used on the RT data, resulting in the removal of 2.54% of trials. Error trials (4.46% of all trials) were also excluded from RT analyses. On average, participants successfully withheld responses on 96.9% of catch trials.

MEMORY TRAINING TASK.

All participants completed memory training in two blocks (the minimum). The average accuracy (proportion of correct target recognition trials) across participants and blocks was 99.25% suggesting that participants were able to effectively commit the required images to memory.

ATTENTION CUEING TASK

Reaction Times

RTs for the cueing task are plotted in Figure 5. As in Chapter 1, studied cues appeared to produce a spatial cueing effect, while non-studied cues did not. To assess the statistical significance of these differences I subjected RTs to a 2 (Cue Type: studied cue vs. non-studied cue) × 2 (Cueing: cued vs. uncued) within-subjects ANOVA. This analysis revealed a main effect of cue type, $F(1,19) = 17.694, p < .001, \eta^2 = 0.482$, but not cueing, $F(1,19) = 0.471, p = .501, \eta^2 = 0.024$. Importantly, a significant two-way interaction between cue type and cueing, $F(1,19) = 6.703, p = .018, \eta^2 = 0.261$, was observed, replicating the main finding from Chapter 1. Planned post-hoc t-tests revealed a statistically significant cueing effect for studied cues, $t(19) = -2.276, p = .035, d = 0.509$, but not non-studied cues, $t(19) = 1.041, p = .311, d = 0.233$. These results replicate the selective cueing effect found in Chapter 1, providing further evidence in favour of contingent attentional capture.
Figure 5. Reaction data for Chapter 2, revealing selective cueing effects. Studied cues produced cueing effects, but not non-studied cues.

Accuracy

To assess speed-accuracy trade-offs, I subjected task accuracy on non-catch trials to the same 2 (Cue Type: studied vs. nonstudied) × 2 (Cueing: cued vs. uncued) ANOVA. This analysis revealed no main effects and no interaction, all $F$-values < 1, suggesting that the differences in RT were not driven by speed-accuracy trade-offs. Importantly, error rates in Chapter 2 are significantly lower overall than in Chapter 1, $t(35) = 7.180$, $p < .001$. Error rates are noted in Table 2 below.

<table>
<thead>
<tr>
<th>Cue Type</th>
<th>Cued</th>
<th>Mean</th>
<th>SE</th>
<th>Uncued</th>
<th>Mean</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studied</td>
<td>4.63</td>
<td>0.40</td>
<td></td>
<td>4.15</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Non-Studied</td>
<td>4.18</td>
<td>0.40</td>
<td></td>
<td>4.87</td>
<td>0.80</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Error rates (%) and standard errors of the mean (SE) for Chapter 2. No effects of cueing, cue type, or their interaction on accuracy were observed. Error rates were significantly lower than in Chapter 1.
DISCUSSION

Chapter 2 replicates the selective cueing effect of Chapter 1, while limiting error rates. Together, both studies employed the conventional cuing task used to study contingent attentional capture, and both experiments observed that participants can adopt ACSs specific to a set of 30 unique visual objects. These findings support the conclusion that ACSs can be specified based on episodic LTM representations.
Chapter 3 was conducted to investigate an additional marker of attentional capture, inhibition of return, in order to determine whether the cueing effects measured in Chapters 1 and 2 are caused by capture, or possibly by voluntary selection. Relative to past attentional capture studies, the current study used more complex stimuli, potentially increasing task demands. Additionally, RTs were slower in Chapters 1 and 2 than those typically seen in cueing tasks (Posner & Cohen, 1984). One might wonder, therefore, if participants used this extra time to volitionally attend to cues despite the fact that they were non-predictive of target location and could not be used to improve task performance. Therefore, Chapter 3 investigated the extent to which the cueing effects reported thus far are due to controlled, volitional attention by increasing the cue-target stimulus onset asynchrony (SOA) and testing for a different marker of attentional capture: inhibition of return (IOR; i.e., a reversal of the cueing effect over time). Whereas the benefits of volitional attention tend to persist for more than a second after cue onset (Posner, Nissen, & Ogden, 1978), the benefits of capture are more transient and reverse at SOAs of about 200–300 ms (Lupianez, 2010; Posner et al., 2007). Thus, the conclusions from Chapters 1 and 2, that the cueing effects reflect attentional capture, provides a clear, falsifiable prediction that the effects should reverse at longer SOAs. If the effect was driven by voluntary attention, facilitation (faster on cued vs. uncued trials) should persist at these long SOAs. Notably, this experiment did not include an ACS manipulation (i.e., the cue frame always included a studied item only) to instead focus on direct tests of capture: whether or not IOR would be observed.

Chapter 3 presented studied cues with a cue-target SOA of 150 ms, 250 ms, or 400 ms. Here, the 400 ms SOA is the primary condition of interest. If the cueing effects observed in Chapters 1 and 2 reflect attentional capture, I should observe IOR for the 400 ms SOA condition in Chapter 3, as transient capture effects typically reverse by this time point. Alternatively, if the cueing effects in Chapters 1 and 2 reflect volitional attention, then cueing effects should persist at the 400 ms SOA given their relatively sustained time course in comparison to capture effects. Otherwise, the experimental design was similar to Chapter 2.
METHODS

PARTICIPANTS

Twenty-five undergraduate students (mean age of 19.07) at the University of Guelph received partial course credit as compensation for their participation. Five participants with error rates at 15% or higher were excluded from the analyses, resulting in a sample of 20 participants. Including all participants did not alter the pattern of statistically significant effects reported below. All participants reported having normal colour vision and normal, or corrected-to-normal, visual acuity.

APPARATUS, STIMULI, AND PROCEDURE

The apparatuses, image database, training task, and cueing task trial layout remained the same as in Chapter 2, except that (a) only studied cue trials were included and (b) the cue frame and target frame were separated by a variable SOA, each with equal probability (displayed in Figure 6). The SOA manipulation was mixed within blocks. As with Chapter 2, on 13%, or 48 of 368 trials (catch trials), both images were nonstudied and from the 2,100 image database. This resulted in a fully crossed, 3 (SOA: 150, 250, or 400 ms) × 2 (Cueing: cued vs. uncued) within-subject design, with four blocks of 105 trials each.
Figure 6. Example trial sequence of the attention cueing task for Chapter 3. The trial sequence was identical to that in Chapter 2, with the exception of the SOA duration (duration of the cue frame and the ISI): either 150, 250, or 400 ms.

RESULTS

Again, the same method of recursive trimming was used on the RT data resulting in the removal or 2.29% of trials. Error trials (5.67% of all trials) were also excluded from RT analyses. On average, participants successfully withheld responses on 96.33% of catch trials; range: 81.25% to 100%. The sphericity assumption was violated for the test of SOA as measured by Mauchly’s test of sphericity, $\chi^2(2) = 6.305, p = 0.043$, thus the Hyunh-Feldt correction, $\varepsilon = 0.827$, was applied to the degrees of freedom; no other violations of the sphericity assumption were evident.

MEMORY TRAINING TASK.
All participants completed memory training in two blocks (the minimum). The average accuracy (proportion of correct target recognition trials) across participants and blocks was 99.58% suggesting that participants were able to effectively commit the required images to memory.

**ATTENTION CUEING TASK**

**Reaction Times**

RTs for the cueing task are plotted in Figure 7. IOR appeared to be present at the 400 ms SOA, while no cueing effects were present at the 150 and 250 ms SOAs. To assess the statistical significance of these differences I subjected RTs to a 3 (SOA: 150, 250, or 400 ms) × 2 (Cueing: cued vs. uncued) within-subjects ANOVA. This analysis revealed no two-way interaction between SOA and cueing, $F(2,38) = 0.548, p = .583, \eta^2 = 0.028$, a main effect of SOA, $F(1.65,38) = 5.612, p = .012, \eta^2 = 0.228$, and no main effect of cueing, $F(1,19) = 2.804, p = .110, \eta^2 = 0.129$. However, given our primary a priori prediction of IOR in the 400 ms SOA condition, I followed this omnibus analysis with planned post-hoc t-tests assessing the cueing effect for each SOA separately. These analyses revealed statistically significant IOR at the 400 ms SOA, $t(19) = 2.158, p = .044, d = 0.483$, indicating faster responses for uncued compared to cued trials, but no detectable differences at the 150 and 250 ms SOAs: $t(19) = 0.417, p = 0.681, d = 0.09$ and $t(19) = 1.145, p = 0.266, d = 0.26$, respectively.
Figure 7. Reaction time data for Chapter 3. Inhibition of return is present at long SOAs (400 ms) suggesting that cueing effects in the present experiments were caused by attentional capture rather than volitional attention.

**Accuracy**

To assess speed-accuracy trade-offs, task accuracy on non-catch trials was subjected to the same 3 (SOA: 150, 250, or 400 ms) × 2 (Cueing: cued vs. uncued) ANOVA. This analysis revealed no main effect of SOA, no main effect of cueing, and no interaction, all $F$-values < 1.19, suggesting that the differences in RT were not driven by speed-accuracy trade-offs. Error rates are noted in Table 3 below.

<table>
<thead>
<tr>
<th>Chapter 3 Mean Error Rates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cueing</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>150</td>
</tr>
<tr>
<td>250</td>
</tr>
<tr>
<td>400</td>
</tr>
</tbody>
</table>

Table 3. Error rates (%) and standard errors of the mean (SE) for Chapter 3. No effects of cueing, SOA, or their interaction on accuracy were observed.
DISCUSSION

The primary purpose of Chapter 3 was to use an IOR paradigm to investigate the mechanism behind the cueing effects I observed in Chapters 1 and 2. After extending the SOA between cues and targets to 400 ms, a reversal of the cueing effect – that is, IOR – was observed. This shift from facilitation to inhibition suggests that the observed cueing effects are unlikely the result of volitional orienting. It is interesting to note that a cueing effect was not observed in the 150 ms SOA condition of the present experiment, which appears at odds with the significant cueing benefit observed at the same SOA in Chapter 1, and replicated in Chapter 2. Despite the absence of this cueing effect, the fact that cues produced IOR in Chapter 3 allows for confidence in the notion that cues captured visual spatial attention (see Pratt, Hillis, & Gold, 2001).

As well, two aspects of Chapter 3 differed from Chapters 1 and 2: only studied cues were presented in Chapter 3 and additional SOAs were introduced. The absence of non-studied cues is unlikely to have altered short SOA cueing effects, which are more strongly determined by properties of the target display than the types of cues that are presented (e.g., Bacon & Egeth, 1994; Folk et al., 1992). Rather, it is likely that the absence of the cueing benefit at the 150 ms SOA is a consequence of our use of a mixed-block design in Chapter 3, which introduced a large range of SOAs (150 to 400 ms) within every block. Work by Cheal and Chastain (2002) has noted the sensitivity of IOR to within-block SOA manipulations. As the range of SOAs within a block increases, IOR effects emerge at shorter SOAs, and the transition from cueing benefits to IOR can happen at least as early as SOAs of 150 ms. Importantly, the predicted IOR effect is present at the long SOA, providing evidence that volitional attention is contributing at most minimally to the observed cueing effects.
Chapter 4: Converging Evidence for Selective Capture by LTM ACS-Matching Items While Controlling for Perceptual Masking

Chapter 4 was motivated by similar factors as Chapter 3: the spatial overlap of cue and target stimuli. Having cue stimuli occupy the same physical space as the target on cued trials could potentially cause perceptual masking, a factor that also may have increased task demands in the present study. Chapter 4 therefore focused on another direct test of capture while controlling for potential sensory effects: whether selective cueing effects are replicated when cue stimuli are presented in both target locations.

On every trial a studied cue appeared in one placeholder, and a non-studied cue in the other. While this paradigm does not afford separate measurement of cueing effects by studied and non-studied items (as was possible in Chapters 1 and 2, since non-studied trials with two non-studied cues would equate the saliency of both spatial locations) and thus cannot be used to verify that non-studied cues do not produce measurable cueing effects, it does equate potential perceptual masking at both target locations. Observing faster RTs when the target appears at the location of the studied cued, rather than the non-studied cue, would provide a replication of the selective cueing effects observed in Chapters 1 and 2, while controlling for masking.

