The Nature of Working Memory: Separate, Flexible Resources for Location- vs. Feature-based Representations in Visual Short-term Memory

by

Maha Adamo

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Department of Psychology
University of Toronto

© Copyright by Maha Adamo 2011
The nature of working memory: Separate, flexible resources for location- vs. feature-based representations in visual short-term memory

Maha Adamo
Doctor of Philosophy
Department of Psychology
University of Toronto
2011

Abstract

Working memory, or the ability to maintain and manipulate information in mind when it is no longer physically present, is a pervasive yet severely capacity-limited component of cognition. Visual working memory, also known as visual short-term memory (VSTM), is limited to three or four items on average, with individual differences that range from roughly two to up to six items. Despite agreement that capacity is functionally limited, the current literature is split with respect to the nature of VSTM representations on two key questions: (1) What information is maintained in VSTM? (2) How is information stored in VSTM? The studies presented here address these questions using an event-related potential (ERP) task and a series of behavioural experiments that incorporate attentional selection via retrospective cueing (retro-cues). Experiment 1 manipulated both the number of features and the number of locations to be remembered in a lateralized change-detection task, with differences in the amplitude and topography of the resulting contralateral delay activity (CDA) indicating separate stores for features and locations. Experiments 2a-c established the basic effects of retro-cues on change-detection tasks, showing that attentional selection operated on one system at a time, with overall shorter response times and increased capacity estimates once baseline capacity was exceeded. Experiments 3a-b demonstrated that retro-cues biased VSTM resources to the cued item at the expense of
representational strength of the other, non-cued items, showing flexible reallocation of resources. Experiments 4a-b presented multiple retro-cues to further examine the flexible reallocation of resources in VSTM, showing that capacity benefits depended on spatial specificity of retro-cues and that VSTM weights could be reallocated multiple times before probe comparison. Experiment 5 discounted the potential role of a general alerting mechanism in boosting capacity estimates, showing again that the retro-cue benefit required specificity of the cue. Experiment 6 showed that flexible reallocation of resources within one system did not change the online maintenance of representations within the other system. Thus, the studies collectively address the questions of (1) what and (2) how information is stored by supporting a two-system model of VSTM in which (1) locations and features are stored (2) independently via flexibly allocated resources.
Acknowledgments

This dissertation could not have been completed without the guidance and support of many people, most of all Susanne Ferber. I am very thankful to have had Susanne as my graduate supervisor. Her willingness to take every step with me, her generosity of time, and her genuine interest in my academic and professional success have carried me through my graduate studies. More importantly, her commitment to research excellence has made me a better scientist, and I have garnered many lessons from our discussions throughout the past six years that I will carry forward into my career and beyond.

Many thanks also to Jay Pratt and Adam Anderson for their guidance on my dissertation and other research throughout my graduate studies. I have learned a lot about how to conduct and share meaningful research, not only from their feedback for me but also by example. I am lucky to have had the opportunity to learn from them.

I am also fortunate to have been in the company of the best lab members imaginable, Stephen Emrich, Carson Pun, and Kristin Wilson. Without them, none of this research would have been accomplished, but they have enriched my life far beyond the lab and I am ever grateful for their friendship. I am also lucky to have worked with many talented undergraduates who have contributed to this and other work, most of all Areeba Adnan, Raffles Cowan, Peter Lenkic, and Simon Wozny.

I am deeply grateful to Ann Lang for her support and kindness throughout graduate school.

My sincerest gratitude and affection goes to the friends who have provided constant encouragement and the oft-needed break from work, especially Naseem Al-Aidroos, Leigh Botly, Kate Brown, Hanah Chapman, Karen Emrich, Norman Farb, Tharsni Kankesan, Charles Kassardjian, Dan Lee, Taylor Schmitz, Taryn Sirove, and Ruthann Thomas.

Thanks to my mentor, Frank Haist, for motivation and support before, during, and after.

All love and gratitude to my amazing parents, Jabir and Sarab, and to my beautiful sisters, Reem, Haila, and Luma. I am me because of you.

And to Josh Susskind: ><
# Table of Contents

Acknowledgments...................................................................................................................... iv

Table of Contents..................................................................................................................... v

List of Figures ........................................................................................................................... x

List of Appendices ................................................................................................................... xiv

1 Introduction............................................................................................................................ 1
  1.1 General introduction ........................................................................................................ 1
  1.2 Working memory: An overview ..................................................................................... 2
  1.3 Capacity limits of VSTM ................................................................................................ 3
  1.4 Representations in visual short-term memory ................................................................. 4
  1.5 Neural correlates of visual short-term memory storage .................................................. 8
  1.6 Bridging the gap between fMRI and ERP findings ......................................................... 11
  1.7 Visual short-term memory and attention ....................................................................... 11
  1.8 Retrospective cueing and the nature of representation in VSTM ................................... 15

2 Study 1: Dissociating feature and location load using ERP ................................................ 18
  2.1 Introduction ...................................................................................................................... 18
  2.2 Methods........................................................................................................................... 19
    2.2.1 Participants................................................................................................................ 19
    2.2.2 Apparatus ................................................................................................................ 19
    2.2.3 Stimuli and procedure ............................................................................................. 20
    2.2.4 ERP processing ....................................................................................................... 24
    2.2.5 Analyses .................................................................................................................. 25
  2.3 Results .............................................................................................................................. 25
    2.3.1 Load effects .............................................................................................................. 25
      2.3.1.1 Behaviour ......................................................................................................... 25
2.3.1.2 ERP .......................................................... 27
2.3.2 Feature difficulty effects .......................................................... 32
  2.3.2.1 Behaviour .......................................................... 32
  2.3.2.2 ERP .......................................................... 36
2.4 Discussion .............................................................................. 41
3 Study 2: Establishing basic retro-cue effects for features and locations .............................................................................. 43
  3.1 Introduction ............................................................................ 43
  3.2 Methods ................................................................................ 46
    3.2.1 Participants ....................................................................... 46
    3.2.2 Apparatus ........................................................................ 46
    3.2.3 Stimuli and procedure ....................................................... 46
    3.2.4 Analyses ........................................................................... 51
  3.3 Results .................................................................................... 52
    3.3.1 Exp 2a: Colors .................................................................... 52
      3.3.1.1 Capacity ..................................................................... 52
      3.3.1.2 Response time ............................................................ 55
    3.3.2 Exp 2b: Shapes ................................................................... 55
      3.3.2.1 Capacity ..................................................................... 55
      3.3.2.2 Response time ............................................................ 57
    3.3.3 Exp 2c: Locations ............................................................... 57
      3.3.3.1 Capacity ..................................................................... 57
      3.3.3.2 Response time ............................................................ 59
    3.3.4 Comparing retro-cue effects across stimulus types: Colors vs. Shapes vs. Locations ........................................................ 60
      3.3.4.1 Capacity ..................................................................... 60
      3.3.4.2 Response time ............................................................ 61
3.4 Discussion........................................................................................................................................63

4 Study 3: Manipulating validity to assess relative VSTM weighting..............................................66

4.1 Introduction......................................................................................................................................66

4.2 Methods........................................................................................................................................67

4.2.1 Participants...............................................................................................................................67

4.2.2 Apparatus...............................................................................................................................67

4.2.3 Stimuli and procedure.............................................................................................................67

4.3 Results...........................................................................................................................................68

4.3.1 Exp 3a: 80% validity................................................................................................................68

4.3.1.1 Capacity.............................................................................................................................68

4.3.1.2 Response time....................................................................................................................70

4.3.2 Exp 3b: 60% validity ................................................................................................................71

4.3.2.1 Capacity.............................................................................................................................71

4.3.2.2 Response time....................................................................................................................73

4.3.3 Comparing retro-cue effects across validity rates: 100% vs. 80% vs. 60% ..............73

4.3.3.1 Capacity.............................................................................................................................74

4.3.3.2 Response time....................................................................................................................77

4.4 Discussion......................................................................................................................................78

5 Study 4: Using multiple location cues to examine cue specificity and reallocation of
resources...............................................................................................................................................80

5.1 Introduction......................................................................................................................................80

5.2 Methods........................................................................................................................................81

5.2.1 Participants...............................................................................................................................81

5.2.2 Apparatus...............................................................................................................................81

5.2.3 Stimuli and procedure.............................................................................................................81

5.3 Results...........................................................................................................................................84
5.3.1 Exp 4a: Simultaneous .......................................................... 84
  5.3.1.1 Capacity ................................................................. 84
  5.3.1.2 Response time ....................................................... 87
5.3.2 Exp 4b: Sequential .......................................................... 87
  5.3.2.1 Capacity ................................................................. 87
  5.3.2.2 Response time ....................................................... 89
5.4 Discussion ........................................................................... 90
6 Study 5: Verifying the specificity of retro-cues ............................................ 92
  6.1 Introduction ......................................................................... 92
  6.2 Methods ............................................................................. 92
    6.2.1 Participants ................................................................... 92
    6.2.2 Apparatus ..................................................................... 92
    6.2.3 Stimuli and procedure ................................................ 93
  6.3 Results ................................................................................ 93
    6.3.1 Capacity ....................................................................... 93
    6.3.2 Response time .......................................................... 96
  6.4 Discussion ........................................................................... 96
7 Study 6: Testing whether location retro-cues affect feature resolution .......... 98
  7.1 Introduction ......................................................................... 98
  7.2 Methods ............................................................................. 100
    7.2.1 Participants ................................................................... 100
    7.2.2 Apparatus ..................................................................... 100
    7.2.3 Stimuli and procedure ................................................ 100
  7.3 Results ................................................................................ 103
    7.3.1 Capacity ....................................................................... 103
    7.3.2 Response time .......................................................... 107
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.4 Discussion</td>
<td>108</td>
</tr>
<tr>
<td>8 General discussion and conclusions</td>
<td>110</td>
</tr>
<tr>
<td>8.1 The nature of representation in VSTM</td>
<td>110</td>
</tr>
<tr>
<td>8.2 Summary and integration of findings</td>
<td>110</td>
</tr>
<tr>
<td>8.3 Relation of current results to previous theories</td>
<td>112</td>
</tr>
<tr>
<td>8.4 Future directions</td>
<td>115</td>
</tr>
<tr>
<td>8.5 Conclusion</td>
<td>118</td>
</tr>
<tr>
<td>References</td>
<td>119</td>
</tr>
<tr>
<td>Appendices</td>
<td>124</td>
</tr>
</tbody>
</table>
List of Figures

Figure 2.1 An example of how three simple-feature items were combined into one multi-feature item for Experiment 1. ................................................................. 21

Figure 2.2 Illustration of event sequences for (a) Low load, (b) Location load, and (c) Feature load trials in Experiment 1. ........................................................................................................... 22

Figure 2.3 (a) Accuracy and (b) response time according to load type in Experiment 1. .......... 26

Figure 2.4 Scalp topography for difference waves, shown on the right hemisphere, for (a) Low load, (b) Location load, and (c) Feature load trials in Experiment 1. ......................................................... 28

Figure 2.5 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated over all posterior sites. Displayed waveforms are smoothed across a sliding 50-ms window, and the gray bars represent the memory array and probe onset, respectively. ................................................ 29

Figure 2.6 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated at the topographical peak of the CDA, PO7/PO8. ................................................................. 30

Figure 2.7 Scalp topography for (a) the Location effect, generated by subtracting Low load difference waves from Location load difference waves, and for (b) the Feature effect, generated by subtracting Low load difference waves from Feature load difference waves in Experiment 1. (c) The p-values of the pairwise t-test between the Location and Feature effects plotted as a scalp topography, with areas in blue indicating greater CDA amplitudes for Feature load relative to Location load and areas in red showing the reverse. ................................................................. 31

Figure 2.8 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated at the site where Location load had a greater effect relative to Feature load, P1/P2. ............................................ 33
Figure 2.9 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated at the site where Feature load had a greater effect relative to Location load, P7/P8.

Figure 2.10 (a) Accuracy and (b) response time according to stimulus type in Experiment 1.

Figure 2.11 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated over all posterior sites.

Figure 2.12 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated at PO7/PO8.

Figure 2.13 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated at P1/P2.

Figure 2.14 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated at P7/P8.

Figure 3.1 Illustration of event sequences for (a) retro-cue and (b) simu-cue trials in Experiment 2a: Colors.

Figure 3.2 Illustration of event sequences for (a) retro-cue and (b) simu-cue trials in Experiment 2b: Shapes.

Figure 3.3 Illustration of event sequences for (a) retro-cue and (b) simu-cue trials in Experiment 2c: Locations.

Figure 3.4 Group mean (a) capacity and (b) response time across set sizes in Experiment 2a: Colors, with error bars representing the standard error of each mean.
Figure 3.5 Individual capacity estimates for Experiment 2a: Colors, with participants rank-ordered by baseline capacity and the retro-cue benefit shown added to the baseline capacity estimate. ................................................................. 54

Figure 3.6 Group mean (a) capacity and (b) response time across set sizes in Experiment 2b: Shapes, with error bars representing the standard error of each mean. ........................................... 56

Figure 3.7 Group mean (a) capacity and (b) response time across set sizes in Experiment 2c: Locations, with error bars representing the standard error of each mean. ......................... 58

Figure 3.8 (a) Capacity estimates compared across the three tasks, with and the retro-cue benefit shown added to the baseline capacity estimate, and (b) RT slopes compared across the three tasks in Study 2. ........................................................................................................ 62

Figure 4.1 Group mean (a) capacity and (b) response time across set sizes in Experiment 3a: 80% validity, with error bars representing the standard error of each mean. ....................... 69

Figure 4.2 Group mean (a) capacity and (b) response time across set sizes in Experiment 3b: 60% validity, with error bars representing the standard error of each mean. ......................... 72

Figure 4.3 (a) Capacity estimates compared across Experiments 2a, 3a, and 3b to examine effects of validity rates, with and the retro-cue benefit shown added to the baseline capacity estimate, and (b) RT slopes compared across Experiments 2a, 3a, and 3b. ......................................................... 76

Figure 5.1 Illustration of event sequences for (a) one-cue and (b) two-cue trials in Experiment 4a: Simultaneous. .................................................................................................................. 82

Figure 5.2 Illustration of event sequence for trials with two retro-cues, in this case a first two-cue trial, in Experiment 4b: Sequential. ............................................................. 85

Figure 5.3 Group mean (a) capacity and (b) response time across set sizes in Experiment 4a: Simultaneous, with error bars representing the standard error of each mean ......................... 86

Figure 5.4 Group mean (a) capacity and (b) response time across set sizes in Experiment 4b: Sequential, with error bars representing the standard error of each mean ......................... 88
Figure 6.1 Illustration of event sequences for (a) spatially informative retro-cue and (b) spatially non-informative ready-cue trials in Experiment 5. ................................................................. 94

Figure 6.2 Group mean (a) capacity and (b) response time across set sizes in Experiment 5, with error bars representing the standard error of each mean ................................................................. 95

Figure 7.1 Illustration of event sequences for (a) same-category and (b) cross-category change trials in Experiment 6. ........................................................................................................... 102

Figure 7.2 Correlations between individual capacity estimates for (a) colors on simu-cue trials and shapes on simu-cue, same-category change trials, (b) colors on simu-cue trials and shapes on simu-cue, cross-category change trials, and (c) shapes on simu-cue, same-category change trials and shapes on simu-cue, cross-category change trials in Experiment 6, with fits for the regression lines shown at the top of each scatter plot. .................................................................................. 104

Figure 7.3 Group mean (a) capacity and (b) response time across set sizes in Experiment 6, with error bars representing the standard error of each mean ........................................................................... 106
List of Appendices

Appendix 1 Basic demographic information, including age, sex, and handedness, for all experiments.  ................................................................................................................................................................................................. 124

Appendix 2 Stimuli used in the experiments. (a) All simple shapes, colors, and orientations used in Experiment 1, with (b) samples of multi-feature stimuli created by combining simple features. (c) Simple colored circles used in Experiments 2a, 3a, 3b, 4a, 4b, 5, and 6. (d) Achromatic shapes used in Experiments 2b and 2c (first two columns only). (e) Colored shapes used in Experiment 6; half the participants saw these stimuli, and for the other half the colors were reversed. ................................................................................................................................................................................................. 125
1 Introduction

1.1 General introduction

Working memory is a crucial component of everyday cognition. Commonly characterized as the ability to hold information in mind when it is not physically present, its role in life is pervasive, involved in tasks ranging from the quotidian, such as holding a conversation, crossing the street, or preparing a meal, to the more rare, such as performing complex mathematical operations or making medical diagnoses. It is not surprising, then, that working memory capacity has been linked to general intelligence (Conway, Kane, & Engle, 2003), with evidence that individual estimates of working memory capacity are predictive of success across several domains, such as academic achievement (Cowan et al., 2005), reading ability (Daneman & Carpenter, 1980), language comprehension (Daneman & Merikle, 1996), and reasoning (Gilhooly, Logie, Wetherick, & Wynn, 1993; Toms, Morris, & Ward, 1993). Indeed, some theories of consciousness claim that working memory is essential for awareness, such that we can only be consciously aware of percepts that are available to working memory (e.g., Prinz, 2005). It may seem counterintuitive, then, that working memory is limited to a capacity of roughly three to four items (Cowan, 2001). Despite our rich phenomenological experience, only up to four items can be held in mind at any given time, making a thorough understanding of the nature of working memory capacity and its limits imperative.

The work presented in this dissertation addresses the nature of visual working memory, or the ability to maintain visual information after it is no longer visible. Using an event-related potential (ERP) study and a series of behavioural experiments that incorporate a retrospective cueing technique to examine attentional selection of items held in working memory, the present
findings converge on the thesis that visual working memory stores features and locations independently and flexibly. More specifically, the studies show that information in visual working memory is represented in separate systems that correspond to location and feature content, that resources within these two systems have distinct capacity limitations but can be flexibly allocated, and that the operations within one system are integrated with the other only at the point of behavioural output.

1.2 Working memory: An overview

The notion of “working memory” has been advanced as a more fitting alternative for the concept of a short-term store that bridges initial perception and long-term memory; in particular, working memory allows for a dynamic state in which the activity across a neuronal population reflects the active maintenance of mental representations without necessitating any structural changes to the neurons themselves, as occurs during the consolidation of long-term memories. Baddeley and Hitch (1974) first proposed a multi-component working memory model to replace the then-popular model of short-term memory (Atkinson & Shiffrin, 1968), emphasizing that such a system allows for manipulation of information beyond simple maintenance and thus accentuating its functional role (Baddeley & Hitch, 1994). Baddeley’s model is comprised of two slave systems corresponding to verbal and visuospatial input (the phonological loop and the visuospatial sketchpad, respectively), another slave component linking to long-term memory (the episodic buffer), and a controlling central executive. Although primarily a cognitive model, some aspects of the various components have been mapped to specific brain regions (Baddeley, 2003), and the model has provided a successful theoretical framework for many phenomena, such as the phonological similarity effect, the word length effect, articulatory suppression, and a release from dual-task interference when tasks involve different modalities (Baddeley, 2001; Baddeley & Hitch, 1994). Nonetheless, as Postle (2006) indicates, the number and variety of empirical
findings regarding multiple components in working memory warrant some adaptation of the standard model. As he posits, working memory emerges as a result of sustained activity during retention in the brain areas that represent initial percepts for various modalities, with attention mediating the control of those sustained representations. Rather than attributing working memory functions to any domain-general brain region, this view of emergent working memory holds that input-specific areas are themselves the sites of interest when examining working memory. While Postle’s (2006) view is primarily concerned with the mechanisms that control working memory, this dissertation is well-aligned with his conceptualization of multiple, domain-specific storage systems and makes advances in understanding how those various sources of sustained activity do or do not interact with one another. As the studies here concern the nature of storage within visual working memory, or the short-term maintenance of visual representations, rather than the executive control of working memory, the term visual short-term memory (VSTM) will be used.

1.3 Capacity limits of VSTM

One of the most prominent features of currently accepted models of VSTM is that it has limited capacity, and a corollary to this is the role of individual differences in capacity. Although the precise derivation of capacity limits remains a point of debate, there is general consensus that the limit is approximately four items (Cowan, 2001; Irwin & Andrews, 1996; Luck & Vogel, 1997; Todd & Marois, 2004; Vogel & Machizawa, 2004). In the case of visual change-detection tasks in which participants must determine whether any single element of a memory array has changed after a brief retention delay, such as will be reported here, a simple formula can be applied to estimate capacity based on responses for both matching and non-matching trials. Based on a model initially formulated by Pashler (1988), Cowan (2001) proposed that capacity, \( K \), can be estimated for any given set size, \( S \), using the following formula: \( K = (H + CR - 1)S \), where \( H \) and \( CR \) correspond to the proportion of correct responses for matching (hits) and non-matching
(correct rejection) trials, respectively. The key benefit of applying such a formula, as opposed to using an unscaled proportion of correct responses, is that it better characterizes the point at which measured capacity reaches a plateau, or no longer increases with increasing load. For instance, an individual might respond correctly on 85% of trials in which four items must be remembered but only 68% of trials in which five items must be remembered. While the performance scores expressed in percentage correct seem to indicate that performance declines as set size increases, scaling by the set size indicates that in either circumstance the individual is retaining 3.4 items. This formula has been widely adopted by researchers who employ change-detection tasks to study VSTM capacity (e.g., Alvarez & Cavanagh, 2004; Todd & Marois, 2004; Vogel & Machizawa, 2004) and will be particularly appropriate in the tasks reported here in which one item is probed at a time with equal probability of being a match or a non-match to the corresponding item from the memory array. It is also useful because it allows for an estimate of the maximum capacity that any individual can reach across set sizes, which is crucial to characterizing individual differences in capacity. As Vogel and Awh (2008) discuss, accounting for inter-individual variability can help validate neurophysiological measures, relate seemingly disparate constructs, and dissociate apparently similar constructs. Thus, they promote analyses that incorporate individual differences to better elucidate and constrain cognitive theories. Such an approach is important for the tasks reported here that manipulate capacity in various conditions, as the relationships between these changes and individual capacity may better determine the nature of VSTM storage.

1.4 Representations in visual short-term memory

When discussing the “nature” of VSTM, there are essentially two distinct but interrelated questions that must be addressed, the first being what information is being stored? In other words, what is the basic unit of measurement for VSTM? It could be the individual features
and/or locations that contribute to initial perception, or it could be the whole, integrated objects that are typically the product of conscious report (Duncan, 1984). As will be discussed shortly, current views of VSTM storage have yet to reach some consensus on this first question, which is intrinsically linked to the second question; how is information being stored? The existing literature is also split with respect to this issue, with some support for the idea of “fixed slots” and other evidence for the notion of “flexible resources.” The first camp purports that VSTM capacity is determined by an absolute number of items that can be maintained (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001; Zhang & Luck, 2008) and tends to be aligned with the view that VSTM stores whole objects regardless of feature complexity, while the second contends that resources can be allocated over VSTM representations in a variable manner (Alvarez & Cavanagh, 2004; Bays & Husain, 2008; Wilken & Ma, 2004). This typically translates to the estimated capacity being determined by the complexity of the items to be remembered such that a pool of resources is distributed over the VSTM representations, but a corollary that only recently has begun to be explored is that resources may be unevenly divided across representations, such that some items in VSTM have greater weights relative to others (Barton, Ester, & Awh, 2009; Bays & Husain, 2008).