METHODS

Participants

Twenty-one undergraduate students (mean age of 19) at the University of Guelph received partial course credit as compensation for their participation. One participant did not complete the study, and one with an error rate at 15% or higher was excluded from the analyses, resulting in a sample of 19 participants. Including all participants who completed the study did not alter the pattern of statistically significant effects reported below. All participants reported having normal colour vision and normal, or corrected-to-normal, visual acuity.
**APPARATUS, STIMULI, AND PROCEDURE**

The apparatus, image database, training task, and cueing task trial layout remained the same as in Chapter 3, except that (a) on all trials the cue frame and target frame were separated by a 150 ms SOA and (b) two cue images were present on each trial: one studied and one non-studied (displayed in Figure 8). This resulted in a single within-subjects condition of Cueing (studied cue: target appeared in the same location as the studied cue vs. non-studied cue: target appeared in the same location as the non-studied cue), with 12 blocks of 31 trials each (for a total of 372 trials). As in Chapters 1 to 3, on approximately 13%, or 48 of 372 trials (catch trials), both images were non-studied and from the 2,100 image database.

**Figure 8.** Example trial sequence of the attention cueing task for Chapter 4. The trial sequence was identical to that in Chapter 2 with the exception of the cue frame: One task-irrelevant object appeared in each of the placeholder boxes, one studied and one non-studied.
RESULTS

Again, the same method of recursive trimming was used on the RT data resulting in the removal of 2.49% of trials. Error trials (5.42% of all trials) were also excluded from RT analyses. On average, participants successfully withheld responses on 97.26% of catch trials; range: 87.50% to 100%.

MEMORY TRAINING TASK.

All participants completed memory training in two blocks (the minimum). The average accuracy (proportion of correct target recognition trials) across participants and blocks was 99.83% suggesting that participants were able to effectively commit the required images to memory.

ATTENTION CUEING TASK

Reaction Times

RTs for the cueing task are plotted in Figure 9. A cueing effect (faster when the target appeared at the location of the studied vs. non-studied cue) appeared to be present. To assess the statistical significance of this difference RTs were subjected to a paired samples t-test (Cueing: studied cue vs. non-studied cue), revealing a significant effect of cueing, \( t(18) = 3.508, p = .003 \), \( d = 0.804 \). These results replicate and extend the facilitation effects observed in Chapters 1 and 2.
Figure 9. Reaction time data for Chapter 4. The 150 ms SOA cueing effect for studied cue images is replicated using a modified, two-cue paradigm.

**Accuracy**

To assess speed-accuracy trade-offs, task accuracy on non-catch trials was subjected to the same paired samples t-test (Cueing: studied vs. non-studied cue). This analysis revealed no effect of cueing on accuracy, $t(18) = -0.375, p = .712$, suggesting that the differences in RT were not driven by speed-accuracy trade-offs. Error rates are noted in Table 4 below.

<table>
<thead>
<tr>
<th></th>
<th>Studied Cue</th>
<th>Non-Studied Cue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.30</td>
<td>5.54</td>
</tr>
<tr>
<td>$SE$</td>
<td>0.45</td>
<td>0.45</td>
</tr>
</tbody>
</table>

*Table 4.* Error rates (%) and standard errors of the mean (SE) for Chapter 4. No effects of cueing were observed.
DISCUSSION

Together with Chapters 1 to 3, the results of Chapter 4 provide strong, converging evidence that support the conclusion of facilitatory capture at short SOAs, and more generally, that LTM representations can support the implementation of the ACSs that guide the capture of attention. While these results are consistent with the prediction that capture can be guided by an initial, rapid stage of episodic retrieval, it is unclear whether this episodic retrieval is accomplished through recollection or familiarity (i.e., the recall of prior learning with vs. without awareness of the source of learning; Yonelinas, 2002). This question is more fully evaluated in Chapter 5 using a conventional source manipulation (i.e., studied items were divided into “Set A” and “Set B”, and only one was designated as the targets in the attention task) to determine which type of LTM retrieval process supports these capture effects. Furthermore, these findings are extended through the use of the rapid serial visual presentation (RSVP) task – an alternate experimental paradigm used to measure attentional capture (Folk et al., 2002; Folk et al., 2008; Wyble et al., 2013).
Chapter 5: Episodic LTM ACSs are supported by source-specific recollection memory

Building on the evidence from Chapters 1 to 4 that participants can specify ACSs based on episodic LTM representations, the central aim of Chapter 5 was to elucidate the type of episodic long-term retrieval being used. Namely, a list of 30 visual objects could potentially be maintained through either of two distinct neural processes that both underlie episodic memory: familiarity or recollection (Yonelinas, 2002). Familiarity refers to the feeling that a particular piece of information has been encountered before, but no specific contextual information is recalled (e.g., recognizing a former colleague on the bus, but not recalling how you know them). Recollection on the other hand involves recall of specific contextual information (i.e., details of the source of information in LTM). If participants can adopt ACSs for a contextually specific set of visual objects, this would help to clarify the type of episodic memory participants are using (i.e., in favour of recollection rather than familiarity) while also providing further evidence against the contribution of semantic memory to these tasks. Accordingly, the memory training task was modified by dividing the 30 images into two sets of 15 images (see Figure 10); one of these sets was designated as the targets for the entirety of the subsequent attention task, while requiring participants to ignore the other set. By separating studied information into distinct sets of information, participants are forced to recall the source (i.e., the specific subset) of studied items in order to accurately complete both the memory training and the attention task.

In addition to introducing a source manipulation (Radvansky, 2011), in Chapter 5 the attention task was adapted in order to gain converging evidence for the role of episodic memory in contingent attentional capture. While the majority of contingent capture investigations have employed attention cueing tasks, rapid serial visual presentation (RSVP) tasks have also become a common measure of capture (Folk et al., 2002; Folk et al., 2008; Wyble et al., 2013). On each trial of an RSVP task, participants are presented with a sequential stream of visual objects that rapidly appear every 100 ms, and one of the objects is a target that must be identified upon completion of the stream (see Figure 11). To measure attentional capture, a distractor stimulus precedes the target, appearing spatially adjacent to the RSVP stream. If, and only if, the
distractor captures spatial attention, it produces a spatial blink: impaired target identification when the distractor precedes the target by 100 to 700 ms, but no impairment otherwise (Folk et al., 2002; Folk et al., 2008). Using RSVP tasks, it has been demonstrated numerous times that attentional capture by such distractors is contingent on ACSs.

Here participants first memorized two sets of 15 object images that were designated as set A and set B. Participants then completed the RSVP attention task, for which they were required to report any target from set A, while ignoring targets from set B (set assignment was counterbalanced across participants). If participants can establish ACSs based on the recollection of episodic LTM representations, then they should be able to establish an ACS specific to set A images. Such set-specific effects are a classic operationalization used to demonstrate the contribution of recollection of episodic memories in list-learning studies (Jacoby, 1991). If participants instead establish ACSs based on familiarity LTM, which is source agnostic, the ACS will include both the task-relevant set A images and task-irrelevant set B images. To assess the status of participants’ ACSs, distractors from set A, set B, and non-studied distractors were presented with equal frequency. To preview the results, only set A distractors (i.e., distractors from the target set) produced a spatial blink when the target followed the distractor by 300 ms (i.e., reduced accuracy on trials with a set A distractor as compared to those with a non-studied distractor), suggesting that participants were able to use recollection-based episodic LTM when establishing ACSs, thereby excluding set B items despite their familiarity.

METHODS

PARTICIPANTS

Twenty undergraduate and graduate students were recruited through on-campus advertisements or from an introductory Psychology course at the University of Guelph (mean age of 22.45 years). Participants received either monetary compensation or course credit for their participation. All participants reported having normal colour vision and normal, or corrected-to-normal, visual acuity. Two participants were removed from analyses because they did not follow the task instructions, resulting in a sample of 18 participants.
APPARATUS, STIMULI, AND PROCEDURE

The experiment was conducted on a 27 inch ViewSonic Monitor with a 1920 × 1080 resolution LCD display at a screen refresh rate of 60 Hz, and responses were made on a standard keyboard. The image database used for the stimuli remained the same as the previous experiments.

MEMORY TRAINING TASK

The memory training task was similar to previous experiments, but differed in a few critical ways. The set of 30 object images studied by each participant was randomly separated into two sets of 15—set A and set B—for the duration of the experimental session (see Figure 10). Passive viewing began with the label “Set A” followed by the 15 set-A objects one at a time, and then the label “Set B” followed the 15 set-B objects. Participants were instructed that they would have to recall all objects, including which set each object belonged to. During the test phase of training, object images were presented one at a time, and participants were instructed to indicate whether the object belonged to set A, set B, or was not previously studied. All set A and set B images were presented during the test phase, in addition to 30 never-seen-before images selected without replacement from the pool of 2100 images. To ensure that all participants could confidently discriminate set A and set B objects, participants were required to achieve 90% accuracy or above on this memory test on four consecutive blocks before they could move on to the RSVP task.
Figure 10. Trial sequence of the memory training task for Chapter 5. Left panel. Study phase: participants were shown two sets of 15 objects for 3000 ms each, with an inter-stimulus interval of 950 ms, and were instructed to commit these objects to memory. Each set was explicitly designated as separate from the other, and the presentation of each was separated by an interval of time controlled by participants. Right panel. Recollection phase: participants were shown 45 trials containing one object each and were required to indicate whether the object was from the first set of studied objects (set A), the second set (set B) or was not previously studied. Both phases were repeated until participants achieved 90% accuracy on four consecutive recognition phases.

**RSVP TASK**

Following memory training, participants completed a modified RSVP task (see Figure 11). They were instructed to monitor the centre of the display for the appearance of images from one set while ignoring the appearance of images from the other set, henceforth referred to as the target set and non-target set. The target and non-target set identities were consistent throughout the entirety of the RSVP task. Each trial of the task began with a fixation frame: a black central
fixation cross measuring $1 \times 1^\circ$ presented for 500 ms. This was followed by a blank screen for 200 ms, and 20 centrally located images subtending $5 \times 5^\circ$, presented sequentially for 100 ms each with an inter-stimulus interval of 50 ms. One of the images, between the 11th and 16th frames, was an image from participants’ designated target set, while another image within these frames was from their non-target set. Additionally, two, five, or eight frames (i.e., lags) before the target image, the central stimulus was flanked by two additional images subtending $7 \times 7^\circ$, one $5^\circ$ below the central image, and the other $5^\circ$ above.

One of these images was always a non-studied image selected randomly without replacement from the large image database, while the other was either an image from the target set, an image from the non-target set, or, was a non-studied image. Non-studied distractor images were included to serve as a baseline for stimuli that do not capture attention. Accordingly, distractor images either matched the participants’ ACSs (target set distractor) or not (non-target set distractor or non-studied distractor). These flanking images remained onscreen during the ISI and were removed from the screen once the next frame began. This frame was designated as the distractor frame. Participants were told that the distractors were irrelevant to their task, and should be ignored. All other images in the stream were non-studied images from the pool of 2,100 images. Participants were asked to fixate their vision on the central images during the image stream, and to report the identity of the target once the stream completed.

Following the RSVP stream, a probe frame appeared in which eight images were presented around a black central fixation cross measuring $1 \times 1^\circ$. At this point participants were required to select the target image they had seen in the stream. The surrounding images all subtended $3 \times 3^\circ$ and were presented at eight separate locations on an imaginary $10 \times 10^\circ$ square centred on fixation. One of the images was the target from the RSVP stream, another the non-target image from the RSVP stream; the remaining six images consisted of different images from the target set (three images) and the non-target set (three images). Notably, the critical distractor image was never present as a response option in the probe frame to prevent participants from confusing the distractor image with the target, especially when the distractor image is also from the target set and appears closely in time before the target. The use of non-target set images in the stream and probe screen, as well as the additional target and non-target images in the probe screen, served as markers of false alarm responses (i.e., if participants make an incorrect
response, do they choose another target set item, reflecting accurate memory for target set items in turn suggesting an incorrect response was due to a spatial blink, or are they equally likely to choose a non-studied item, reflecting impaired memory for target set items). They also ensured equal levels of familiarity with each image type. The probe frame remained on screen until participants responded by selecting the number on the keyboard number pad that corresponded to the target’s spatial location. Participants were encouraged to be as accurate as possible in their responses. For incorrect or response-absent trials, a 500 Hz tone was generated for 50 ms.

**Figure 11.** Trial sequence of the rapid serial visual presentation (RSVP) task in Chapter 5. Participants were instructed to use the probe screen to report the identity of either the target set (e.g., set A) or non-target set (e.g., set B) item (counterbalanced across participants) presented within the RSVP stream on each trial. Two, five, or eight frames prior to the target, two distractors appeared flanking the RSVP stream. One of the flanking objects was always a non-studied object, while the other could be an additional non-studied object, a studied object from the target set, or a studied object from the non-target set.
The RSVP task consisted of four blocks of 81 trials each (324 trials total), with all trials presented in a random order. All levels of distractor type (target studied set, non-target studied set, or non-studied) and distractor lag (two, five, or eight frames prior to the target) were fully crossed within subjects, and each unique combination of conditions occurred with equal frequency. The location of each image within the probe screen was randomly determined on each trial.

RESULTS

MEMORY TRAINING TASK
All participants completed memory training in four blocks (the minimum for two memory sets). Average accuracy (proportion of trials in which the image source was correctly identified) was 97.28% suggesting that participants were able to effectively commit the required images to memory, as well as the source of encoding.

RSVP TASK
Accuracy for the RSVP task is plotted in Figure 12. A spatial blink effect (i.e., an impairment in task performance at lag 2) appeared to be present for target set distractors but not non-target set distractors or non-studied distractors. To assess the statistical significance of these differences, accuracy rates were subjected to a 3 (Distractor Type: studied target, studied non-target, or non-studied distractor) x 3 (Lag: two, five, or eight) within-subjects ANOVA. No violations of the sphericity assumption were evident as measured by Mauchly’s Test of Sphericity. This analysis revealed a statistically significant main effect of distractor type, $F(2, 34) = 7.745, p = .002, \eta^2 = 0.313$, but not lag, $F(2, 34) = 0.214, p = .809, \eta^2 = 0.012$. Importantly, this analysis revealed a significant two-way interaction between distractor type and lag, $F(4, 68) = 3.052, p = .023, \eta^2 = 0.152$, as would be expected if the spatial blink (i.e., lag effect) was contingent on distractor type. Given our primary a priori prediction of a selective spatial blink effect (decreased accuracy on
lag 2 trials) for target distractors, the omnibus analysis was followed with planned within-subjects t-tests comparing performance across lag 2 for each distractor type. These analyses revealed that performance at lag-2 was significantly worse for trials with studied target-set ($t(17) = 4.547, p < 0.001, d = 0.457$) but not studied non-target set ($t(17) = 1.021, p = 0.321$) distractors, as compared to trials with non-studied distractors, and that the difference in performance between studied target and studied non-target distractors at lag-2 was statistically significant ($t(17) = 3.181, p = .005, d = 0.577$), indicating that studied target set distractors selectively captured attention.

These findings motivate the question: When participants made an incorrect response, did they fail to detect the target (e.g., due to a spatial blink), or did they fail to discriminate between target set and non-target set images? Consistent with the interpretation that participants could accurately recall the source of each image set, on error trials, participants were more likely to choose an incorrect target set item (8.7% of all trials) than either the non-target set item from the RSVP stream (1.2% of all trials), $t(17) = 6.745, p < .001$, or, one of the other non-target set items (1.1% of all trials), $t(17) = 7.508, p < .001$. 


Figure 12. Accuracy data for Chapter 5, revealing a contingent spatial blink (i.e., impairment of target detection at lag 2), that occurred only for studied target distractors.

DISCUSSION

Chapter 5 clarifies the mechanisms supporting contingent attentional capture effects found in Chapters 1 to 4. By having participants memorize two sets of images, but only search for one during the attention task, the contribution of recollection vs. familiarity to participants’ ACSs was examined. The finding that only target set images captured attention demonstrates that participants could establish ACSs based on memory representations tied to a specific context (i.e., set), a signature of recollective LTM. Moreover, I employed a different paradigm, the RSVP task, in which attentional capture is measured by the presence of a spatial blink. Consequently, these data improve the generalizability of the results in Chapters 1 to 4, extending them beyond attention cueing tasks.
In Chapters 1 to 5 I asked whether individuals can adopt ACSs based on representations stored in episodic LTM, as evidenced by selective spatial capture by LTM ACS-matching items. Although recent work has demonstrated a role for semantic LTM in storing ACS items, the majority of work has focused on working memory as the candidate memory system. Episodic LTM, a form of memory historically associated with the guidance of voluntary shifts in spatial attention, would serve to support ACSs built on more flexible, yet durable and context-rich representations, and would add to growing evidence in favour of a rapid, two-stage model of episodic recollection. In all Chapters participants first memorized 30 images of complex, naturalistic objects—a task that requires episodic LTM given the storage capacity and duration limitations of working memory. These items were then designated as the targets in different paradigms used to assess the presence and use of ACSs: namely, whether items in episodic LTM that are designated as targets selectively impact stimulus-driven attention. Chapters 1 and 2 used variants of the attention cueing task and found attentional capture, measured by standard facilitatory cueing effects (faster responses when the cue is in the same spatial location as the subsequent target), but only when the cue stimulus matched the LTM ACS. Chapter 3 used an attention cueing task with a cue-target SOA manipulation to test for standard inhibitory marker of spatial capture: inhibition of return, or slowed target responses on cued trials at longer SOAs. Inhibition of return to ACS-matching items was observed, confirming the reliability of the effects in Chapters 1 and 2 (i.e., that it is highly unlikely these effects are driven by voluntary attention). To address the potential role of perceptual masking in Chapters 1 to 3, and the possibility that task demands were therefore increased, Chapter 4 used a two-cue attention cueing task with a standard short (150 ms) SOA. Again, selective cueing effects were present, extending the generalizability of these findings across variants of the cueing task. Lastly, Chapter 5 used a set-manipulation in the memory training task, and designated only one of the sets as the targets in the subsequent attention task to determine the type of episodic memory supporting selective spatial capture: recollection vs. familiarity. The ability to differentiate large amounts of previously learned information based on the original learning context is unique to episodic LTM. An RSVP paradigm, another standard task for measuring capture, was used to again extend the
reliability of the findings in previous Chapters, as the RSVP task is also used for assessing spatial orienting. Here, capture was selective to studied target set matching RSVP stream flanking distractors, rather than both studied target and studied non-target set matching distractors, suggesting episodic LTM ACSs can be supported by context-specific, recollection memory.

Together, the studies in these chapters provide multiple, converging pieces of evidence that ACSs can be specified by source-specific, episodic long-term memory representations. The implications of these findings for our understanding of the role of memory in ACSs will be discussed in greater detail in the General Discussion section of the thesis. However, at present, it is worth highlighting that the evidence demonstrated in Chapters 1 to 5 extends the role of memory in ACSs beyond visual working memory and semantic long-term memory to episodic long-term memory. The extension into episodic LTM highlights a different form of control over spatial capture than currently acknowledged in the ACS literature; ACSs can also be specified by the vast number of representations stored in episodic LTM that, by definition, involve the ability to rapidly associate contextual and even arbitrary information in the formation of these representations.
CHAPTER 6: INTER-TRIAL FEATURE PRIMING DOES NOT EXPLAIN
ACS CONTINGENT CAPTURE, BUT LONGER-LASTING PRIMING MIGHT

What allows for certain episodic memory representations at certain points in time to guide selection, such as the components of your outfit before a party, while all of the others in your vast store of episodic memory do not? Proponents of the contingent capture model suggest this ability is accomplished through top-down control: We can choose which episodic memories will be part of our ACS, and, in turn, which aspects of the environment will capture attention. However, this notion has recently been challenged by the claim that ACSs are adopted through bottom-up feature priming (Awh et al., 2012; Belopolsky et al., 2009; Theeuwes, 2013). According to the feature priming account of ACSs, instead of choosing the contents of ACSs, every time target stimuli are selected for a response during a trial, those stimuli become implicitly primed and, in turn, this priming creates an ACS. As noted in the general introduction, the rationale behind the priming account of capture comes from the use of individual, unchanging targets in the majority of ACS studies; when the task target is instead variable, contingent capture effects are eliminated, suggesting cueing effects are more related to the identity of the target primed on the previous trial. However, given that in studies with variable targets participants are required to adopt a new ACS on a trial-by-trial basis, the lack of contingent capture is confounded by the possibility that participants are simply unable to effectively adopt ACSs, and so capture is driven entirely by bottom-up saliency (Lien et al., 2010; Lien et al., 2014; Neo & Chua, 2006).

Looking back at Chapters 1 to 5, two aspects of the observed contingent capture effects are informative for assessing the feature-priming account of ACSs. First, in these tasks, switching target stimuli from trial to trial occurred without the concomitant switch in ACSs, avoiding the confound from past priming studies. Specifically, although participants maintained the same ACS throughout the experiment, the target set contained more than one item (between 15-30 items) and consequently, when ACS-matching cues/distractors appeared, more often than not the target on the preceding trial was a different stimulus (up to ~96% of the time), greatly limiting the potential contribution of priming. Second, studied and non-studied cues (Ch. 1 to 4)
and distractors (Ch. 5) were randomly selected from the same stimulus database; thus, the featural similarity within studied and non-studied sets of items was no greater than that across these sets. Therefore, on the majority of trials, the target on one trial should have equivalently primed the features of both studied and non-studied cues/distractors on the subsequent trial, again limiting the contribution of feature priming. These two factors not only address the potential confound of past ACS studies using feature-based targets, they satisfy the established criteria for evaluating inter-trial feature priming: a. varying the cue/distractor feature value on a trial-by-trial basis and b. preventing distractor-target feature overlap and thus within trial priming confounds (Awh et al., 2012). Importantly, given the described aspects of the methodology used in Chapters 1 to 5, the feature-priming account would predict minimal evidence for contingent attentional capture, yet across all chapters I observed contingent capture reliably.

Might there be ways to reconcile the results of Chapters 1 to 5 with the feature-priming account? After all, these experiments were not designed for testing this theory, rather for testing the memory systems that support the maintenance of ACS items. One potential way to reconcile these results with the priming account of ACSs is to consider that, over time, participants still repeatedly attended to and selected the same 15 to 30 target items, rendering it possible that the long-term priming of higher-level object representations contributed to the observed ACS-contingent effects. To contrast with the inter-trial feature priming account, I refer to this alternative explanation as the longer-lasting priming account. Does continuously responding to the same set of stimuli over the course of an experiment act to bias selection of such stimuli in subsequent encounters, without invoking the need for a top-down control account of how ACSs are implemented? Priming is typically denoted as a transient effect; evidence in favour of the notion that it can in fact be long-lasting is relatively novel (Cosman & Vecera, 2013a, 2014; Hutchinson & Turk-Browne, 2012; Kruijne & Meeter, 2016; Leber et al., 2009; Thomson, Willoughby, & Milliken, 2014), although supported by the well-known, consistent, and automatic impact of prior learning episodes on on-going behavior (Hommel, 1998, 2004; Logan, 1990, 1998; Shiffrin & Schneider, 1977). Evaluating the potential role of longer-lasting priming is necessary for understanding how the results of Chapters 1 to 5 inform the feature-priming account, and more generally how ACSs are established.
The present study tests this potential role of the long-lasting priming of higher-level object representations in implementing ACS by using the RSVP paradigm from Chapter 5, and by incorporating a direct manipulation of priming. To evaluate the role of longer-lasting priming, participants first memorized two sets of 16 complex, naturalistic images of everyday objects; one of these sets was then designated as the target set, allowing for a replication of the source-specific capture effects of Chapter 5 and, under a voluntary implementation account, to encourage participants to establish an ACS for this set of images. Critically, unbeknownst to participants, only half of the target set items ever appeared as the RSVP target (i.e., only half the target set items were primed through target selection). All target set items, though, appeared as the critical distractor with equal frequency, allowing for a measure of the impact of long-lasting priming. Under a longer-lasting priming account, only the target set items that actually appear as targets in the RSVP task will capture attention when they appear as the critical distractor, whereas the other target-set items will not. Alternatively, under a voluntary implementation account, all target items will capture attention when they appear as the critical distractor, regardless of whether they have appeared as the RSVP target or not (i.e., have been primed or not), since all studied items are included in participants’ ACS without the need to respond to them first.