Studies that support the view that VSTM representations are stored as integrated objects typically find that participants can report several aspects of multi-feature objects as easily as they can report single feature objects. In a series of experiments, Luck and Vogel (1997) presented participants with objects of varying numbers of features, from simple colored squares in one task to bars that each contained up to four features (bar color, bar orientation, gap presence, and gap size), and in all cases found that group performance peaked at roughly four objects. For instance, in one task with bars that each had two properties, color and orientation, Luck and Vogel observed that participants’ performance was equivalent whether only color changes were
reported and only orientation changes were reported; critically when either property could change and participants thus had to retain both features, capacity was comparable to that of reporting either feature alone. Similarly, for four-feature items, performance was roughly equivalent for any one of those features independently and for the conjunction of all features. These findings were verified and extended by Vogel, Woodman, and Luck (2001), and collectively provided evidence that capacity for single features corresponds to capacity for conjunctions, from which the authors concluded that the basic units of VSTM representations are whole objects.

The idea that capacity is constant regardless of feature complexity was challenged by Alvarez and Cavanagh (2004), who used speed of visual search in addition to change-detection performance to catalogue VSTM capacity for a range of simple to complex objects (colored squares, letters, simple line drawings, Chinese characters, grey polygons, and shaded cubes). In their study, they tested the processing speed for items from these varying categories of complexity by first presenting a single item and then measuring the time taken to detect that item in a subsequent array of items drawn from the same category. Alvarez and Cavanagh argued that the increasing response times for more complex objects indexed the greater information load of those objects, and further observed that across categories, information load was inversely related to the set size at which change-detection performance fell to 75%. That is, the more complex the category, the fewer items from that category could be maintained in VSTM at a time. It is important to note that the authors modeled an upper limit to the number of items that could be maintained (assuming items of the lowest information load) at about four or five items (Alvarez & Cavanagh, 2004), suggesting a relative limit based on the amount of visual information, capped by an absolute limit based on the number of items. This could also be characterized as two forms of capacity limitations, one based on feature load and one based on number of objects,
or locations. This is a very important implication which will be revisited at various points throughout this dissertation.

To reconcile the apparent discrepancy between Alvarez and Cavanagh’s (2004) findings and those of proponents of the fixed slots model, Awh, Barton, and Vogel (2007) devised a change-detection task in which they attempted to isolate VSTM capacity for complex objects from comparison errors due to increased similarity between remembered items and the probe item. After first verifying that categories such as Chinese characters and shaded cubes did indeed require longer processing time and typically generated lower capacity estimates relative to colored squares, Awh et al. examined change-detection performance for these more complex categories when the probe was selected from the same category compared to when the probe item was from the other category. For example, a non-matching trial in which a shaded cube changed could involve either a different shaded cube at the probe (within-category change) or one of the Chinese characters (cross-category change). The authors predicted that capacity estimates would be relatively low for within-category changes, consistent with previous findings, because the high degree of similarity between the memory item and the probe would generate a greater proportion of comparison errors. Cross-category changes, on the other hand, would involve probes that were highly dissimilar from the memory items, and in this case they anticipated capacity estimates that would be comparable to those observed for simple objects like colored squares. The findings from Awh and colleagues were consistent with their hypotheses, supporting their theory that VSTM capacity is determined by a fixed number of slots with limited resolution in each slot: While complex items may not be stored with a high degree of resolution, making them prone to comparison errors, an equivalent number of simple or complex items can be maintained (Awh et al, 2007). Follow-up studies from Barton, Ester, and Awh (2009) further specified that the degree of resolution is fixed across all items in VSTM, even
when a memory array is composed of objects with differing complexity. Although this group of researchers makes a firm conclusion that VSTM capacity is limited to a fixed number of slots, with one whole object occupying one slot at a time (Awh et al., 2007), they still concede that change-detection performance ultimately is limited by resolution, or by the relative complexity of the objects in those slots. Indeed, Awh et al. (2007) make reference to a two-system model that incorporates a fixed number of objects in one system and feature complexity in the other (Xu & Chun, 2006; see below). The current thesis maintains that while each of these systems may have its own capacity profile, both of these factors are essential in determining the functional capacity limit of VSTM.

1.5 Neural correlates of visual short-term memory storage

Converging evidence from electrophysiological and neuroimaging data lend support to the notion of multiple systems interacting in visual change detection. Parallel findings from functional magnetic resonance imaging (fMRI; Todd & Marois, 2004) and from event-related potentials (ERP; Vogel & Machizawa, 2004) indicate that regions in the posterior parietal cortex (PPC) show activation profiles that correspond to increases in memory load and that plateau when capacity has been reached within individuals. While the role of frontoparietal networks in working memory maintenance had already been established using fMRI (e.g. Postle, Zarahn, & D’Esposito, 2000), Todd and Marois (2004) were the first to use fMRI to demonstrate that the processing bottleneck that determines the capacity of VSTM storage can be localized to parietal areas. Using a change-detection task with memory arrays containing one to eight colored circles and a single item probe, they observed activity bilaterally in intraparietal sulcus (IPS) during the delay period of the change-detection task that was commensurate with the number of items to be remembered. That is, the degree of activation in IPS increased with increasing memory load and capped out at the same set sizes for which behavioural estimates of capacity (i.e., \(K\)) reached a
plateau. A subsequent study using the same task design further examined the activation profiles of PPC in relation to individual $K$ functions, showing a strong correlation between PPC activity and VSTM capacity across individuals (Todd & Marois, 2005).

Vogel and colleagues have demonstrated much the same effects using event-related potentials and a modified change-detection task in which participants were directed to attend and remember simple colored squares on one side of a bilateral display at a time, allowing the researchers to measure contralateral delay activity (CDA; McCollough, Machizawa, & Vogel, 2007; Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005). They found that the amplitude of the CDA, reflecting greater negativity over contralateral than ipsilateral posterior parietal sites, corresponded to the number of items to be remembered, with the maximum amplitude reached at the same set size at which capacity peaked. Again, CDA amplitude and $K$-estimates were strongly correlated across individuals, and the authors established that the CDA provided a temporal index of the contents of VSTM that was highly responsive to online uploading of items to VSTM (Vogel, McCollough, & Machizawa, 2005). In a later study, Woodman and Vogel (2008) used multi-feature objects (colored bars with different orientations) in a CDA task to show that participants could selectively upload only the relevant feature of such objects: The CDA (as measured over PO1/PO2) had greater amplitude when the change detection was made on the basis of orientation, which was more behaviourally challenging (i.e., yielded slightly lower performance), than on the basis of color. When either color or orientation could change and participants were required to maintain the conjunction of features, the CDA was only as responsive as for the more difficult of the two features, with CDA amplitude increasing with the number of objects in all cases. Woodman and Vogel conclude from these data that participants can selectively upload only relevant features of objects to VSTM, but of equal interest is the conclusion that the CDA indexes the number of objects over the total number of features. From
these studies, it would appear that activity in PPC, a likely locus of VSTM capacity limits, can be tracked using whole objects as the unit of measure.

While the foregoing neuroimaging and electrophysiological studies may seem to favor a fixed slots model of VSTM, a more recent series of fMRI studies demonstrates that activation in PPC reflects both fixed and variable profiles, corresponding to the number of objects and to feature complexity, in distinct areas (Xu & Chun, 2006; Xu, 2007; Xu 2008). Using objects that were either single features or combinations of two features, Xu and Chun (2006) reported diverging activations in distinct areas around the intraparietal sulcus (IPS), such that activity in inferior IPS (iIPS) corresponded to the number of objects to be remembered while activity in superior IPS (sIPS), in addition to a region within the lateral occipital complex (LOC), corresponded to the total number of features present. While capacity estimates were lower for the complex objects than the single features, only the sIPS and LOC reflected the pattern of behavioural output; the iIPS continued to index the number of locations at which information appeared. An additional task presented simple objects simultaneously at different locations, sequentially at different locations, or sequentially at one central location. The fMRI results verified that the iIPS responded primarily to the number of locations at which memory items appeared, with lower activation when objects appeared at one location. Meanwhile, activation in the sIPS continued to increase as the number of different features increased, regardless of location (Xu & Chun, 2006). In subsequent studies, Xu (2007; 2008) showed that the sIPS stores the total number of features to be remembered in a flexible manner, which contributes to object identification, while the iIPS corresponds to object individuation by coding for the locations at which items must be remembered regardless of feature information.
1.6 Bridging the gap between fMRI and ERP findings

The fMRI evidence thus far indicates that two systems are at play in VSTM maintenance, one concerned with features and the other with locations, which is well aligned with suggestions following from behavioural data that capacity may be determined by the integration of two systems (Alvarez & Cavanagh, 2004; Awh et al., 2007). How can this be reconciled with CDA findings to date, which seem to indicate that electrophysiological activity in PPC is predominantly driven by the number of locations with only a secondary effect of feature complexity? The ERP experiment reported in the next chapter addresses this apparent discrepancy, namely by emphasizing that all three features of compound objects be maintained independently and juxtaposing this condition to one in which the three features are presented at separate locations and to one in which single features are presented singly. To preview the key findings, some sites reflect primarily the number of features maintained independent of the number of locations and some correspond primarily to locations, while others index both the feature load and the number of locations at which information appears. This study provides the ERP equivalent of Xu’s (2007; 2008; Xu & Chun, 2006) fMRI work showing dissociable VSTM stores in PPC. Nonetheless, this evidence for separate stores for features and locations does not unequivocally determine whether resources within those stores are flexibly allocated or restricted to a certain number of fixed slots, or how those separate stores are integrated. To achieve this, the later chapters report a series of behavioural experiments that take advantage of retrospective cueing, through which the nature of VSTM representations is elucidated via attentional selection.

1.7 Visual short-term memory and attention

The idea that selective attention and working memory interact is certainly not novel. Some theories of working memory posit that only items within the current focus of attention can be
accurately reported (e.g., McElree, 2001; Oberauer, 2002), and that attention is responsible for making representations accessible to working memory and thus available for conscious awareness (Prinz, 2005). A prominent theory of VSTM maintenance concerns attention-based rehearsal (Awh, Anllo-Vento, & Hillyard, 2000; Awh & Jonides, 2001; Awh, Jonides, & Reuter-Lorenz, 1998; Awh, Vogel, & Oh, 2006), whereby representations are sustained by internally directing spatial attention to the locations at which the items had appeared, with disruptions to that attentional selection resulting in diminished VSTM performance. Still more studies show that the functional and neural bases of working memory and attention are at least partially overlapping (Corbetta, Kincade, & Shulman, 2002; Fougnie & Marois, 2006; Jha, 2002; Johnson, Hollingworth, & Luck, 2008; Jonides, Lewis, Nee, Lustig, Berman, & Moore, 2008; LaBar, Gitelman, Parrish, & Mesulam, 1999; Magen, Emanouil, McMains, Kastner, & Treisman, 2009; Smith & Jonides, 1997), and that the contents of VSTM affect attentional selection (Desimone, 1996; Makovski & Jiang, 2008; Olivers, Meijer, & Theeuwes, 2006; Park, Kim, & Chun, 2007; Soto, Hodsoll, Rotshtein, & Humphreys, 2008). Attentional selection, likewise, can play a role in VSTM maintenance. Schmidt, Vogel, Woodman, and Luck (2002) demonstrated that providing cues for perceptual input results in preferential transfer to and storage of those items in VSTM, which is informative for how items may be prioritized for storage but does not particularly elucidate the nature of representations once those items have gained access to VSTM.

A more beneficial approach is to examine how attentional selection influences items that are already maintained in VSTM, using the outcome of such selection as an indicator of the flexibility of representations. This sort of retrospective cueing (or retro-cueing) has been employed by other research groups. Nobre and colleagues have a now well-established program of research revealing the similarities between attentional selection of perceptual representations
and selection of internal VSTM representations. Starting with the work of Griffin and Nobre (2003), and quickly followed by further studies (Lepsien, Griffin, Devlin, & Nobre, 2005; Lepsien & Nobre, 2006; Nobre, Coull, Maquet, Frith, Vandenberghe, & Mesulam, 2004), this research group has utilized an experimental method referred to as “search in VSTM.” Rather than using a standard change-detection task, in which probe arrays are organized similarly to the initial memory arrays, these tasks follow the memory array with a single probe presented at fixation, necessitating a search through the items held in memory. This process can be compared and contrasted with a perceptual search in which a target item is known in advance of the presentation of a search array. Nobre and colleagues compare the effects of orienting attention to a particular location in the VSTM/perceptual search array, achieved via a centrally presented, symbolic cue that appears during the delay. This method, in conjunction with ERP (Griffin & Nobre, 2003) and fMRI (Nobre et al., 2004), reveals that the neural mechanisms supporting search in VSTM are largely overlapping with those for perceptual search, and that the beneficial effects of the retrospective cue attenuate the typical effects of increasing the search size (Lepsien et al., 2005). Another fMRI experiment using stimuli known to activate specific perceptual processing areas (i.e., faces and scenes) demonstrated that a retrospective cue altered the amount of activity supporting a representation already in memory (Lepsien & Nobre, 2007). Meanwhile, follow-up ERP studies have provided additional, more sophisticated evidence that spatially informative, retrospective cues bias VSTM representations such that a component that reflects the difficulty of search is diminished following cues (Nobre, Griffin, & Rao, 2008), and that the spatial layout and attentional selection of items maintained in VSTM is equivalent to the spatial organization of perceptual search (Kuo, Rao, Lepsien, & Nobre, 2009). These studies cumulatively suggest that attentional processes operate on VSTM representations in much the same way that they would influence perceptually available information; however, they are
relatively agnostic with respect to questions concerning the fixed or flexible nature of VSTM representations, with the exception that attention can be flexibly allocated within visual working memory.

The change-detection approach taken in this dissertation more closely resembles another series of studies using retrospective cues, in which responses are dependent both on locations and on features (Makovski & Jiang, 2007; Makovski, Sussman, & Jiang, 2008). Makovski and Jiang (2007) presented an array of one to six colored circles and used a single circle probe, and participants determined whether the probe was the same or different color as the circle that had occupied the same position in the memory array. They found that change detection was improved when a central arrow cue retrospectively indicated the location that was about to be probed, particularly when this cue appeared before the onset of perceptual masks that interfered with performance. The authors also compared this boost in performance following single cues to change detection when multiple cues were presented either before the onset of the memory array (pre-cues) or during the retention interval (retro-cues). While boosts in performance on pre-cue conditions were proportional to the number of pre-cues, consistent with the findings of Schmidt et al. (2002), change detection was improved on retro-cue conditions only when one cue appeared. Thus Makovski and Jiang argued that the retro-cue is only effective when attention is focused on a single VSTM representation, and in a subsequent study they proposed that this effect is due specifically to a reduction in probe interference (Makovski, Sussman, & Jiang, 2008). While this line of research begins to approach the issue of how items are maintained in VSTM and how selection may flexibly bias these representations, it does not fully address what aspect of these representations are being selected and how location-based selection integrates with feature-based storage. The studies reported in the later chapters build on the foundation
provided by Makovski and colleagues and more thoroughly establish how location-based and feature-based VSTM systems are affected by retrospective attentional selection.

1.8 Retrospective cueing and the nature of representation in VSTM

In the retrospective cueing experiments presented in this dissertation, both location information and feature information must be integrated in order to complete the probe comparison. However, the single probe item is presented such that the two aspects of the comparison can be teased apart. By default, the probe matches one item from the memory array with respect either to location or to feature. This allows participants to narrow their search to only the relevant item and then indicate whether the other aspect (feature or location, respectively) also matches. The theory that emerges from the various experiments is that providing a retrospective cue allows that first aspect of the probe comparison to proceed within VSTM representations prior to the full appearance of the probe. This sort of attentional selection within VSTM representations could be seen either as favoring a fixed slots account of capacity, wherein the retrospective cue moves the pertinent item into a slot so that aspect of probe comparison is guaranteed to be successful, or it can be taken as evidence for the flexible allocation of a pool of VSTM resources, wherein the cue generates a shift in weights toward the pertinent location/feature. To anticipate the data, the latter model is better supported.

The chapters in the dissertation correspond to the various experimental stages in demonstrating that information is stored in both location-based and feature-based VSTM systems and that attentional selection within either of these systems allows for flexible reallocation of resources within one system at a time that is integrated with the other system at the point of probe comparison. Chapter 2 presents an ERP study (Experiment 1) in which the number of features and the number of locations of items to be remembered are varied independently, the results of
which demonstrate that the CDA indexes features and locations, depending on the specific sites measured. After establishing this basic parallel to fMRI studies that also show two VSTM systems, Chapter 3 presents the initial retrospective cueing tasks (Experiments 2a-c) showing that attentional selection can operate on one system at a time. The experiments reported in Chapter 3 show that capacity estimates are indeed improved following retrospective cues as the point at which this improvement is apparent across individuals is determined by baseline capacity, and that response times are altogether shorter when a retrospective cue allows for the advanced comparison of one aspect of the probe, independent of capacity. Chapter 4 (Experiments 3a-b) builds on this foundation and demonstrates that the retrospective cue biases VSTM resources to the cued item at the expense of representational strength of the other, non-cued items. While weights in a VSTM system are redistributed following a retrospective cue, non-cued items are not completely eliminated from memory. Chapter 5 uses multiple retrospective cues (Experiments 4a-b) to further examine the flexible reallocation of resources in VSTM, showing that spatial specificity is necessary for benefits in capacity but that response time continues to display some degree of bias toward cued representations, and that VSTM weights can be reallocated multiple times before probe comparison. Chapter 6 (Experiment 5) demonstrates that the boost to response time is indeed due to spatial specificity of the retrospective cue rather than to a general alerting mechanism provided by the appearance of a cue. Finally, the thesis culminates in the study presented in Chapter 7 (Experiment 6) showing how the flexible reallocation of VSTM resources within one system does not change the online maintenance of representations within the other system, which are integrated at the point of probe comparison. The findings from all experiments are then aggregated in the concluding chapter, which returns to the issue of the nature of VSTM representation and relates the
argument that capacity is determined by both feature-based and location-based systems to other models of multiple modalities in attention and working memory.
2 Study 1: Dissociating feature and location load using ERP

2.1 Introduction

The opening chapter presented various and somewhat contradictory views of VSTM storage with two important questions still unanswered: what is being stored and how? This chapter will address the first question. While all the theories presented agree that VSTM is constrained by limitations in storage capacity, such that estimated capacity reaches a plateau with increasing load, there is debate as to what exactly defines VSTM load. Load may be determined by the number of locations at which to-be-remembered information appears, or it may depend on the number and/or complexity of individual features that must be maintained at those locations. As noted, recent fMRI studies have demonstrated that both aspects of visual information contribute to activation profiles within dissociable regions in the posterior cortices (Xu, 2007, 2008; Xu & Chun, 2006). Yet the most influential ERP studies on the topic, namely those that have characterized the sustained negativity known as the contralateral delay activity (CDA) whose amplitude corresponds to differential load within individuals and to differences in capacity across individuals, have focused primarily on the role of the number of objects, or locations (McCollough, Machizawa, & Vogel, 2007; Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005).

Woodman and Vogel (2008) showed selectivity for which features were stored in VSTM based on task demands (as opposed to obligatory storage for all features of an object). They instructed participants to remember colors, orientations, or both colors and orientations of items in memory arrays of either two or four items, using the same multi-feature stimuli in all cases. Limiting their analyses to a single electrode pair (PO1/PO2), they found that CDA amplitude was modulated
primarily by the number of items retained but also by task demands, with amplitude for
conjunctions limited by the more difficult of the feature classes (i.e., CDA was equivalent for
orientations only and color-orientation conjunctions, and smaller for colors only). They seem to
discredit a feature-based account of VSTM storage in favor of one in which multi-feature objects
are stored as consolidated wholes when all features are potentially relevant. In order to observe
the effects of feature selectivity, Woodman and Vogel intentionally presented the same multi-
feature stimuli across all task conditions, thus circumventing the issue of varying feature loads in
different conditions. However, that is precisely the issue addressed here: Under obligatory
storage of as many features as are presented, to what degree is VSTM load determined by the
number of features presented and to what extent is it influenced by the number of objects? To
test this, the CDA was measured as an index of active representations in VSTM when three
features of varying difficulty were presented either together in one object or at three separate
locations, allowing a comparison between the effect of feature load and that of location load.

2.2 Methods

2.2.1 Participants

Fifteen participants were tested from the University of Toronto community using procedures
approved by the Ethics Review Board, including the provision of written, informed consent at
the start of the test session. All participants had normal or corrected-to-normal vision and were
not color-blind. Basic demographic information, including age, sex, and handedness, is reported
for all experiments in the Appendix.

2.2.2 Apparatus

The experiment was programmed and displayed using Presentation® software (Version 14.0,
www.neuro-bs.com) running on a desktop PC, with a ViewSonic 21-inch monitor (1600x1200
resolution, 60 Hz refresh rate) placed at a viewing distance of approximately 57 cm. Participants made responses with the right index and middle fingers positioned over the left-arrow and down-arrow keys, respectively, using a standard keyboard placed at a comfortable distance on the table; the spacebar was used to initiate trials. Continuous, unreferenced EEG was recorded at a sampling rate of 512 Hz using a BioSemi ActiveTwo system with 64 scalp electrodes in standard 10-20 placement and an additional electrode at each mastoid, at the outer canthus of each eye, and below each eye.

2.2.3 Stimuli and procedure

All stimuli were presented against a white background with a black, central fixation cross (subtending 1°) present throughout each trial. Left- or right-pointing black arrows (2° wide) were presented 2° above the central fixation to direct attention to either side of the bilateral display. Memory and probe stimuli were one of four colors (red, blue, green, yellow), one of four unfamiliar grey shapes, one of four blocks of oriented grey lines (0°, 45°, 95°, 135°), or combinations of one feature from each category (adapted from “blob” stimuli used by Barense, et al., 2005; see Figure 2.1 for a sample and see Appendix for all features). Stimuli were all scaled to 2.43° by 1.69° and were presented randomly in 4° by 8° rectangular areas centered 5° to the left and right of the central fixation, with the constraint that potential locations were all at least 2° apart.

Each trial was initiated when the participant pressed the spacebar, after which the fixation appeared alone for 300 ms, followed by the presentation of a left or right arrow for 200 ms that directed the participant to attend only that side for the duration of the trial (while maintaining central fixation, see Figure 2.2). After a variable delay from 300 to 400 ms, a memory array appeared containing one or three single- or multi-feature items at random locations on each side.
Figure 2.1 An example of how three simple-feature items were combined into one multi-feature item for Experiment 1.
Figure 2.2 Illustration of event sequences for (a) Low load, (b) Location load, and (c) Feature load trials in Experiment 1.
of the display (see conditions below). The memory array was visible for 200 ms, followed by a delay for 1000 ms, after which a single probe item appeared on each side for 2000 ms, and the participant was to indicate whether the probe was a match or a non-match for the item at that location in the memory array. Each intertrial interval (ITI) was for as long as the participant needed to blink and reset for the next trial.