METHODS

PARTICIPANTS

Thirty-two undergraduate students (mean age of 18.62 years) at the University of Guelph received partial course credit as compensation for their participation. Participants with overall error rates of 30% or higher (5 of 33 participants), who failed to follow task instructions (6 of 33 participants), and one participant who failed to complete the experiment were excluded from the analyses, resulting in a sample of 21 participants. I selected a high error threshold for inclusion to accurately reflect the overall performance of the study sample relative to Chapter 5 (i.e., our previous study using an RSVP task); this decision was made post-hoc using the grand average of task accuracy across all experimental conditions. Of note, however, including all participants in the analyses did not alter the pattern of statistically significant effects reported below. All
participants reported having normal colour vision and normal or corrected-to-normal visual acuity.

**APPARATUS, STIMULI, AND PROCEDURE**

The experiment was conducted on a 18-inch ViewSonic monitor with a 1280 x 1024 resolution CRT display at a refresh rate of 75 Hz, and responses were made on a standard keyboard. The image databases were identical to all previous Chapters.

**MEMORY TRAINING TASK**

The experimental session began with a memory task. This task (see Figure 13 for a schematic representation of conditions) was identical to that in Chapter 5, with the following change: participants studied two sets of 16 images each, one designated as “set A” and the other “set B”.

![Figure 13. Schematic of RSVP distractor conditions and corresponding target manipulations across Chapters 6 (left panel) and 7 (right panel). The distractor conditions in the boxes represent our conditions of interest, relative to non-studied distractor conditions, for discerning between feature priming and voluntary implementation accounts of ACSs.](image)

**RSVP TASK**

Following memory training, participants completed a modified RSVP task (see Figure 13 for schematic representation of conditions). The task was identical to that in Chapter 5, except with the following changes. Unbeknownst to participants, the set of 16 images designated as the target set (i.e., either set A or set B, randomly selected) was divided into two subsets of 8 images: the primed target set and the never-primed target set. The target image on a given trial was
always randomly selected from the primed target set, whereas images from the never-primed target set were never presented as targets, and thus were never primed (via selection). In Chapter 5, the central RSVP stream always contained one target set item that participants had to report, and one non-target set item to avoid reporting. In that chapter, non-target set items were included in the RSVP stream to equate participants’ exposure to these items with their exposure to the target items. In the present study, I adopt a similar approach, and also account for the difference between primed and never-primed target sets. Specifically, the set of 16 non-target items was divided into two subsets of 8 items: primed non-target items which appeared in the RSVP stream, and never-primed non-target items which did not. While technically both sets of items were not be primed (i.e., items from neither set of stimuli were selected and responded to), this labelling indicates how each set relates to the two target sets in terms of exposure.

Importantly, the critical distractor could be from any of the image sets: primed targets, never-primed targets, primed non-targets, never-primed non-targets, or non-studied images. Accordingly, as in previous chapters using the RSVP task, the critical distractor either matched the participants’ presumed ACS (target set distractor) or not (non-target set or non-studied distractor). Critically, though, a matching distractor could also either be primed or never primed, unbeknownst to participants.

The RSVP task consisted of four blocks of 81 trials each (324 trials total), with all trials presented in a random order. All levels of distractor type (target studied set primed, target studied set never primed, non-target studied set primed, non-target studied set never primed, or non-studied) and distractor lag (two, five, or eight frames prior to the target) were fully crossed within subjects, and each unique combination of conditions occurred with equal frequency. The location of each image within the probe screen was randomly determined on each trial.

RESULTS

MEMORY TRAINING TASK

All participants completed memory training in four blocks, indicating that all participants reached the accuracy threshold by the minimum number of testing rounds. Average accuracy was
89.0% for the first round (first two blocks), and 96.0% for the second round (second two blocks), suggesting participants were able to effectively commit the required images to memory, and had source knowledge of the studied images.

**RSVP Task**

**Accuracy**

Accuracy for the RSVP task is plotted in Figure 14. A spatial blink (i.e., reduced task accuracy at lag-2) was selectively present for primed target set distractor trials, while no evidence of a spatial blink was present for never-primed target set, primed non-target set, never-primed non-target set, and non-studied distractor trials. To assess the statistical significance of these differences, task accuracy rates were subjected to a 5 (Distractor Type: primed target set, never-primed target set, primed non-target set, never-primed non-target set, or non-studied) x 3 (Lag: 2, 5, or 8) within-subjects ANOVA. No violations of the sphericity assumption were evident as measured by Mauchly’s Test of Sphericity. This analysis revealed a statistically significant main effect of distractor type, $F(4, 88) = 5.863, p < 0.001, \eta^2 = 0.227$, but not of lag, $F(2, 40) = 2.203, p = 0.124, \eta^2 = 0.099$. Critically, there was a significant 2-way interaction between distractor type and lag, $F(8, 160) = 2.725, p = 0.008, \eta^2 = 0.120$, suggesting that the spatial blink effect visible in the data was different across distractor types. This interaction is consistent with the conclusion that performance was significantly worse on lag-2 trials only when the critical distractor was a primed studied target set image.

To determine whether distractors in each condition captured attention, the omnibus analysis was followed by planned within-subjects $t$-tests. For these tests, accuracy at lag-2 on non-studied-distractor trials was used as a baseline for the absence of capture, and, thus, lag-2 accuracy for each of the other four distractor conditions was compared against this baseline to test for capture. These analyses revealed that performance at lag-2 was significantly worse for trials with primed target set ($t(20) = 3.313, p = 0.003, d = 0.723$) and never-primed target set ($t(20) = 2.082, p = 0.050, d = 0.454$) distractors, as compared to trials with non-studied distractors, but not for trials with primed non-target set ($t(20) = 1.496, p = 0.150, d = 0.326$) and never-primed non-target set ($t(20) = 1.905, p = 0.071, d = 0.416$) distractors, indicating that target set distractors selectively captured attention, replicating our findings from Chapter 5.
Based on our a priori question of whether ACS-selective spatial capture effects observed in previous research are driven by priming, I compared lag-2 accuracy for primed vs. never-primed target set distractor trials. This test revealed significantly worse performance for primed target set distractor trials, \( t(20) = -2.172, p = 0.042, d = 0.474, \) suggesting that primed items captured attention significantly more than never-primed items. Thus, the present finding, that primed target set items have the strongest impact on spatial capture, appears to indicate that priming plays a role in establishing ACSs; however, as will be seen in the next analysis, this conclusion may be premature. Moreover, given that items that had never been primed (i.e., never appeared as RSVP targets) captured attention, it appears that factors other than priming can also contribute to the implementation of ACSs.

Figure 14. Accuracy data for Chapter 6: a selective spatial blink (impairment of target detection on lag-2 trials) for trials with a studied target set critical distractor only. This selective spatial blink was modulated by priming: primed target set items captured attention significantly more than never primed target set items.
Timecourse analyses

From the spatial blink analyses it seems that priming significantly impacts the likelihood that ACS-matching items will capture attention. However, these analyses are restricted to averaged performance over the course of the experimental task. If I can see evidence for increases in capture over time by primed targets, this would provide converging evidence for a long-lasting priming account. If ACSs are implemented through priming, one might predict that every time a stimulus is primed the ACS gets stronger, and the likelihood of capture by that stimulus increases with this accumulating selection history. While this is not a falsifiable prediction of the long-lasting priming account (e.g., it may be the case that one instance of selecting a stimulus for response is adequate to fully prime that stimulus), observing the accumulation of the effects of priming over time would provide confirmatory evidence for this account.

Can I observe patterns of long-lasting priming if I look across individual trials of our experimental sessions? Of course, given that the spatial blink relies on comparisons across conditions, I am unable to directly measure spatial capture on any one trial; however, I can measure capture indirectly through task accuracy. Specifically, if capture increases the more times a particular stimulus has been presented as a target (i.e., with repetition count), then I should observe a negative relationship when regressing task accuracy onto repetition count (i.e., primed items are more likely to capture attention, resulting in poorer task accuracy). Additionally, although all primed targets occurred with equal frequencies throughout the entire experiment, the randomized nature of the trial structure means that the amount of time between the last appearance of a primed target as the RSVP target and its next appearance as the critical distractor necessarily varies. As this time increases—measured as the number of trials separating these two events, and referred to as distance—the selection bias provided by priming (via target selection) may be reduced in comparison to that provided by more recent priming. If distance impacts the bias provided by priming, then I should observe a positive relationship when regressing task accuracy onto distance count (i.e., distantly primed items are less likely to capture attention, resulting in better task performance). Once again, while it is not a falsifiable prediction of the longer-lasting priming account, such a pattern could provide confirmatory evidence. To address these predictions, a generalized mixed-effects logistic regression evaluating the predicted likelihood that a particular distractor image from the primed target set will lead to
an incorrect (or correct) trial response on lag-2 trials, using repetition count and distance as model predictors (model coefficients), while controlling for between-participant variance, was performed. The model (depicted in Figure 15 and summarized in Table 5) provided no evidence that either repetition count or distance count were significant predictors of accuracy. That is, participants were no more likely to be incorrect when items that had been primed more frequently or more recently in the course of the experiment appeared as a distractor on a given trial. In fact, although the trend in the data for distance count was in the direction predicted by a priming account, the opposite was true for repetition count: accuracy increased with frequency. Of note, repetition count is confounded with time in the experiment, and so may be affected by practice effects or fatigue (and indeed appear to be); this potential confound is addressed in more detail in the Discussion of Chapter 7.

Figure 15. The predicted likelihood of accurate vs. inaccurate performance on all lag-2 primed target set distractor trials given repetition count i.e., the number of times a particular distractor image has previously appeared as a target (Left panel) and distance count i.e., the amount of time in trials since the last appearance of a particular distractor image as a target (Right panel). Neither factor was found to significantly predict the likelihood of an inaccurate response on a given trial.
Summary of Chapter 6 Logistic Regression of Repetition Count and Distance Count of lag-2 Primed Target Set Distractors on Task Accuracy.

|                             | β     | Standard Error | z-ratio | P(>|z|) | e^β (odds ratio) |
|-----------------------------|-------|----------------|---------|---------|-----------------|
| Repetition count            | 0.017 | 0.013          | 1.337   | 0.181   | 1.018           |
| Distance count              | 0.017 | 0.018          | 0.926   | 0.354   | 1.017           |
| Intercept (Constant)        | 1.308 | 0.013          | 1.337   | 0.002** | 3.700           |

Model χ^2 = 2.649 (df=2), p = 0.266
n = 375

Significance codes. 0 '****' 0.001 '***' 0.01 '**' 0.05 '*' 0.1 ' ' 1
Notes. n = 21 (observations = 375). The dependent variable in this analysis is accuracy, coded so that 0 = inaccurate and 1 = accurate.

Table 5. Logistic regression of repetition count and distance count for lag-2 primed target set distractor trials in Chapter 6. Neither factor was a significant predictor of accuracy.

Discussion

A timecourse approach to the analysis of priming therefore revealed that factors that might be expected to influence the amount of priming did not influence the amount of capture. Although the selection of target images seemed to modulate the amount of capture, as indicated by a significantly larger spatial blink on trials with primed as compared to never-primed target set distractors, I failed to observe converging evidence for this account upon performing the regression analyses.

The results thus far seem to be at odds. A long-lasting priming account cannot be ruled out given that primed target items modulated contingent capture, yet there is a lack of evidence in support of priming from additional tests of priming (i.e., relationships between capture and the frequency and recency of priming, as measured through repetition count and distance count,
respectively). However, the conclusions from the timecourse analyses performed are limited as they were post-hoc in nature: Repetition count and distance count were not systematically manipulated; they were simply analyzable as a consequence of the trial design. To address this limitation, in Chapter 7 the frequency of priming was directly manipulated (frequently primed versus infrequently primed target set items) and the impact of frequency on the magnitude of capture (i.e., spatial blink effect) was evaluated; under a longer-lasting priming account, this relationship ought to be proportional.
CHAPTER 7: LONGER-LASTING PRIMING DOES NOT EXPLAIN ACS CONTINGENT CAPTURE EITHER; ACSs REFLECT FLEXIBLE CONTROL OVER CAPTURE

The purpose of Chapter 7 was to evaluate directly the impact of priming frequency on the magnitude of capture effects. To do this, the studied target set items were again divided into two groups: This time, however, all items appeared as targets, and the frequency with which they were primed throughout the experiment was manipulated. Participants first memorized a set of 18 images that were then designated as the targets in the RSVP attention task. Unlike Chapter 6, participants were not required to memorize a second set of “non-target” objects; because the set-specific capture effects from Chapter 5 were already replicated in Chapter 6, non-target distractor conditions were omitted from Chapter 7 to increase power in the target-set conditions. The RSVP task was identical to that of Chapter 6. Half of these targets were then, unbeknownst to participants, frequently primed, each appearing approximately every 12 trials (or on 75% of all trials), while the other half were infrequently primed, each appearing approximately every 36 trials (or on 25% of all trials). In following with the results from Chapter 6, I can expect that both types of distractors will capture attention somewhat. The open question is whether the magnitude of capture will be proportional to the magnitude of priming: a larger spatial blink for trials with frequently primed distractors relative to trials with infrequently primed distractors.