On Low load trials, the memory array contained one single-feature item chosen from the same feature category on each side (i.e., one color on each side, or one shape on each side, or one orientation on each side, with equal likelihood of each feature class appearing). The probe was then a single-feature item from the same feature class as the memory array and was equally likely to be a match or non-match. On Location load trials, three single-feature items were presented on each side in the memory array, with one feature from each category (i.e., one color and one shape and one orientation, on each side). The probe was a single-feature item from the same feature class as the item that had appeared at that location in the memory array and participants were to indicate with a button press whether or not the probe item exactly matched a previously presented item. The probe item was equally likely to be a match or non-match. On Feature load trials, the memory array contained one multi-feature item comprised of a feature from each category presented on each side (i.e., one particular shape made of a particular orientation of lines in a particular color, on each side). The probe was also a multi-feature object that was equally likely to be a match on all three features or a non-match on any one of the three features (e.g., the shape would change but the color and orientation would be the same). In any case, there were no constraints regarding overlap between features on each side (e.g., a blue stimulus may or may not appear on both sides), and the match/non-match trials were counterbalanced on each side and fully crossed with the side that had to be attended, such that
the decision for the attended side was completely independent of the unattended side (see Figure 2.2). All trial conditions were intermixed and presented in random order.

After obtaining informed consent and preparing for EEG recording, participants viewed all possible stimuli and were informed of the various task conditions. They were instructed to attend only the side indicated by the arrow on each trial while keeping their eyes fixated on the cross at the center of the display. Participants were told that when items changed after the delay, they would differ by only one feature. The emphasis was placed on judging as accurately as possible whether the probe was the same or different, and participants were asked to make a “best guess” on trials for which they were uncertain of the correct response.

2.2.4 ERP processing

EEG datasets were converted into EEProbe format (version 3) using PolyRex (Kayser, 2003). Each individual dataset was filtered using a finite impulse response (FIR) filter high-passed at 0.01 Hz and low-passed at 30 Hz. The continuous data were then re-referenced to the average of the two mastoids. Additionally, lateral eye movements were computed by re-referencing the right outer eye to the left outer eye. Rejection markers were based on activations in excess of 30 standard deviations over a 200 ms sliding window and were generated over all eye channels in addition to frontopolar channels and posterior sites contributing to analyses. Each continuous dataset was averaged to generate individual ERPs time-locked to the onset of the memory array, with the 200 ms prior to array onset serving as the baseline and the rejection window spanning from the arrow onset until probe onset. Average waveforms for each individual were then imported to Matlab (version 7) for analyses using custom scripts, including topographical plots of activations and waveform amplitudes.
2.2.5 Analyses

Two sets of analyses were conducted. The first established how the effects of location and feature load differed by comparing trials according to Load \((Low \text{ vs. } Location \text{ vs. Feature})\), which were expected to elicit differences in CDA amplitude and/or topography such that \textit{Location} and \textit{Feature} load elicited a greater response than \textit{Low} load albeit with different topographical foci. The second analysis examined the role of feature difficulty by comparing \textit{Low} load trials that had different types of single features (color, shape, orientation) and further contrasted these to \textit{Feature} load trials in which a conjunction of features was to be remembered, for which differences in CDA amplitude were anticipated (i.e., larger amplitude with increasing feature difficulty as first determined by the behavioural results). ERP results were generated for all artifact-free trials, irrespective of behavioural performance for each trial, and focused on difference waves over eight pairs of posterior sites (P1/P2, P3/P4, P5/P6, P7/P8, P9/P10, PO3/PO4, PO7/PO8, O1/O2). These were obtained by first mirroring the individual waveforms for trials in which attention was directed to the right and then averaging these right-sided trials with left-sided trials, after which differences were calculated between each left-right pair of electrodes to obtain individual contralateral minus ipsilateral waveforms in each condition that were submitted to group analyses.

2.3 Results

2.3.1 Load effects

2.3.1.1 Behaviour

Figure 2.3 displays the group results for the percentage of trials with correct responses and the associated reaction times (RTs) according to Load. Individual participants’ measures of accuracy (hits and correct rejections) and response times were submitted to separate, one-way repeated measures ANOVAs with the single factor of Load type \((Low \text{ vs. } Location \text{ vs. Feature})\). Load
Figure 2.3 (a) Accuracy and (b) response time according to load type in Experiment 1.
type had a significant effect on accuracy, $F(2,28) = 62.25, MSE = 14.67, p < 0.0001$, with a significantly greater proportion of correct trials for Low load trials compared to both Location and Feature loads, and significantly greater percentage correct for Feature load relative to Location load. The Load effect was also significant for RT, $F(2,28) = 93.11, MSE = 1661.58, p < 0.0001$, with longest RTs for Location load relative to both Low and Feature loads, with significantly shortest RTs for Low load. Overall, Load type affected performance on the change-detection task, with effects for both Location and Feature load.

2.3.1.2 ERP

The contralateral delay activity (CDA) was defined as the mean amplitude of the contralateral-ipsilateral difference wave from 400 to 1000 ms after the onset of the memory array, generated separately for each Load type. Figure 2.4 displays the topography of group mean amplitudes (collapsed onto the right hemisphere). The effect of Load type was examined first by collapsing across all posterior sites (see Figure 2.5 for group average waveforms and mean amplitudes); when submitted to a one-way, repeated measures ANOVA, this revealed an effect of Load type, $F(2,28) = 4.34, MSE = 0.31, p < 0.05$, such that either type of load modulated the CDA: CDA amplitude was not significantly different for Location load and Feature load, but larger for both relative to Low load. When only the topographical peak of the CDA was examined, at PO7/PO8 (see Figure 2.6), the effect of Load type was no longer significant, $F(2,28) = 0.81, MSE = 1.64, p > 0.4$, despite a significant omnibus test showing that any Load type elicited CDA amplitude greater than zero, $F(1,14) = 14.19, MSE = 5.39, p < 0.005$. That is, the CDA was equally present at PO7/PO8 for any level of load in the VSTM task. The different pattern of load dependence according to the site of CDA calculation indicated that the topographical extent of the CDA could show different effects of Load type. Figure 2.7a-b
Figure 2.4 Scalp topography for difference waves, shown on the right hemisphere, for (a) Low load, (b) Location load, and (c) Feature load trials in Experiment 1.
Figure 2.5 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated over all posterior sites. Displayed waveforms are smoothed across a sliding 50-ms window, and the gray bars represent the memory array and probe onset, respectively.
Figure 2.6 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated at the topographical peak of the CDA, PO7/PO8.
Figure 2.7 Scalp topography for (a) the Location effect, generated by subtracting Low load difference waves from Location load difference waves, and for (b) the Feature effect, generated by subtracting Low load difference waves from Feature load difference waves in Experiment 1. (c) The $p$-values of the pairwise $t$-test between the Location and Feature effects plotted as a scalp topography, with areas in blue indicating greater CDA amplitudes for Feature load relative to Location load and areas in red showing the reverse.
shows the isolated effects for Location load and for Feature load obtained by subtracting the Low load condition from each, which further indicated different CDA extent according to Load type. A direct comparison was obtained by calculating a pairwise t-test on Location and Feature difference waves at each site (significance levels are plotted topographically in Figure 2.7c), which highlighted two sites of interest, P1/P2 which showed greater CDA amplitudes for Location load and P7/P8 which showed greater CDA amplitude for Feature load. Additional ANOVAs tested the effects at these sites (see Figures 2.8 and 2.9 for group average waveforms and mean amplitudes). For P1/P2, the Load effect was significant, $F(2,28) = 4.49, MSE = 0.88, p < 0.05$, with the largest amplitude for Location load. For P7/P8, the Load effect was again significant, $F(2,28) = 3.89, MSE = 2.02, p < 0.05$, but in this case the largest amplitude was that of Feature load. This confirmed that the topographical extent of the CDA varied according to the type of load, with discernible differences for Location and Feature load relative to one another and to Low load.

2.3.2 Feature difficulty effects

2.3.2.1 Behaviour

Figure 2.10 displays the group results for accuracy and RT for the different single-feature items (shape, color, orientation) as well as the conjunction of all three (i.e., Feature load trials). Separate, repeated measures ANOVAs examined the effect of stimulus type. A significant effect was observed on accuracy, $F(3,42) = 36.37, MSE = 10.70, p < 0.0001$, such that participants were most accurate for colors, which differed only marginally from shapes. Orientation and conjunctions showed the lowest accuracies, with no significant difference between the two. The same pattern was observed on RT: A main effect of stimulus type, $F(3,42) = 36.54, MSE = 1448.58, p < 0.0001$, with RTs for colors and shapes significantly smaller than those for
Figure 2.8 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated at the site where Location load had a greater effect relative to Feature load, P1/P2.
Figure 2.9 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across load type in Experiment 1, generated at the site where Feature load had a greater effect relative to Location load, P7/P8.
Figure 2.10 (a) Accuracy and (b) response time according to stimulus type in Experiment 1.
orientations and conjunctions. Thus, the stimulus type that had to be remembered determined the
difficulty of the change-detection task, even when only one item had to be held in VSTM. These
behavioural results served as the index of feature difficulty for the ERP analyses.

2.3.2.2 ERP

Guided by the topographical results for the effects of Load type, separate repeated measures
ANOVA analyses were conducted on four sets of mean amplitudes. The CDA collapsed across all
posterior sites (see Figure 2.11 for group average waveforms and mean amplitudes) showed an
effect of feature difficulty, $F(3,42) = 5.92$, $MSE = 0.44$, $p < 0.005$, such that amplitudes ranged
from largest for conjunctions and orientations, followed by shapes, and finally smallest for
colors. When analyzed only at the topographical peak of the CDA at PO7/PO8 (see Figure 2.12),
the effect of feature difficulty was again significant, $F(3,42) = 3.23$, $MSE = 2.44$, $p < 0.05$, here
with the greatest amplitude for orientations, followed by conjunctions and shapes, and the
smallest amplitude for colors. Thus, measured across all posterior sites or just at the peak site,
CDA amplitude was larger for items with greater feature difficulty as determined by the
behavioural results. When examined only at P1/P2 (see Figure 2.13), which was the site of the
Location load effect, feature difficulty had no significant effect on CDA amplitude, $F(3,42) =$
0.55, $MSE = 1.88$, $p > 0.6$. When analyzed at the site of the Feature load effect, P7/P8 (see
Figure 2.14), the effect of feature difficulty also failed to reach significance, $F(3,42) = 2.05$, $MSE$
= 5.27, $p > 0.1$, despite the graphical differences in amplitudes, most likely owing to large
variability in the data (standard deviations were all at least 1.7 µV). Nonetheless, CDA amplitude
overall distinguished among the differing levels of difficulty of items as determined by
differences in behavioural performance.
Figure 2.11 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated over all posterior sites.
Figure 2.12 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated at PO7/PO8.
Figure 2.13 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated at P1/P2.
Figure 2.14 (a) Grand average waveforms and (b) group mean amplitudes, with error bars representing the standard error of each mean, across stimulus type in Experiment 1, generated at P7/P8.
2.4 Discussion

The key finding of the behavioural data is that both location load and feature load impact VSTM performance, with differences depending on the features of the stimuli to be remembered even when a single object was held in memory. This was corroborated by the ERP data, which showed that the CDA was affected by both location and feature loads, with distinct effects in the topographical extent of activation during the retention interval. Thus the behavioural response to different task demands and the activity of parietal areas during VSTM storage were well aligned. While this interpretation is in general accord with the existing literature concerning ERPs during VSTM tasks, it presses further to specify that VSTM load was not determined simply by the absolute number of objects to be remembered, but critically was affected by the number of features to be remembered and by the difficulty of those features, whether difficulty is determined by the relative salience of different feature categories or by the ease of discrimination within a feature class. Furthermore, the present experiment better characterizes the nature of VSTM storage than did the Woodman and Vogel (2008) study by varying the visual load across trials and by taking into account topographical differences for different aspects of visual stimuli (for which inclusion of multiple electrode sites is crucial).

The finding that location and feature loads elicited divergent effects on CDA amplitude was particularly interesting given recent fMRI studies showing distinct posterior areas that correspond either to object individuation or identification (Xu, 2007; 2008). Given the coarse spatial resolution of ERP, the topographies of the CDA for locations and for features were largely overlapping, to the degree that collapsing across all posterior sites exhibiting the CDA showed that either sort of load had comparable effects on amplitude. For that matter, looking only at the peak site of CDA amplitude showed no load-dependent differences, making the case that in order to understand what aspects of visual information contribute to VSTM load, the
distribution of activity must be examined. While the findings of this experiment could not be
mapped directly onto the structures observed by Xu and Chun (2006) given the obvious
challenges with localizing ERP sources, the sites showing statistically reliable differences in
amplitude for location versus feature load corresponded to those areas strikingly well, supporting
the theory that locations and features are stored in separate VSTM systems. It is unclear,
however, when these separate VSTM systems become integrated into a full representation, which
will be addressed in the following chapters.
3 Study 2: Establishing basic retro-cue effects for features and locations

3.1 Introduction

In the previous chapter, the different contributions of feature-based and location-based VSTM systems were characterized electrophysiologically, showing that storage capacity is determined by both the number of features and the number of locations as reflected in the activity of parietal areas. While the evidence that CDA indexes the number of features held in memory is novel, others have argued previously that ultimate performance on a VSTM task should be determined by the lower limit of the two capacity-limited systems (e.g., Awh, Barton, & Vogel, 2007; Xu & Chun, 2006). When simple features are presented in distinct locations, the feature-based and location-based systems have comparable response profiles, and thus the behavioural capacity limit corresponds to activations within either system. When remembering more complex stimuli that are composed of multiple or difficult-to-discriminate features, however, the feature-based system will maximize available resources across a fewer number of items, and this will determine the lower limit for determining capacity. This account is well-aligned with existing behavioural and neural evidence, yet it does not provide an explanation for how/when the two systems are integrated and, more specifically, whether the response profiles within the systems interact online during VSTM maintenance or are simply integrated at the point of behavioural output (i.e., at probe comparison).

By using a change-detection task with a single probe that is by default a match on one aspect of the comparison (i.e., either the same location or the same feature as one item in the memory array) and that is sampled without replacement (i.e., a non-matching probe had not been seen in
the memory array), the contributions of the location-based and feature-based VSTM systems can be disentangled. For instance, a memory array may present various shapes in various locations, with a single shape probe appearing after the delay. If feature identity is being tested, the probe can be presented at the same location as one of the memory items, and the participant must determine whether that probe is the same shape as what had previously appeared at that location. A non-matching probe would be a new shape presented at an old location. Conversely, if location is being tested, the probe can be presented at any location in the display, and the participant must determine whether that shape is presented where it previously appeared. In this case, a non-matching probe would be an old shape presented at a new location. This is the basic principle underlying the experiments reported in this and subsequent chapters, where participants were responding to one of two questions, either (1) “Is this the item that appeared at this location before?” or (2) “Is this the location where this item appeared before?” Essentially, this method breaks the probe comparison down into its two components: In the former scenario, location-matching must be ascertained before any feature comparison can proceed at the location in question, and vice versa for the latter scenario, where the feature must be selected before location can be assessed.

Modifying this basic change-detection paradigm to include a retrospective cue, or retro-cue, provides an optimal method of assessing the effects of one storage system on the other. More precisely, this method allows the examination of how selection within one system affects output by the other. As discussed in the first chapter, cueing a representation within VSTM generates attentional selection that is somewhat comparable to that obtained by cueing a perceptual stimulus, seen as larger estimates of capacity relative to baseline conditions in which no informative cues are provided (e.g., Griffin & Nobre, 2003). For example, Makovski and colleagues (Makovski & Jiang, 2007; Makovski, Sussman, & Jiang, 2008) presented an array of
stimuli, such as colored circles, to be maintained in VSTM, and then presented a single stimulus after a retention interval. While the location of that single probe item was guaranteed to occur in a location that had previously been stimulated in the memory array, participants had to determine whether the feature matched the item that was previously in that location. Responses to this feature probe were more accurate when a retro-cue appeared during the retention interval to indicate which location was about to be probed, relative to when there was no advanced cue to allow attentional selection of the probe location. Note that those studies did not explicitly consider the relative contributions of independent VSTM systems for features and locations. In the change-detection tasks reported here, cues contain either feature information or location information, and the effects of retro-cues on behavioural output are measured both in terms of estimated storage capacity (K) and of correct response times (RT).

This initial study employing the retro-cue paradigm established the basic effects of attentional selection within one VSTM system at a time: The effects of location-based selection on feature VSTM were examined using both simple features (Exp 2a: Colors) and more complex stimuli (Exp 2b: Shapes), and the effects of feature-based selection on VSTM for location were determined (Exp 2c: Locations). Although overall capacity estimates for Colors and Shapes were expected to differ based on existing literature on VSTM for varying complexity (e.g., Alvarez & Cavanagh, 2004), the magnitudes of the retro-cue benefits across the experiments were not expected to differ, as spatial selection within the location-based system was hypothesized to be independent of storage within the feature-based system. Similarly, the overall estimates for Colors and Locations were expected to be comparable, since the capacity profiles for simple features and locations correspond closely to a 1:1 ratio (Xu & Chun, 2006).
3.2 Methods

3.2.1 Participants

Independent samples of 24 participants were tested for each of Exps 2a-c, all recruited from the University of Toronto community and tested under procedures approved by the Ethics Review Board, including the provision of written, informed consent at the start of the test session. All participants had normal or corrected-to-normal vision and were not color-blind. Basic demographic information, including age, sex, and handedness, is reported for all experiments in the Appendix.

3.2.2 Apparatus

All experiments were programmed and displayed using Presentation® software (Version 12.0 or newer, www.neuro-bs.com) running on a desktop PC, with a ViewSonic 21-inch monitor (1600x1200 resolution, 60 Hz refresh rate) placed at a viewing distance of 57 cm. Participants made responses using a standard keyboard placed at a comfortable distance on the table, and responses were made with the right index and middle fingers positioned over the left-arrow and down-arrow keys, respectively.

3.2.3 Stimuli and procedure

All experiments in Study 2 were presented against a black background. Exp 2a: Colors used colored circles that were 1.31° in diameter and filled with brown, red, orange, yellow, green, light blue, dark blue, purple, or pale grey (see Appendix for RGB values). Exp 2b: Shapes used nine unfamiliar white shapes that subtended 1.31° (see Appendix), with a subset of six of these shapes used for Exp 2c: Locations. For the Colors and Shapes experiments, stimuli could be presented in any of six locations evenly spaced along an imaginary circle centered on the display with a diameter of 5°, and cues were white arrows that originated at the center of the display and
terminated at a corresponding location along an imaginary concentric circle with a diameter of 3°. For the Locations experiment, stimuli could be presented in any of 12 locations evenly spaced along a 3x4 grid, with rows 4° below, on, or 4° above the horizontal meridian and columns 3° or 7° to the left or the right of the vertical meridian. Cues appeared within a central box subtending 1.5° at the center of the screen.

On each trial, a memory array with anywhere from one to six different items to be remembered was presented for 1000 ms, followed by a blank delay of 1000 ms. At this point, one of two possible events occurred (see Figure 3.1): In simu-cue trials, a single probe item appeared along with a cue and remained present on the screen until a response was made, which terminated the trial and initiated an inter-trial interval (ITI) of 1000 ms. In retro-cue trials, only a cue appeared for 100 ms after the initial delay, followed by a second delay interval of 400 ms and finally a single probe item, which remained on the display until a response was made and the ITI began (see Figures 3.1-3.3). The probe item was equally likely to be a match or a non-match to the cued item, with non-matching items randomly selected from the remaining items that were not present in the memory array (i.e., items were sampled without replacement). For each experiment, each cue type (i.e., simu-cue and retro-cue) was fully crossed with set size (i.e., between one and six items) and with whether the probe was matching or non-matching, with 25 repetitions for each combination leading to a total of 600 trials which were randomly intermixed.

Following the informed consent process, participants were seated comfortably at the test computer, where they received full instructions and completed a brief practice session with the experimenter present in the testing room, after which the experimenter monitored general compliance from outside the room via a webcam. Participants were introduced to stimuli prior to receiving detailed instructions in order to verify that stimuli were perceptually discriminable
Figure 3.1 Illustration of event sequences for (a) retro-cue and (b) simu-cue trials in Experiment 2a: Colors.
Figure 3.2 Illustration of event sequences for (a) retro-cue and (b) simu-cue trials in Experiment 2b: Shapes.
Figure 3.3 Illustration of event sequences for (a) retro-cue and (b) simu-cue trials in Experiment 2c: Locations.
from one another. No participant in any experiment reported difficulty in discriminating stimuli when all were displayed simultaneously, and all participants were aware of the various task conditions prior to task onset. They were instructed to remember all items presented during the memory array and were told that the cue would always indicate which memory item was being tested, making the cues imperative. In other words, participants were told that the cue was 100% accurate, and thus were advised that the cue could help them when it appeared in advance of the probe. Although cautioned against taking too long to respond, the emphasis was placed on judging as accurately as possible whether the probe matched the cued item, and participants were told to make a “best guess” on trials for which they were uncertain of the correct response. Each experiment was divided into 10 blocks (60 trials per block), and participants were allowed to take a self-timed break between each block of trials.

3.2.4 Analyses

VSTM performance for each task was measured in the following two ways: (1) reaction times (RTs), averaged over all correct responses (i.e., hits and correct rejections) per condition; and (2) estimated storage capacity, $K$, calculated by taking the difference between the proportion of correct responses for match trials and the proportion of incorrect responses to non-matching trials and weighting that by the number of elements to be remembered. Furthermore, each participant’s baseline capacity was estimated as the highest $K$ achieved on simu-cue trials and was used as a covariate to establish the degree to which observed group effects were modulated by individual differences. SPSS was used to conduct repeated measures ANOVAs on RT and capacity estimates within each of Exp 2a-c separately, and mixed model ANOVAs were conducted to compare effects across the three experiments.
3.3 Results

3.3.1 Exp 2a: Colors

3.3.1.1 Capacity

*K*-estimates were generated for each set size and each cue type (see Figure 3.4a for group means), and baseline capacity for each individual was estimated as the best *K* achieved for any set size in the *simu-cue* condition. The average baseline capacity across individuals ranged from 2.08 to 5.04 (mean ± standard deviation = 3.39 ± 0.81). All participants showed improvements in capacity following a *retro-cue* in at least three of the six set sizes, and all but three showed an overall greater capacity for *retro-cue* trials relative to *simu-cue* trials (average improvement in *K* of 0.75 ± 0.12). A repeated measures ANOVA (Cue type X Set size) verified that capacity was overall greater for *retro-cue* trials, *F*(1, 23) = 77.93, *MSE* = 0.16, *p* < 0.0001. There was also a significant effect of Set size, *F*(5, 115) = 99.28, *MSE* = 0.45, *p* < 0.0001, with *K* increasing with each additional item up to a set size of 4 on average, after which capacity reached a plateau. The interaction between Cue type and Set size was also significant, *F*(5, 115) = 11.18, *MSE* = 0.13, *p* < 0.0001, such that the relative benefit of *retro-cue* trials over *simu-cue* trials appeared to increase as the number of items to be remembered increased. Notably, this interaction was modulated by baseline capacity: When baseline capacity was entered as a covariate to the ANOVA, the Cue type X Set size interaction no longer achieved significance, *F*(5, 110) = 0.94, *MSE* = 0.13, *p* > 0.4, demonstrating that the *retro-cue* was effective only once an individual’s baseline capacity had been exceeded. That is not to say, however, that all individuals achieved a similar *K* following *retro-cues* or that an individual’s baseline capacity determined how effective the *retro-cue* was overall, as can be seen in the individual estimates presented in Figure 3.5.