METHODS

PARTICIPANTS
Twenty-nine undergraduate students (mean age of 18.39 years) at the University of Guelph received partial course credit as compensation for their participation. Participants with error rates of 30% or higher (1 of 29 participants) were excluded from the analyses, resulting in a sample of 28 participants. Despite higher overall performance of the sample in this experiment as compared to that in Chapter 6, the high error threshold from Chapter 6 was maintained for consistency. Of note, including all participants in the analyses did not alter the pattern of statistically significant
effects reported below. All participants reported having normal colour vision and normal or corrected-to-normal visual acuity.

**APPARATUS, STIMULI, AND PROCEDURE**

The apparatuses and stimuli used were identical to those in Chapter 6.

**MEMORY TRAINING TASK**

The experimental session began with a memory task. This task (see Figure 13 for schematic representation of conditions) was identical to that in Chapters 1 to 4, with the following change: participants studied one set of 18 images. Given that I replicated the source-specific capture effect seen in previous work in Chapter 6, I used a single studied set in Chapter 7 to decrease the number of experimental conditions and therefore increase statistical power. The study and test phase procedure was identical to that of Chapters 1 to 4.

**RSVP TASK**

All instructions and trial design specifications were identical to that of Chapter 6 except that all target set images appeared as targets (see Figure 13 for a schematic representation of conditions). The contents of the probe frame were also modified: In addition to the RSVP target, the probe frame included three additional target set items, and four non-studied items, to serve as markers for false alarm responses. To equate participants’ exposure to frequently primed and infrequently primed target set items, if the target happened to be an infrequently primed image, the three additional target set images on the probe screen comprised one other infrequently primed image and two frequently primed images, and vice versa for trials in which the target was a frequently primed image. As in Chapter 6, the critical distractor and target were never the same image; as well, the distractor on a given trial was never present as a response option in the probe frame.

The RSVP task consisted of four blocks of 81 trials each (324 trials total), with all trials presented in a random order. All levels of distractor type (frequently primed studied target set, infrequently primed studied target set, and non-studied) and distractor lag (2, 5, or 8 frames prior to the target) were again fully crossed within-subjects, and each unique combination of lag and
distractor type occurred in combination with an infrequently primed target image on 25% of trials, and a frequently primed target image on 75% of trials.

RESULTS

MEMORY TRAINING TASK

All participants completed memory training in two blocks, except for one who completed training in three blocks, indicating that the majority of participants reached the accuracy threshold by the minimum number of testing rounds. Average accuracy was 95.8% for the first blocks, 99.6% for the second blocks, and 97.22% for the third, suggesting participants were able to effectively commit the required images to memory.

RSVP TASK

Accuracy

Accuracy for the RSVP task is plotted in Figure 16. A spatial blink (i.e., reduced task accuracy at lag-2 relative to non-studied distractors) was present for both frequently and infrequently primed studied target set distractor trials. To assess the statistical reliability of these differences accuracy rates were subjected to a 3 (Distractor Type: infrequently primed target set, frequently primed target set, or non-studied distractor) x 3 (Lag: 2, 5, or 8) within-subjects ANOVA. The sphericity assumption was violated for the test of lag as measured by Mauchly’s test of sphericity, $\chi^2(2) = 7.096, p = 0.029$, thus the Huynh-Feldt correction, $\varepsilon = 0.846$, was applied to the degrees of freedom. No other violations of the sphericity assumption were evident. This analysis revealed a statistically significant main effect of distractor type, $F(2, 52) = 7.587, p < 0.001, \eta^2 = 0.226$, and a significant main effect of lag, $F(1.69, 52) = 21.385, p < 0.001, \eta^2 = 0.451$. Critically, there was a significant 2-way interaction between distractor type and lag, $F(4, 104) = 3.690, p = 0.008, \eta^2 = 0.124$, suggesting that the spatial blink visible in the data was indeed modulated by distractor type. Based on visual inspection of Figure 16, this interaction may reflect a pattern where performance was significantly worse on lag-2 trials when the critical
distractor was a primed studied set image, whether frequently or infrequently primed, as compared to when it was a non-studied image. To confirm whether this was the case, the omnibus analysis was followed with planned within-subject t-tests (2-tailed) comparing lag-2 performance between frequently primed target set and non-studied distractors, and between infrequently primed target set and non-studied distractors. This revealed a significant decrease in accuracy for both frequently (t(26) = 3.472, p = 0.002, d = 0.668) and infrequently primed (t(26) = 2.969, p = 0.006, d = 0.571) distractor trials, suggesting both distractor types captured attention. To evaluate whether this contingent capture was modulated by priming frequency, I conducted a planned within-subject t-test (2-tailed) comparing frequently and infrequently primed distractor trials. The comparison between frequently (M = 77.46%) and infrequently (M = 79.67%) primed distractors, t(26) = 1.006, p = 0.324, d = 0.194 revealed a null result. Therefore, the results of these analyses align with the conclusions from the regression analyses presented in Chapter 6 (i.e., failure to observe converging evidence for a priming account) in that frequency did not significantly modulate the amount of capture.
Figure 16. Accuracy data for Chapter 7: a selective spatial blink (i.e., impairment of target detection on lag-2 trials) for trials with a studied target set distractor, whether frequently or infrequently primed.

Despite the absence of a significant difference in capture effects for frequently and infrequently primed distractors, the average accuracy for infrequently primed distractor trials was numerically higher than that of frequently primed distractor trials (79.7% vs. 77.5%). Is it possible that there are differences across the priming conditions that were not captured by the ANOVA analyses? It may be the case that even infrequently primed targets are primed “enough” by the end of the experiment, such that priming plateaued in both infrequently and frequently primed conditions. Or, similarly, target set items may only need to be primed once or on several occasions in order to bias attention. Such an explanation would also be able to account for the overall primed vs never-primed difference for studied target set items, but lack of frequency effects, in Chapter 6. As in Chapter 6, I investigate this explanation using regression analyses.
**Timecourse analyses**

To address the aforementioned alternative explanations for the overall spatial blink effects, I performed timecourse analyses of frequency and recency using a generalized mixed-effects logistic regression evaluating the predicted likelihood that individuals are inaccurate on lag-2 trials, across frequently and infrequently primed distractor conditions. The model (depicted in Figure 17 and summarized in Table 6) revealed that, as in Chapter 6, neither repetition count nor distance count were significant predictors of accuracy. The outcomes of this model strengthen confidence in the lack of statistically significant differences in accuracy between frequently and infrequently primed distractor trials.

*Figure 17.* The predicted likelihood of accurate vs. inaccurate performance on all lag-2 primed target set distractor trials given repetition count i.e., the number of times a particular distractor image has previously appeared as a target (Left panel), and distance count i.e., the amount of time in trials since the last appearance of a particular distractor image as a target (Right panel). Neither factor was found to significantly predict the likelihood of capture on a given trial.
Summary of Chapter 7 Logistic Regression of Repetition Count and Distance Count of lag-2 Target Set Distractors (Frequently and Infrequently Primed) on Task Accuracy.

|                      | β     | Standard Error | z-ratio | P(>|z|)          | Odds ratio |
|----------------------|-------|----------------|---------|-----------------|------------|
| Repetition count     | 0.020 | 0.019          | 1.040   | 0.298           | 1.020      |
| Distance count       | 0.001 | 0.003          | 0.085   | 0.932           | 1.001      |
| Intercept (Constant) | 1.484 | 0.228          | 6.519   | 7.09e-11***     | 4.411      |

Model $\chi^2 = 2.5618$ (df=3), $p = 0.4642$

$n = 1553$

Significance codes. 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Notes. $n = 27$ (observations = 1553). The dependent variable in this analysis is accuracy, coded so that 0 = inaccurate and 1 = accurate.

Table 6. Logistic regression of repetition count and distance count for lag-2 primed target set distractor (infrequently and infrequently primed) trials in Chapter 7. Neither factor was a significant predictor of accuracy.

DISCUSSION

With the results from both Chapters 6 and 7 together, what can we conclude? On the one hand, priming seems to play a role, as capture in Chapter 6 was significantly modulated by the presence vs. absence of priming. On the other hand, the impact of other modulations of priming on capture do not seem to be supported; i.e., the frequency and recency of priming do not significantly affect performance. Moreover, participants somewhat included never-primed objects in their ACSs, even though these objects had never appeared as targets in the attention task. How might we reconcile these results?

As mentioned, it is possible that the amount of priming needed for an item to be included in an ACS may be fairly low; in other words the impact of priming may max out fairly quickly, and this is why there are no detectable differences across frequency conditions using the current
measures. For example, perhaps the act of selecting and responding to target stimuli in the initial memory training task is sufficient priming for these objects to be included in participants’ ACSs during the RSVP task. While this would explain how never-primed stimuli capture attention, a contribution of priming from training seems unlikely given the lack of capture by non-target studied distractors in Chapter 6. That is, the memory training phase was identical for target and non-target distractors in Chapter 6, yet only target distractors captured attention.

There is, however, an alternative explanation for these results that does not rely on a priming account. Is it possible that instead of ACSs being implemented through priming, participants start with an ACS that is instantiated through top-down control, but then forget the target items that they never respond to? Given the current state of episodic memory recollection literature, however, a forgetting account is unlikely: Individuals are able to memorize thousands of complex, naturalistic images of visual objects, and recall studied items with above 80% accuracy over the course of a 5.5 hour experiment (Brady et al., 2008). Perhaps instead, the original ACS is refined over the course of the experiment with task experience: Experience selecting and responding to certain items (i.e., primed items) and not others (i.e., never-primed items) may determine which items remain in the ACS, and which items can be removed, respectively, what I will refer to as a flexible top-down account. Again, the initial inclusion of particular items in an ACS would still be top-down, but this inclusion may not be permanent, with items being removed from the ACS in response to changes in the environment over time.

To evaluate the possibility of this model as a contender for explaining the data in the current study, it is necessary to evaluate the impact of selection history directly by evaluating accuracy over time (i.e., on a trial to trial basis) for both primed and never-primed items. Under a priming account, accuracy on trials with primed distractors ought to decrease over time (i.e., capture gets more likely as priming increases), and accuracy is expected to be consistent over time for never-primed items (i.e., they are never primed so their status never changes over time, making them no more or less likely to capture attention over time). Under an adaptive top-down account, however, the opposite pattern of results would be expected. For primed target set items, capture should remain consistent over time: top-down control instantiates the ACS at the start of the experiment and the presence of these items as targets simply confirms these items truly are targets, and thus belong to the ACS. Never-primed target items start off in the same state as
primed target set items: Participants believe these items are just as likely to occur as targets. However, as the experiment goes on and these items never appear as targets, the likelihood of capture by these items, and therefore a bias towards incorrect responses, should decrease over time as the original ACS adapts based on trial history: The absence of certain items may lead them to be “pruned” from the original ACS. Support for this model would also help to explain the null effects of the frequency analyses. The best way to test these differing predictions is with the Chapter 6 data, as these opposing conditions already exist: primed and never-primed studied target set items.

Changes in accuracy over time can be assessed by regressing accuracy onto trial number. Trial number happens to be an important factor for discerning between a priming and flexible top-down account given that the regression co-efficients previously used (i.e., repetition count and distance count) cannot be applied to trials with never primed distractors as these items have never appeared as targets. However, because trial accuracy is being used to indirectly measure the amount of capture, the interpretation of these analyses warrant caution. In particular, factors other than capture may also affect accuracy over time, namely, practice effects and fatigue. In order to gauge the contribution of these potential confounds, the regression of accuracy onto trial number for non-studied distractor trials can be used as a baseline. On these trials, capture does not change over time as distractors are neither studied nor primed, thus any differences over time can only be a consequence of phenomena such as practice effects and/or fatigue.
Figure 18. The predicted likelihood of accurate vs. inaccurate performance on all lag-2 primed target set distractor trials never-primed target set distractor trials, and non-studied distractor trials over time. Time was a significant predictor of accuracy for never-primed and non-studied critical distractor trials; the likelihood of making an accurate response increased over time.
Summary of Chapter 6 Logistic Regressions of Time of lag-2 Primed Target Set, Never Primed Target Set, and Non-Studied Distractors on Task Accuracy.