Although there was no correlation between baseline capacity and the average benefit following a *retro-cue* (*r* = -0.04, *p* > 0.8), *K*-estimates for the two cue types were highly correlated (*r* = 0.80,
Figure 3.4 Group mean (a) capacity and (b) response time across set sizes in Experiment 2a: Colors, with error bars representing the standard error of each mean.
Figure 3.5 Individual capacity estimates for Experiment 2a: Colors, with participants rank-ordered by baseline capacity and the retro-cue benefit shown added to the baseline capacity estimate.
indicating a relatively constant benefit for retro-cue trials once capacity was exceeded.

3.3.1.2 Response time

RTs were calculated for all correct responses for each set size and each cue type (see Figure 3.4b for group means) and submitted to a repeated measures ANOVA (Cue type X Set size). The effect of Cue type was significant, \( F(1, 23) = 352.93, MSE = 7280.57, p < 0.0001 \), with shorter RTs for all correct retro-cue responses relative to all correct simu-cue responses. The effect of Set size was also significant, \( F(5, 115) = 25.17, MSE = 4920.68, p < 0.0001 \), with RTs increasing as the number of items to be remembered increased, and the interaction between Cue type and Set size reached significance, \( F(5, 115) = 12.17, MSE = 5146.16, p < 0.0001 \), such that RTs increased less with increasing set size for retro-cue relative to simu-cue trials.

3.3.2 Exp 2b: Shapes

3.3.2.1 Capacity

As in Exp 2a, baseline K-estimates from simu-cue trials varied across individuals, in the case of shapes from 1.60 to 3.80 (2.70 ± 0.62), and all but three participants showed an average improvement in K following a retro-cue (group average improvement of 0.79 ± 0.60; see Figure 3.6a for group means). A repeated measures ANOVA (Cue type X Set size) showed greater capacity estimates for retro-cue relative to simu-cue trials, \( F(1, 23) = 65.17, MSE = 0.21, p < 0.0001 \), with estimates increasing overall as set size increased, \( F(5,115) = 70.70, MSE = 0.29, p < 0.0001 \), reaching a plateau at a set size of 3 items. Again, the interaction between Cue type and Set size was significant, \( F(5,115) = 9.87, MSE = 0.17, p < 0.0001 \), and was modulated when baseline capacity was entered as a covariate, \( F(5,110) = 1.59, MSE = 0.16, p > 0.15 \). That is, the retro-cue was most effective once baseline capacity for each individual had been exceeded.
Figure 3.6 Group mean (a) capacity and (b) response time across set sizes in Experiment 2b: Shapes, with error bars representing the standard error of each mean.
Estimates for *simu-cue* and *retro-cue* trials were correlated \((r = 0.57, p < 0.005)\), and in the case of shapes there was a moderate correlation between the size of the *retro-cue* benefit and baseline capacity \((r = -0.42, p < 0.05)\), such that individuals with the lowest baseline capacity showed the greatest improvements.

### 3.3.2.2 Response time

A repeated measures ANOVA (Cue type X Set size) for RTs for all correct responses showed significant effects of Cue type, \(F(1, 23) = 306.09, MSE = 6901.81, p < 0.0001\), with shorter RTs for all correct *retro-cue* responses relative to all correct *simu-cue* responses, and of Set size, \(F(5, 115) = 74.03, MSE = 5383.99, p < 0.0001\), with RTs increasing as the number of items to be remembered increased (see Figure 3.6b for group means). The Cue type X Set size interaction was also significant, \(F(5, 115) = 9.85, MSE = 3976.28, p < 0.0001\), with increasing set size resulting in greater increases to RT for *simu-cue* relative to *retro-cue* trials.

### 3.3.3 Exp 2c: Locations

#### 3.3.3.1 Capacity

Baseline \(K\)-estimates from *simu-cue* trials ranged across individuals from 1.44 to 6 (3.84 ± 1.14), with all but six participants showing capacity improvements following the *retro-cue* (group average improvement of 0.51 ± 0.69; see Figure 3.7a for group means). Note that some participants were already performing near or at ceiling, which partly accounts for the large variability in the *retro-cue* benefit. Nevertheless, a repeated measures ANOVA (Cue type X Set size) demonstrated a significant *retro-cue* benefit overall, \(F(1, 23) = 24.64, MSE = 0.24, p < 0.0001\), as well as a significant effect of Set size, \(F(5,115) = 110.74, MSE = 0.51, p < 0.0001\), with \(K\)-estimates reaching a plateau across participants and Cue type at 5 locations (i.e., pairwise contrasts between set sizes 2 and up all were significantly different, \(p < 0.05\), except for sets 5
Figure 3.7 Group mean (a) capacity and (b) response time across set sizes in Experiment 2c: Locations, with error bars representing the standard error of each mean.
and 6, \( p > 0.1 \)). There was a significant interaction between the retro-cue benefit and the number of locations to be remembered, \( F(5,115) = 4.57, \ MSE = 0.17, \ p < 0.001 \). In this case, the Cue type X Set size interaction was only slightly modulated when the baseline capacity was included as a covariate but was still significant, \( F(5,110) = 3.75, \ MSE = 0.16, \ p < 0.005 \). That is, baseline capacity still determined the point at which the retro-cue was effective in increasing \( K \)-estimates, but those participants whose baseline capacity was near or at ceiling had little to no room for improvement, thus in this sample individuals with lower capacity showed much greater retro-cue benefits. As with feature probes, when spatial locations were cued and the response to the probe ultimately rested on whether the feature at that location matched or not, capacity estimates for locations following simu-cues and retro-cues that selected for features were highly correlated (\( r = 0.80, \ p < 0.0001 \)), and the correlation between baseline capacity and the retro-cue benefit was also significant (\( r = -0.52, \ p < 0.01 \)).

3.3.3.2 Response time

A repeated measures ANOVA (Cue type X Set size) for RTs for all correct responses showed significant effects of Cue type, \( F(1, 23) = 527.72, \ MSE = 5384.96, \ p < 0.0001 \), with shorter RTs for all correct retro-cue responses relative to all correct simu-cue responses, and of Set size, \( F(5, 115) = 50.96, \ MSE = 2280.96, \ p < 0.0001 \), with RTs increasing as the number of items to be remembered increased (see Figure 3.7b for group means). The interaction between Cue type and Set size was significant, \( F(5, 115) = 6.34, \ MSE = 1635.48, \ p < 0.0001 \), with increasing set size resulting in greater increases to RT for simu-cue relative to retro-cue trials.
3.3.4 Comparing retro-cue effects across stimulus types: Colors vs. Shapes vs. Locations

While the independent analyses for each of Exps 2a-c indicate that retro-cues were effective in increasing $K$ once baseline capacity had been exceeded and in shortening RTs overall, a direct comparison between the three experiments is warranted to determine whether the magnitude of these effects were indeed comparable for various features probed by location and for locations probed by features. Thus data from all 72 participants from across the three experiments were combined.

3.3.4.1 Capacity

Figure 3.8a displays the baseline capacity estimates and average improvements across participants in each of the three experiments. A mixed design ANOVA, with Task type as a between-subjects factor and with Cue type and Set size as within-subjects repeated measures, verified that retro-cues conferred a comparable benefit on color, shape, and location probes. There was a main effect of Task type, $F(2,69) = 6.63, MSE = 2.62, p < 0.005$, with significant differences in overall capacity estimates for shapes relative to colors and locations ($p < 0.03$ and $0.003$, respectively) but no difference between colors and locations ($p > 0.99$). In other words, participants had comparable $K$s for colors and locations, with lower $K$s for shapes. The effect of Cue type was significant overall, $F(1,69) = 152.62, MSE = 0.20, p < 0.0001$, and did not interact with Task type, $F(1,69) = 2.32, p > 0.10$, indicating that the retro-cue benefits collapsed over all set sizes were comparable across the three tasks. The effect of Set size was also significant, $F(5,345) = 271.27, MSE = 0.42, p < 0.0001$, and it did interact significantly with Task type, $F(10,345) = 7.93, p < 0.0001$, confirming that $K$-estimates reached a plateau at lower set sizes for shapes relative to colors and locations. Finally, the Cue type X Set size interaction was significant, $F(5,345) = 22.77, MSE = 0.15, p < 0.0001$, which did not interact with Task type,
\( F(10,345) = 0.54, p > 0.80 \), verifying that the Task type had no effect on the observed retro-cue benefits. Across the three experiments, correlations followed the same pattern as observed separately for each experiment, with a negative correlation between baseline capacity and the retro-cue benefit \((r = -0.38, p < 0.001)\) and a stronger, positive correlation between \(K\)-estimates for retro-cue and simu-cue trials \((r = 0.79, p < 0.0001)\). That is, participants with lower baseline capacity had greater room for improvement, but all participants showed a relatively constant boost in \(K\) following a retro-cue once baseline capacity had been exceeded.

### 3.3.4.2 Response time

A mixed ANOVA (Task type X Cue type X Set size) for RTs from all correct trials showed no overall effect of Task type, \( F(2,69) = 2.26, MSE = 227903.94, p > 0.10 \), indicating that the overall response time for colors, shapes, and locations across the three experiments was comparable. The effect of Cue type was significant, \( F(1,69) = 1141.95, MSE = 6605.83, p < 0.0001 \), with no interaction with Task type, \( F(2,69) = 1.82, p > 0.15 \). Thus, RTs were shorter overall for retro-cue relative to simu-cue trials, irrespective of whether participants were responding to colors, shapes, or locations. The main effect of Set size, \( F(5,345) = 141.30, MSE = 4015.70, p < 0.0001 \), did interact with Task type, \( F(10,345) = 7.46, p < 0.0001 \), due to differential rates of increasing RT for increasing set sizes in the three experiments (see Figure 3.8b for slopes of RT functions in each experiment, calculated by taking the difference in RTs for set sizes 6 and 1 and dividing by 5). While each additional item to be remembered added roughly 25 ms to the RT at probe comparison for both colors and locations, each additional shape added roughly 50 ms to the comparison RT, with the range of RTs for shapes encompassing those for colors and locations. That is, participants responded comparably quickly when a single shape was remembered as when a single color was remembered (630 ms vs. 636 ms, respectively), but responding when six shapes were held in memory took nearly 100 ms
Figure 3.8 (a) Capacity estimates compared across the three tasks, with and the retro-cue benefit shown added to the baseline capacity estimate, and (b) RT slopes compared across the three tasks in Study 2.
longer than when six colors were held in memory (871 ms vs. 773 ms, respectively). This was further reflected in the Cue type X Set size effect, $F(5,345) = 28.39, \textit{MSE} = 3470.63, \ p < 0.0001$, which significantly interacted with Task type, $F(10,345) = 2.42, \ p < 0.01$, which was driven in part by greater slopes for shape RTs overall and in part by proportionately smaller slopes for \textit{simu-cue} trials for locations.