### Primed

|               | β   | Standard Error | z-ratio | P(>|z|)  | Odds ratio |
|---------------|-----|----------------|---------|----------|------------|
| Time (Trial Number) | 0.003 | 0.001          | 1.775   | 0.080**  | 1.003      |
| Intercept (Constant) | 1.328 | 0.415          | 3.198   | 0.001**  | 3.775      |

Model $\chi^2 = 3.0598$ (df=2), $p = 0.080$

$n = 375$

Significance codes. 0 '***' 0.001 *** 0.01 '**' 0.05 '*' 0.1 ' ' 1

Notes. $n = 21$ (observations = 375). The dependent variable in this analysis is accuracy, coded so that 0 = inaccurate and 1 = accurate.

### Never-Primed

|               | β   | Standard Error | z-ratio | P(>|z|)  | Odds ratio |
|---------------|-----|----------------|---------|----------|------------|
| Time (Trial Number) | 0.006 | 0.001          | 4.117   | <0.001***| 1.006      |
| Intercept (Constant) | 1.265 | 0.315          | 4.012   | <.001**  | 3.545      |

Model $\chi^2 = 10.2$ (df=2), $p = 0.001**$

$n = 381$

Significance codes. 0 '***' 0.001 *** 0.01 '**' 0.05 '*' 0.1 ' ' 1

Notes. $n = 21$ (observations = 381). The dependent variable in this analysis is accuracy, coded so that 0 = inaccurate and 1 = accurate.
Generalized mixed-effects logistic regressions evaluating the impact of trial number on the likelihood of inaccurate responses, separately for primed, never-primed, and non-studied distractor image trials, were performed (depicted in Figure 18 and summarized in Table 7). Overall, there do appear to be practice effects as accuracy for non-studied distractor trials significantly increases over time ($\beta = 0.004$, $p = 0.008$). With this in mind, I can interpret the model summaries for primed and never primed distractor trials. According to the priming account, accuracy over time for primed distractor trials (i.e., the slope of the primed distractor line in Figure 18) should decrease; however, there was no statistically significant change in accuracy: Numerically, accuracy actually increased relative to non-studied distractor trials (primed: 8.45% vs. non-studied: 7.25% increase). As well, accuracy should remain consistent over time for never-primed distractor trials, yet accuracy significantly increased over time, and this increase was numerically larger than that of non-studied trials (i.e., than the increase afforded by factors like practice effects; never-primed: 17.25% vs. non-studied: 7.25% increase). Together, these analyses are consistent with an account in which both sets of items initially capture attention (i.e., accuracy for never-primed distractor trials initially looks like accuracy for

Table 7. Logistic regression of time for lag-2 primed target set distractor (Top), never-primed target set (Middle), and non-studied (Bottom) distractor trials in Chapter 6. Time was a significant predictor of accuracy for never-primed target set and non-studied distractor trials only.

|               | $\beta$ | Standard Error | $z$-ratio | $P(>|z|)$ | Odds ratio |
|---------------|---------|----------------|-----------|-----------|------------|
| Time (Trial Number) | 0.004   | 0.001          | 2.640     | 0.008**   | 1.004      |
| Intercept (Constant) | 2.095   | 0.308          | 6.791     | <0.001*** | 8.127      |

Model $\chi^2 = 6.941$ (df=2), $p = 0.008**$

$\text{n} = 756$

**Significance codes.** 0 ‘***’ 0.001 ‘***’ 0.01 ‘**’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Notes.** $\text{n} = 21$ (observations = 756). The dependent variable in this analysis is accuracy, coded so that 0 = inaccurate and 1 = accurate.
primed distractor trials); however, capture by never-primed items decreased over time (i.e., by the end of the experiment, accuracy for never-primed distractor trials looks more like accuracy for non-studied distractor trials).

Unfortunately, this interpretation of the impact of priming conditions relative to baseline (i.e., increases in accuracy due to practice effects) is limited by inferential statistics. To fully interpret whether the primed and/or never-primed distractor trial regression lines differ from the non-studied distractor trial regression line, a regression model evaluating the interaction between each pair of distractor types is necessary. These regressions were performed; however, neither of the interaction terms were statistically significant (summarized in Table 8). Therefore, these analyses provide no evidence in favour of a flexible top-down account. Although the statistical analyses reported on the interactions between distractor type conditions (i.e., the null results of the distractor type interactions) limit our confidence in the support for a flexible top-down account of the implementation of ACSs, the increases in accuracy over time for never-primed distractor trials provide preliminary evidence that never-primed items lose their ability to capture attention over time. Regardless of this speculative flexible top-down account, however, the data at present better conform to a model of ACSs in which ACSs are implemented through voluntary control: both never-primed and primed items (Chapter 6), whether infrequently or frequently primed (Chapter 7), capture attention. While the average accuracy analyses of Chapter 6 seem to suggest longer-lasting priming is an important factor, given that primed distractors significantly modulated contingent capture, there is a lack of support for the influence of priming on capture both in the timecourse analyses of Chapter 6, and the average and timecourse analyses of Chapter 7. It remains an open question, however, as to whether ACSs that have been implemented through voluntary control can also be flexibly refined over time, as well as whether priming does, in fact, contribute to the inclusion of items into an ACS.
## Primed vs. Non-Studied

|                      | β  | Standard Error | z-ratio | P(>|z|) | Odds ratio |
|----------------------|----|----------------|---------|---------|------------|
| Time (Trial Number)  | 0.003 | 0.001 | 3.145   | 0.002** | 1.003      |
| Distractor Type      | -0.538 | 0.181 | -2.977  | 0.003** | 0.584      |
| Time x Distractor Type | -0.001 | 0.001 | -0.766  | 0.444   | 0.999      |
| Intercept (Constant) | 1.710 | 0.300 | 5.692   | <0.001*** | 5.530      |

Model $\chi^2 = 54.881$ (df=2), $p < 0.001***$

$n = 1131$

**Significance codes.** 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Notes.** $n = 21$ (observations = 1131). The dependent variable in this analysis is accuracy, coded so that 0 = inaccurate and 1 = accurate.

## Never Primed vs. Non-Studied

|                      | β  | Standard Error | z-ratio | P(>|z|) | Odds ratio |
|----------------------|----|----------------|---------|---------|------------|
| Time (Trial Number)  | 0.004 | 0.001 | 3.804   | 0.001*** | 1.004      |
| Distractor Type      | -0.268 | 0.130 | -2.065  | 0.039*  | 0.765      |
| Time x Distractor Type | 0.005 | 0.008 | 0.659   | 0.510   | 1.005      |
| Intercept (Constant) | 1.853 | 0.269 | 6.881   | <0.001*** | 6.382      |

Model $\chi^2 = 21.985$ (df=2), $p < 0.001***$

$n = 1137$

**Significance codes.** 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Notes.** $n = 21$ (observations = 1137). The dependent variable in this analysis is accuracy, coded so that 0 = inaccurate and 1 = accurate.
Table 8. Logistic regression of time for lag-2 primed target set vs. non-studied (Top), and never-primed target set vs. non-studied target set (Bottom) distractor trials in Chapter 6. Time by distractor type was not a significant predictor of accuracy in either test.
INTERIM SUMMARY: CHAPTERS 6 AND 7

Chapters 6 and 7 contrasted two accounts of how ACSs are implemented: the top-down account and feature-priming account. Chapter 6 started by noting that the feature priming account is difficult to reconcile with the results of Chapters 1 to 5. In particular, in the Chapter 1-to-5 studies, inter-trial feature priming was almost always equivalent for ACS matching and non-matching cues/distractors, yet participants successfully adopted ACSs that only allowed matching stimuli to capture attention. Thus, implementing this ACS appears to be possible without relying on inter-trial feature priming. To build on this finding, Chapters 6 and 7 investigated a broader account of priming: If inter-trial priming of low-level features cannot account for the implementation of ACS in the present studies, perhaps longer-lasting priming of higher-level object representations can. To address this broader account, the presence and frequency of priming of target set items was manipulated, and the impact of these priming manipulations on the presence and magnitude of capture was measured, again through the spatial blink. Specifically, in Chapter 6, half of the target set items were never actually presented as targets in the RSVP attention task and were therefore never primed, and in Chapter 7, half of the target set items were primed frequently (75% of trials) while the other half were primed infrequently (25% of trials). Although the presence of priming modulated contingent capture, the frequency of priming had no statistically significant impact on capture: As long as items appeared as targets sometimes, no matter how frequently, they captured attention more than never-primed items.

Secondary time-sensitive analyses started to paint a somewhat different picture of the role of feature priming in the implementation of ACSs. Since the spatial blink measure is an averaged measurement of performance across an entire experimental session, it cannot reveal how capture may develop over time: If priming accumulates over time I ought to observe increases in the amount of attentional capture across the experimental session. Therefore, in both chapters, the likelihood of capture (indirectly measured as the likelihood of making an incorrect response) was evaluated, with the frequency and recency of priming as predictors. In both Chapters, neither factor was a significant predictor of priming. In fact, when the impact of time on capture for primed vs. non-primed target set items is evaluated (time being the only timecourse measure that
applies to both primed and never-primed items), the data point towards a speculative alternative flexible top-down account in which experience responding to particular targets and not others refines an already established top-down set; that is, priming may help to reinforce items as belonging to the ACS, rather than contributing to the initial implementation of the ACS. This was evidenced by consistent accuracy over time for primed target set items, and increased accuracy over time for non-primed target set items. That is, over time, non-primed items lost their ability to capture attention, suggesting they may have been removed from the ACS following the experience that these items are consistently absent from the task (i.e., their status as ACS items is never reinforced). However, these conclusions are confounded by additional factors influencing accuracy over time (i.e., practice effects) and will require further inquiry. At present, the experiments in Chapters 6 and 7 indeed provide support for a voluntary, top-down account of the implementation of ACSs, while pointing towards open questions regarding our understanding of ACSs in general. Specifically, the maintenance of ACSs may be more flexible in response to changing environmental demands than previously acknowledged by both pure top-down control and pure priming accounts.
GENERAL DISCUSSION

SUMMARY

The current thesis aimed to extend our understanding of the ACSs that constrain spatial capture by focusing on the flexibility of ACSs, both in the types of representations by which ACSs are specified and in evaluating the role of top-down control in the implementation of ACSs. Across all chapters, this was accomplished by requiring participants to memorize a set of complex, naturalistic images of objects that exceeded the capacity of visual working memory, a task that relies on episodic LTM storage. These objects were then designated as participants’ targets in subsequent tasks used to measure spatial capture.

The first aim of this thesis (Chapters 1 to 5) was to determine whether ACSs can be specified based on multiple items stored in episodic LTM. Chapters 1 to 4 utilized variants of the attentional cueing paradigm used for assessing spatial capture, and found that, in line with classic contingent capture studies, only cues that matched participants’ studied target items captured attention, as measured by both standard cueing effects and inhibition of return, hallmark measures of spatial capture. Chapter 5 then used a standard episodic source manipulation and an RSVP stream paradigm to assess capture via the spatial blink effect, thus demonstrating that episodic LTM ACSs can be recollection-based. These chapters provide multiple pieces of converging evidence supporting the notion that ACSs can indeed be specified by recollection-based episodic LTM representations, thus extending the representational flexibility of ACSs; rather than constricted to individual items held online in visual working memory or even semantic representations, ACSs can draw on the vast amount of behaviourally relevant and context-rich representations stored in episodic memory.

The second aim of this thesis (Chapters 6 and 7) was to further evaluate the flexibility of ACSs by addressing an on-going debate over the role of top-down control vs. feature priming in how ACSs are implemented. Across both chapters, the RSVP paradigm used in Chapter 5 was combined with manipulations of feature priming. Here, manipulations of priming had a limited impact on measures of spatial capture, which itself was still contingent on ACS-matching items;
the results from Chapters 6 and 7 conform to a flexible, voluntary account of ACS implementation, and present us with important open questions regarding the influence of selection history on the maintenance and refinement of ACSs over time. The findings across all chapters contribute to the on-going specification of the nature of ACSs, in particular, flexibility in the representations by which ACSs are specified and in the implementation of ACSs, while highlighting open questions over the role of selection history in the maintenance and refinement of ACSs over time.

**SPECIFYING THE MECHANISMS OF ACSs I: THE INTERFACE BETWEEN EPISODIC LONG-TERM MEMORY AND ATTENTION**

How might the observation of episodic LTM ACSs be reconciled with the body of work demonstrating that ACS items are stored in working memory? Relevant to this question is work by Oberauer (2002) investigating the relationships between attention, working memory, and LTM. This work provides evidence for three states of representation in guiding processes of visual attention: the focus of attention (a single, active representation in working memory), direct access (all other representations in capacity-limited working memory), and activated LTM (a large capacity subset of task-relevant LTM representations). The notion of activated LTM suggests that information in LTM can be held in a background, passive state as compared to information being held directly online in working memory, but is distinguished from the remainder of LTM in that it is held at a higher level of activation, and therefore can still act to rapidly influence behaviour (i.e., when relevant items are encountered and produce familiarity signals) (Miyake & Shah, 1999; Oberauer, 2002). Representations in activated LTM can prove necessary when task-relevant information exceeds the capacity of working memory, but needs to remain in an easy to access state. An ACS for a set of information in episodic LTM may exist in this activated LTM state, task-relevant and available to influence behaviour with working memory acting as a conduit between activated LTM and the familiarity signals produced when ACS items are encountered in the environment (i.e., as the target on a given trial, or as a studied cue/distractor). Encountered items can then be brought into the focus of attention for comparison and subsequent decisions involving the allocation of attention. Under this model, an LTM-based
ACS must therefore rely on working memory as a comparative stage for incoming perceptual information to determine task-relevance. More recent work in visual search (Carlisle & Woodman, 2013; Cunningham & Wolfe, 2014; Drew, Boettcher, & Wolfe, 2016; Drew & Wolfe, 2014; Wolfe, 2012), however, suggests that this critical role of visual working memory may in fact be quite limited, particularly for LTM representations. Instead, direct connections between perceptual and activated LTM areas serve to support the use of task-relevant information in LTM for guiding search (Drew et al., 2016). This is evidenced by search times in hybrid searches (i.e., visual searches for sets of target stimuli stored in LTM) for target sets that contain information either within or exceeding the capacity of visual working memory. If search is reliant on visual working memory, there should instead be significant increases in search times for LTM target sets, given the need to inefficiently bring all items into a capacity-limited active visual working memory state for comparison; however, search times for LTM sets increase logarithmically with set size increases. Of interest, Carlisle and Woodman (2011) found that when a search target remains the same across consecutive search trials, the associated contralateral delay activity dissipates, eventually disappearing by about the seventh trial. This finding demonstrates that, over time, visual-search target templates are transferred out of visual working memory and likely into some form of LTM. Therefore, LTM can support ACSs without invoking the contribution of visual working memory, thus demonstrating that although ACS items can be stored in visual working memory, as evidenced by a number of past studies, it is not necessary for ACS items to be stored in visual working memory (i.e., both candidate memory systems are valid and compatible).

**SPECIFYING THE MECHANISMS OF ACSs II: ACSs REQUIRE MORE THAN STORAGE**

Relevant to the above discussion is the question of whether the storage of target items in some form of memory, be it visual working memory, semantic LTM, or episodic LTM, although necessary, is sufficient for guiding attention (i.e., in the form of an ACS). We can again look at the work on visual working memory-based ACSs to begin to explore the answer to this question. Earlier work on visual working memory-guided capture, by virtue of demonstrating that task-irrelevant items matching the contents of visual working memory selectively captured attention,
began to converge on the idea that perhaps working memory acts as the source of top-down attentional control. Specifically, items held in visual working memory are responsible for producing the top-down signals that bias visual-spatial orienting towards visual working memory-matching, but not non-matching, stimuli in the environment (Soto et al., 2008). Support for this view comes from both behavioural and neural evidence demonstrating that one of the core functions of working memory is to bias attentional resources, a function that is supported by activity in pre-frontal cortex (Chelazzi, Perlato, Santandrea, & Della Libera, 2013; de Fockert & Theeuwes, 2012; Miller & Cohen, 2001).

The aforementioned work thus raises the question: Is the storage of items in visual working memory equivalent to maintaining an internal goal for those items? According to more recent work evaluating visual working memory-based ACSs, these are not equivalent states (Olivers & Eimer, 2011; Woodman et al., 2013). In fact, cognitive control mechanisms beyond visual working memory are critical in the maintenance of ACSs. In 2011, Olivers and Eimer used a version of the additional singleton paradigm in which participants were required to hold an item in visual-working memory for later recall, after completing a visual search task for a consistent target (i.e., allowing for the memory item to remain in the active state) in which a singleton distractor could either match the item in memory or not. Here, however, the order of these tasks (search and memory test) was randomized in half of the blocks (i.e., mixed blocks vs. fixed blocks) such that participants were unable to predict on any given trial which task would occur first. On trials in which the memory test occurred before search, selective memory-matching distractor costs (i.e., the marker of selective attentional capture) were not observed; however, when search occurred first, selective memory-matching distractor costs were in fact, present. Critically, the strength of this capture effect was larger in mixed blocks in which the use of the memory item could be immediate on any given trial. These results were used to conclude that a) items must be active in visual-working memory in order to bias attention (i.e., search has to occur first, otherwise memory items can be forgotten and thus are no longer active) and b) items in visual-working memory only bias attention when they are task relevant (i.e., could require use immediately as in mixed blocks where task order is uncertain), and thus require an element of control beyond storage in order to influence attention.
Converging evidence for the conclusion that storage of target items is not equivalent to having an ACS for target items comes from ERP work by Carlisle and Woodman (2013) in which memory-matching distractors in the same type of visual search task only modulated the N2pc component, a marker of selective attention, when task relevant, again highlighting the requirement of some level of active control. These conclusions are further supported by dual-state models of working memory that contrast active and accessory states of visual-working memory representations (i.e., the active representation is the one that has interfaced with executive control to be brought in a position to bias attention) (van Moorselaar et al., 2014), thus pointing towards a critical role for executive control processes, the higher-level mechanisms responsible for goal maintenance and updating, in the establishment and maintenance of ACSs.

It is possible that if executive control is critical for distinguishing which visual working memory representations are part of ACSs, its role may also extend to ACSs specific by representations in other memory system, including episodic LTM ACSs. An episodic LTM ACS held may be the consequence of an executive control-based mechanism that acts to discriminate ACS representations (i.e., goal-relevant memories that should guide capture through direct and rapid comparison of perceptual representations to memory representations) from the remainder of representations in episodic LTM. In fact, the conclusions from Chapters 6 and 7 (i.e., that voluntary control seems to best explain how ACSs are implemented) may indeed provide support for the notion that some form of executive control is required for separating goal-relevant episodic LTM representations from the rest of the representations stored in episodic LTM and bringing these goal-relevant representations into a state that allows them to guide capture.

**IMPLICATIONS FOR THE IMPLEMENTATION OF ACSs**

Arguments against the role of top-down control in implementing ACSs (i.e., feature priming accounts) are at odds with the aforementioned importance of executive control processes in guiding stimulus-driven attention. The inter-trial priming account of ACSs has recently emerged as an alternative to a top-down control account of the implementation of ACSs (i.e., Belopolsky et al. (2009) emerged less than a decade ago). Given that the majority of work on ACSs focuses on visual working memory as the memory system responsible for the storage of
the targets the comprise ACSs – targets that also happen to be singular and feature-based (e.g., a colour, a shape, singletons, onsets) – the focus on feature-based ACSs stored in working memory has transferred into questions of priming. Feature-based attention can be defined as goal-directed tuning of processing resources towards particular features, in order to efficiently and effectively bias selection towards such features for further processing, while processing of irrelevant features is suppressed. This form of attention supports goal-directed behaviour in complex visual environments. Seminal work by Theeuwes (2013) even goes so far as to claim that the mechanisms of feature-based attention, including ACSs, can be explained almost entirely by inter-trial priming. This is supported by Belopolsky et al’s (2009) work demonstrating a lack of selective capture with trial-by-trial changes in the task target in a cueing paradigm, as well as inter-trial priming from the target on trial \( n-1 \) to the cue on trial \( n \), in the absence of a target template (i.e., capture is selective to primed cue images). Further support for a feature-based priming account comes from work on oculomotor capture (Silvis, Belopolsky, Murris, & Donk, 2015) in which task-irrelevant colours stored in visual working memory biased saccade-based or oculomotor selection of targets in a search task, as measured by selection accuracy across saccade latency bins, when they matched versus did not match properties of the search array (either targets or distractors). This work highlights the importance of considering feature-based priming across capture paradigms, even suggesting that priming is critical to visual working memory-driven capture effects.

Despite the evidence in favour of priming accounts, the role of inter-trial feature priming as critical for selective capture has been brought under question before. Of note, Lien and colleagues (Lien et al., 2010; Lien et al., 2014) have questioned the entire framing of the paradigm used by Belopolsky et al. (2009). Rather than non-selective capture when targets change on a trial-by-trial basis being the result of a lack of inter-trial priming, Lien and colleagues have taken these data as an opportunity to question the strength or boundaries of ACSs. Perhaps ACSs are still implemented in a top-down manner but “break” under certain conditions (i.e., in ever-changing environments, or when changes in the environment require high levels of control or are more difficult to adapt to), and it is in these conditions that priming may play a prominent role in determining the types of stimuli that capture attention. In a first set of experiments (Lien et al., 2010), participants searched for colour targets in a spatial cueing
task; the target colour could switch or repeat from one trial to the next. Despite the presence of general task-switch costs (overall slower RTs on switch trials), cues still only captured attention when they matched the target colour (i.e., participants’ ACS). Notably, inter-trial priming effects were absent. However, in more recent work (Lien et al., 2014) in which the switch in the target-defining feature was across (between singleton and feature search modes) rather than within dimensions (from one colour to another), arguably more akin to complex and demanding “real-world” environments, capture was no longer selective, rather occurred for both ACS-matching and non-matching cue images. Inter-trial priming effects, though, were again absent (i.e., even in the absence of a strong top-down set, inter-trial priming still did not influence behaviour). This led to conclusions focused on the difficulty in establishing or switching between different ACSs, particularly across target dimensions, as the reason for the absence of contingent capture. Importantly, it is perhaps these difficulties that explain Belopolsky et al.’s (2009) findings: The absence of contingent capture when the target-defining feature switches from trial-to-trial may reflect the inability to rapidly engage in or establish new ACSs, rather than the contribution of priming to the implementation of ACSs. As a consequence, in the absence of an ACS, factors like the target feature on a previous trial may more readily influence capture on a subsequent repeat trial, as in Folk and Remington (2008), where priming was only present when there was no specific target-defining feature (i.e., no foundation for an ACS). Lastly, in their recent review, Lamy and Kristjansson (2013) note that despite the presence of inter-trial priming across a variety of experiments and tasks, when the effects of this selection history are controlled for, voluntary goals still significantly influence the direction of stimulus-driven attention.

The studies in Chapters 6 and 7 are able to significantly weigh in on this important debate over the relative roles of top-down control vs. priming in a novel way through the use of episodic LTM ACSs. In particular, these studies underscore the importance of top-down control in initially implementing control sets, while also highlighting the potential importance of task experience in updating these ACSs in changing environments (i.e., keeping necessary items and removing unnecessary items from ACSs). Specifically, Chapters 6 and 7 highlight a speculative consequence of selection history that may have been overlooked thus far in questions of how ACSs are implemented: Whether responding to particular targets and not others can act to
reinforce and/or refine a voluntary ACS (i.e., determine which items remain in and/or which items are removed from the ACS).

A number of studies have provided evidence that past episodes in a capture task can impact or support contingent capture: Cumulative trial history can encourage flexible, context-dependent use of different ACSs (Cosman & Vecera, 2013a), influence the likelihood of selective capture across item-specific proportion congruent conditions (Thomson et al., 2014), and drive visual statistical learning-based ACSs (i.e., the establishment of an ACS based on implicitly learned contingencies between task stimuli) (Cosman & Vecera, 2014). Perhaps the maintenance of top-down established ACSs is more dynamic and flexible than a strict top-down control account. A flexible top-down model would allow individuals to more easily tailor their ACSs to the types of environments that are reflective of the natural world (i.e., changing environments) in order to better support optimal behaviour in such environments. Given the importance of this ability, further research ought to consider task experience as a critical experimental factor, and examine the impact of task experience on the state of ACSs. More concretely, future work should evaluate the fate of never-primed items that are re-introduced late in the timecourse of an experiment. This would allow for a more direct test of how effective potential refinement of ACSs has been, as well as the conditions under which items that may be removed from ACSs can be re-integrated into these refined ACSs (i.e., is this item truly removed from the ACS, might it be re-introduced into the ACS, and if so how long might this take?).

THE REPRESENTATION OF MEMORY ITEMS IN EPISODIC LTM ACSs: DEPENDENT VS. INDEPENDENT

While the studies presented in this thesis provide the first evidence that participants can adopt ACSs for up to 30 items stored in episodic LTM, it is important to note that evidence in favour of ACSs for multiple objects or features has indeed been previously demonstrated. For example, Admo et al. (2008) found that observers can adopt simultaneous ACSs for two colours, each tied to a separate location in space. Whereas I have described our participants as adopting a single ACS for multiple objects, these past studies typically describe participants as adopting multiple ACSs, each for a unique feature or conjunction. Given the function of episodic
memory, which is to retrieve multiple qualitative details bound together within a single episode (Yonelinas, 2013), it may seem unparsimonious that participants in the studies presented in this thesis adopted separate ACSs for each individual object; however, the prior data favour this interpretation (Adamo et al., 2008; Moore et al., 2014; Moore & Weissman, 2010). The best evidence comes from Moore and Weissman’s (2010) set-specific capture effect: When monitoring an RSVP stream for both red and green targets, although distractors of both colours cause capture relative to non-target coloured distractors, viewing a red distractor improves the accuracy of reporting red targets (and vice versa for green distractors and targets), suggesting stimulus-driven activation of the red ACS (following red distractor capture) over the green ACS. According to Moore and Weissman (2014), both ACSs are maintained in working memory; however, the processing of the distractor feature causes the ACS for that feature to enter the focus of attention (Oberauer, 2002). Critically, the activation of one feature over others suggests some independence of representations when maintaining ACSs for multiple stimuli. While the capacity limitations of visual working memory may necessitate that only a single item stored in visual working memory can enter the focus of attention and thus act as the contents of an ACS, the same may not apply to the potentially parallel state in LTM (i.e., activated LTM. At present, however, it remains an open question whether the contents of LTM-based ACSs are represented dependently or independently of each other (i.e., whether LTM-based ACSs better reflect a single ACS for multiple items, or multiple ACSs of one item each).

SPECIFYING THE TIMING OF CONTROL

Central to our understanding of ACSs is a specification of how ACSs actually exhibit control over attentional capture, if at all. As noted, the contingent capture model postulates that top-down control guides initial selection: ACS-matching stimuli bias selection in their favour. The alternative rapid disengagement model contends that top-down control is exerted later in time, after initial selection: Capture is stimulus-driven, and following capture, top-down control is used to rapidly disengage from stimuli that do not match ACSs. In the case of task-irrelevant distractors or cues in capture paradigms, disengagement for ACS-matching stimuli occurs following an evaluation of their potential relevance and eventual rejection of this relevance (i.e.,
since these stimuli are task-irrelevant): Relative to non-matching stimuli, disengagement from ACS-matching stimuli is delayed (Theeuwes, 2010). In the cueing tasks commonly used to measure contingent capture, a delay period (i.e., SOA) between the onset of the cue and the onset of the target provides, under the rapid disengagement model, enough time for disengagement from non-matching stimuli (i.e., 150 ms; disengagement in general requires 80 – 120 ms Wykowska & Schubö, 2011), but not matching stimuli, thus resulting in selective cueing effects (Theeuwes, 2010). The majority of evidence in support of the rapid disengagement account comes from behavioural, neurophysiological, and neuroimaging studies using the additional singleton visual search paradigm in which a task-irrelevant singleton distractor onsets at the same time as targets, thus preventing disengagement prior to target onset (Theeuwes, 2010). Consistent across these studies is the finding that all salient task-irrelevant stimuli capture attention, regardless of whether they match the target or not.

A number of studies have since come to re-claim contingent capture as the predominantly supported model. For example, when cues are presented at the same time as the target in a cueing task (i.e., 0 ms SOA), cueing effects continue to be contingent on ACSs, while non-spatial distractor costs in the additional singleton distractor paradigm are not (Al-Aidroos et al., 2010; Chen & Mordkoff, 2007; Folk & Remington, 1998; Folk & Remington, 2006). As well, an important criticism of the rapid disengagement work is that the additional singleton paradigm task often encourages singleton search mode, rather than feature search mode (Bacon & Egeth, 1994). Singleton search mode induces an ACS for any singleton stimuli, such that any salient singletons, including task-irrelevant distractors, will capture attention (although even singleton search mode has recently been explained in part by inter-trial priming given the failure to induce a strong feature-based ACS when searching for singletons; Lamy & Kristjánsson, 2013). Debate over the implications of various experimental factors in the behavioural paradigms used to measure capture for models of capture continues, and resolution of this debate holds important implications for our complete conceptualization of ACSs and how they operate to control selection: whether ACSs truly do control initial capture. Although the contributions of this thesis exist independently from this debate and are interpretable under either model (i.e., episodic LTM ACSs can exist under both accounts of selective capture, and the top-down implementation of ACSs does not necessitate top-down control over initial capture), resolution of this debate will
provide important constraints on the timing of retrieval of ACS contents given that retrieval must take place by the time control over capture has been exerted. Specifically, a disengagement model would afford more time for memory processes to unfold and, in turn, regulate capture, whereas under a contingent model, these processes would be constrained to a much shorter timeframe.

**UPDATING OUR UNDERSTANDING OF ACSs**

The studies presented here re-conceptualize our knowledge and understanding of ACSs. While the majority of work on the top-down control of stimulus-driven attention has focused on the primary role of visual working memory as the memory system in which active attentional templates or attentional goals are stored, this view severely restricts the number of representations that can guide attention, limiting them to single, or perhaps at most two, items. It may seem intuitive to think that episodic LTM could constrain stimulus-driven attention in ways similar to working memory representations. Why wouldn’t relevant things in your long-term memory grab your attention unexpectedly? While our intuition, in conjunction with extensive work on the importance of LTM in guiding voluntary attention and evidence for the use of episodic LTM in guiding when and which ACSs are used, support the possibility that episodic LTM can guide the top-down control of stimulus-driven attention, the current thesis represents the first systematic evaluation of this possibility. By extending the representations specifying ACSs to episodic LTM, the flexibility of these representations is in turn much greater than if ACSs were restricted to visual working memory or even semantic memory. Episodic LTM, recollection in particular, is by definition flexible and vast in capacity (Brady et al., 2008; Guild et al., 2014), and this information, from finding articles of clothing in your closet, to recalling the details of different experiments at a dissertation defense, is more often than not directly relevant to our behaviour at any given moment. Beyond flexibility, the extension to episodic LTM highlights the sophisticated nature of ACSs: higher-level representations, such as conceptual categories, semantic knowledge, and now contextually-rich episodic memories can be used to guide stimulus-driven selection in behaviourally relevant ways.
The knowledge that ACSs can be specified by episodic LTM representations in general can support on-going work pushing the boundaries of the role of retrieved episodic details in supporting contingent capture. Under the rapid recollection model of episodic LTM retrieval, perceptual detail is critical in supporting the rapid, obligatory match between sensory information and stored representations (Guild et al., 2014; Moscovitch, 2008). A number of studies highlight the importance of perceptual processing to recollection and the intricate links between the neural substrates of perception and episodic recollection (i.e., contained within the medial temporal lobe) (Graham et al., 2006; Yonelinas, 2002): both visual discrimination (Erez, Lee, & Barense, 2013; Rudebeck, Filippini, & Lee, 2013) and object recognition (Lee & Rudebeck, 2010) are impaired following lesions to the medial temporal lobe. Further specifying the degree to which the amount and particular configuration of perceptual details match between the original encoding episode and the time of retrieval (i.e., when encountering ACS-matching stimuli), can speak to just how rigid versus flexible these episodic ACS representations are, and whether the relative degree of rigidity versus flexibility is modulated by task demands. For example, in environments in which objects often appear in different orientations, will ACS representations be more flexible (e.g., when looking for an article of clothing, you can draw on a single representation in which the clothing is in the particular configuration in your closet, or a number of differing instances such as in your laundry hamper, or on you while looking in the mirror, depending on how certain you are of where and how you last left it)?

In extending the study of ACSs to episodic LTM this thesis serves to bridge traditionally separated research areas within the discipline of cognitive psychology by evaluating attentional capture and episodic retrieval interactions in a novel way, thus adding to our growing understanding of attention-LTM interactions more generally. As well, the study of episodic LTM for the purposes of ACSs also informs the study of the nature of episodic memory retrieval, in particular, the conditions under which episodic memories are rapidly retrieved (i.e., through the implementation of an attentional goal) and the purpose that such retrieval can serve (i.e., to guide rapid, transient, stimulus-driven selection). This work dovetails well with growing literature on the rapid recollection model of episodic LTM retrieval.

The studies presented in the current thesis also serve to address challenges to top-down accounts of control over capture; specifically, alternative explanations of typical contingent
capture findings that instead highlight the role of feature priming. By addressing both an inter-trial priming and longer-lasting priming account, and finding a consistent lack of evidence in support of either as solely responsible for contingent capture, these studies counter pure priming accounts of how ACSs are implemented. Interestingly, Chapters 6 and 7 do highlight a potentially important role for selection history (i.e., refinement of the contents of ACSs following experience attending to and responding to particular targets vs. others). This notion falls in line with modern conceptualizations of the role of top-down versus bottom-up attentional mechanisms. Bottom-up factors such as physical saliency, in combination with top-down factors like internal selection goals, compete to bias visual attention; they do not operate in isolation or in an all-or-none fashion. Instead, evidence suggests that information about physical saliency and internal goals is integrated to form priority maps of visual environments in order to guide competition and ultimately selection (Awh et al., 2012; Corbetta & Shulman, 2002; Desimone & Duncan, 1995). Importantly, these priority maps have been recently shown to require the integration of a third category of influence: selection history (Awh et al., 2012; Theeuwes, 2010, 2013). These selection history factors are not strictly top-down, or bottom-up in nature, but nevertheless have been shown to bias selection, and thus provide evidence for a third category of influence. Thus it is possible that the implementation of ACSs involves these multiple categories of influence.

It is through such work, in combination with the foundations set by this thesis, that we can come to fully outline the relative contributions of selection history, in comparison to the contributions of traditionally acknowledged mechanisms of bottom-up saliency and top-down control. In turn this allows for a more thorough understanding of not only how ACSs are implemented, but how they are maintained over time.

CONCLUSIONS

The studies presented in this thesis converge on the conclusion that ACSs are more flexible in nature than suggested by the current state of the literature: Not only are the representations by which ACSs are specified more flexible (i.e., not restricted to active visual working memory representations, or semantic memory), the implementation of ACSs is flexible
in nature. This is evidenced by the demonstration that having a behavioural goal for representations stored in episodic LTM constrains spatial capture (Chapters 1 to 5) and that predictions of a flexible, top-down account of ACS implementation are supported over the predictions of priming accounts (Chapters 6 and 7), respectively. This work has helped to further develop our understanding of ACSs more generally, and at a more specific level can inform models specifying how the vast amount of information and associated details in episodic LTM can be retrieved on a rapid enough timescale to guide involuntary, stimulus-driven selection in efficient, behaviourally-relevant ways. At a broader level, novel knowledge of the mechanisms of ACSs and the top-down control of attention continues to elucidate how we are able to successfully navigate through complex visual environments, and, of course, why the things that are important to us, like the party-appropriate clothing we need to wear, or the landmarks that will get us to that party, seem to appear when we need them!
REFERENCES


Belopolsky, A. V., Schreij, D., & Theeuwes, J. (2009). What is top-down about contingent


Kiss, M., Grubert, A., & Eimer, M. (2013). Top-down task sets for combined features:


103


Olivers, C. N. L., & Eimer, M. (2011). On the difference between working memory and