3.4 Discussion

These three experiments demonstrated the basic effects of \textit{retro-cues}, wherein providing attentional selection for one aspect of VSTM storage improved performance: biasing either a location or a particular feature during a VSTM delay led to both greater capacity estimates and shorter change-detection response times. Whereas previous studies have shown the effects of spatial selection on feature VSTM without making any distinction between two systems (e.g., Makovski, et al., 2008), these experiments demonstrate the reciprocity between location VSTM and feature VSTM such that advanced selection within either system facilitates performance. Comparable capacity estimates were observed for locations and colors, similar to the 1:1 correspondence noted in the existing literature (e.g., Awh et al, 2007), with lower estimates for more complex features such as the unfamiliar shapes. More importantly, the effect of \textit{retro-cues} on capacity estimates was virtually identical across the experiments, suggesting that representations within either system were equally susceptible to being biased by attentional cues and that selection within one system was sufficient to influence behavioural output that was also dependent on the other system. For instance, cueing only the location aspect of VSTM storage affected change-detection performance that was ultimately dependent on a feature comparison. Again, this provides the first demonstration using a single paradigm that location selection affects a feature VSTM task and, vice versa, that feature selection affects a location VSTM task.
All three experiments established robust effects of the retro-cue on response times (largely overlooked in previous studies, e.g., Makovski, et al, 2008), with overall shorter RTs even at the smallest set sizes. On simu-cue trials, when no cue was available during the delay period, the first step at the probe comparison would be to establish which of the locations (in Exps 2a and 2b) or which of the features (in 2c) stored in memory should be compared. Given the consistent effect of the retro-cue on RTs in all three experiments, a likely function of the retro-cue was to allow this initial step of the probe comparison to proceed even before the full probe was provided. When the number of items to be stored and potentially compared was within an individual’s limit, early facilitation of one aspect of probe comparison did not impact the final capacity estimate; once the set size exceeded baseline capacity, however, advanced selection of the location at which features should be compared (or vice versa) impacted storage estimates by reducing the comparison load. That is, the selection aspect of probe comparison proceeded in advance by shifting VSTM resources to one location (or feature) representation via a retro-cue, thus facilitating the full comparison that ensued once the probe appeared. It is important to note, however, that change-detection performance was not perfect following retro-cues, as might be predicted by a single-system, fixed-slots account of VSTM storage. If the retro-cue served to guarantee that the cued representation occupied a slot, then judgments following retro-cues should have been maximally accurate; estimates should not have reached any plateau but should have continued to increase linearly with increasing set sizes, because the item being compared to the probe would always be in a slot. On the contrary, the observed results demonstrate that, despite improved estimates on retro-cue trials, performance still reached a plateau, with a fairly constant benefit once baseline capacity had been exceeded, supporting the view that the retro-cue generated a reallocation of weights among VSTM representations within one system at a time.
To reiterate, *retro-cues* affected response times regardless of baseline capacity, which determined the point at which *retro-cues* conferred any benefit on estimates of capacity. The comparable findings across the experiments indicate reciprocity between separate VSTM systems for location and features, which had not been considered by previous studies using *retro-cues* that treated VSTM representations as unitary. In showing that feature selection and location selection can be dissociated within VSTM, the use of the *retro-cue* paradigm has effectively revealed the organization of VSTM into separate systems, corroborating the ERP results from the previous chapter. The next chapter better examines how weights within VSTM are shifted by *retro-cue* selection to support the notion that VSTM storage is flexible in nature, and this explanation of the *retro-cue* benefit is contrasted with an alternate interpretation in which participants could effectively remove all non-cued items from VSTM after the appearance of the *retro-cue*. 
4 Study 3: Manipulating validity to assess relative VSTM weighting

4.1 Introduction

Having established that a *retro-cue* led to shorter response times at probe comparison and increased estimates of capacity overall, an intuitive explanation is that attentional selection strengthened the cued representation within one VSTM system such that that aspect of the probe comparison could proceed in advance of the appearance of the full probe. In other words, spatial selection provided by a location *retro-cue* allowed the location-based component of probe comparison to be completed before the ultimate matching or non-matching judgment was made on the basis of the features presented in the probe. The work in the previous chapter, however, did not address fully whether the enhancement of the cued representation was achieved due to recruitment of additional resources from outside VSTM or via a redistribution of weights within VSTM. That is, if resources within a VSTM storage system were reallocated to favor one representation, then the benefit to the cued item should have been accompanied by some detriment to the other, non-cued items. In order to assess this, the effects of a *retro-cue* were examined with respect to the other representations which had not been cued, which was achieved by altering the predictive accuracy of the cue and thereby providing the opportunity to probe non-cued items on some trials. If the *retro-cue* is effective by eliminating all other representations from VSTM, then performance following *invalid retro-cues* should drop to chance, even when the set size is within baseline capacity. If, however, *retro-cue* selection reflects a reallocation of weights without complete elimination of non-cued items, then estimates on *invalid retro-cue* trials should be decreased relative to *simu-cue* trials but still follow a curvilinear profile.
4.2 Methods

4.2.1 Participants

Two separate samples of 24 participants each were tested for each of Exps 3a-b, recruited from the same pool as those in Exps 2a-c and tested under the same protocol. Demographic information is available in the Appendix.

4.2.2 Apparatus

The experimental setup was identical to that of Exps 2a-c.

4.2.3 Stimuli and procedure

The experiments in Study 3 were identical to Exp 2a: Colors from the previous chapter with two exceptions. First and critically, rather than having all cues predict with 100% accuracy the location of the impending probe whose color had to be compared to the item from that location in the memory array, the cues in Exp 3a: 80% validity accurately predicted the probe location on only 80% of all trials, leaving 20% of trials in which the cue indicated an item location that was present in the array but the probe ultimately tested a location other than that which had been cued (i.e., 480 valid and 120 invalid trials). Likewise, Exp3b: 60% validity had proportionately more cues that did not predict the probe location, with 40% of trials in which non-cued representations could be examined (i.e., 360 valid and 240 invalid trials). Second, to allow for invalid trials for all set sizes, the number of items in any given memory array ranged from two to six (for a total of five set sizes tested, rather than six). In both the 80% validity and the 60% validity experiments, probes were equally likely to be matching or non-matching regardless of cue validity and trials were equally likely to have a retro-cue or a simu-cue across set sizes from two to six items. In both experiments, participants were informed that the cue would indicate which
location would be tested on most but not all trials, and that they should pay attention to the cue when it appeared in advance.

4.3 Results

4.3.1 Exp 3a: 80% validity

4.3.1.1 Capacity

Baseline capacity was estimated as the best $K$ achieved in valid simu-cue trials (although capacities for valid and invalid simu-cue trials were statistically indistinguishable; see Figure 4.1a for group means). The average baseline capacity across individuals ranged from 1.56 to 4.50 (mean ± standard deviation = 3.28 ± 0.80), with all but three of the 24 participants showing an improvement in $K$ for valid retro-cue trials relative to valid simu-cue trials (group average improvement of 0.70 ± 0.82). When non-cued items were probed, not only did the retro-cue benefit disappear, but $K$-estimates were overall lower for 14 out of the 24 participants for invalid retro-cue relative to invalid simu-cue trials (group average decline of 0.35 ± 0.92), however, performance was still significantly different from chance in separate $t$-tests, (all $p$s < 0.0001). $K$-estimates were submitted to a repeated measures ANOVA with three factors, Cue type (retro-cue vs. simu-cue) X Validity (valid vs. invalid) X Set size. While both the effects of Validity and of Set size were significant, respectively $F(1,23) = 25.13$, $MSE = 0.71$, $p < 0.0001$ and $F(4,92) = 33.93$, $MSE = 0.74$, $p < 0.0001$, the main effect of Cue type did not achieve significance, $F(1,23) = 0.54$, $MSE = 0.63$, $p > 0.45$. That is, retro-cues did not result in overall greater estimates, but valid trials generated overall greater capacity than invalid trials and capacity overall increased with increasing memory load, up to a plateau of roughly three items presented in the memory array. The lack of a main effect of Cue type is explained by the interaction with Validity, $F(1,23) = 43.57$, $MSE = 0.44$, $p < 0.0001$, which confirms the apparent pattern in Figure 4.1a for valid retro-cues to improve capacity while invalid retro-cues diminish capacity, with simu-cues in
Figure 4.1 Group mean (a) capacity and (b) response time across set sizes in Experiment 3a: 80% validity, with error bars representing the standard error of each mean.
between regardless of validity. Although there was no significant interaction between Cue type and Set size, $F(4,92) = 0.71$, $MSE = 0.46$, $p > 0.55$, again because valid and invalid retro-cues generate diverging $K$-estimates, the interaction between Validity and Set size was significant, $F(4,92) = 4.26$, $MSE = 0.48$, $p < 0.005$, as was the interaction between all three factors, $F(4,92) = 5.90$, $MSE = 0.37$, $p < 0.0005$. That is, as the memory load increased, the envelope created by valid and invalid retro-cues widened around simu-cues. This final effect was modulated by baseline capacity entered as a covariate, $F(4,88) = 2.03$, $MSE = 0.34$, $p > 0.09$; although this may still trend toward significance, likely due to a slight imbalance between the benefit for valid retro-cues and the cost for invalid retro-cues, the modulation of the interaction term shows once again that the effects of the retro-cue on capacity were apparent once baseline capacity had been exceeded.

4.3.1.2 Response time

RTs from correct trials were also submitted to a Cue type X Validity X Set size ANOVA (see Figure 4.1b for group means). Cue type exhibited a significant overall effect on RT, $F(1,23) = 58.00$, $MSE = 13291.49$, $p < 0.0001$, as did Validity, $F(1,23) = 153.85$, $MSE = 24728.59$, $p < 0.0001$, and Set size, $F(4,92) = 24.31$, $MSE = 9461.86$, $p < 0.0001$. The interaction between Cue type and Validity was significant, $F(1,23) = 61.26$, $MSE = 13633.12$, $p < 0.0001$, indicating that while RTs for retro-cue trials were overall shorter than those for simu-cues, this was due to proportionately greater effects for valid retro-cues. The Cue type X Set size interaction was not significant, $F(4,92) = 0.84$, $MSE = 6783.89$, $p > 0.5$, showing that RTs following both retro-cues and simu-cues depended on the number of items held in memory. The interaction between Validity and Set size did achieve significance, $F(4,92) = 6.68$, $MSE = 5932.46$, $p < 0.0001$, and the three-way interaction effect was marginally significant, $F(4,92) = 2.41$, $MSE = 7060.04$, $p <$
0.06. This reflects the different pattern of RT for valid retro-cue trials relative to all others, whose rates of increase with increasing load are much larger than that for valid retro-cues.

4.3.2 Exp 3b: 60% validity

4.3.2.1 Capacity

Baseline capacity was estimated again using valid simu-cue trials (see Figure 4.2a for group means). The average baseline capacity ranged from 2.00 to 5.00 (mean ± standard deviation = 3.40 ± 0.84), with all but six of the 24 participants showing an improvement in K for valid retro-cue trials relative to valid simu-cue trials (group average improvement in K of 0.63 ± 0.69). K-estimates again decreased for 15 out of the 24 participants when non-cued items were probed (group average decline of 0.34 ± 0.80), although all invalid retro-cue estimates were still better than chance in separate t-tests (all ps < 0.0001). Capacity was analyzed with a Cue type X Validity X Set size ANOVA, which showed no main effect of Cue type, $F(1,23) = 0.43$, $MSE = 0.44$, $p > 0.50$. Both main effects of Validity and of Set size were significant, $F(1,23) = 25.80$, $MSE = 1.00$, $p < 0.0001$, and $F(4,92) = 30.92$, $MSE = 0.74$, $p < 0.0001$, respectively. The Cue type X Validity interaction was significant, $F(1,23) = 21.60$, $MSE = 0.59$, $p < 0.0002$, due to the divergence of K-estimates for retro-cues, where valid trials produced higher capacity than invalid trials, which was not the case for simu-cues. There was no interaction between Cue type and Set size, $F(4,92) = 0.96$, $MSE = 0.33$, $p > 0.40$, but the interaction between Validity and Set size was significant, $F(4,92) = 5.88$, $MSE = 0.30$, $p < 0.0003$, as was the full interaction between all three factors, $F(4,92) = 2.68$, $MSE = 0.38$, $p < 0.04$. This shows that retro-cue trials again formed an envelope around simu-cue, with the gap between valid and invalid retro-cues widening as the number of items to be remembered increased, with no differences in valid and invalid simu-cues regardless of memory load. This last interaction was modulated by baseline capacity, $F(4,88) = 1.30$, $MSE = 0.37$, $p > 0.25$, indicating that the apparently larger gap between
Figure 4.2 Group mean (a) capacity and (b) response time across set sizes in Experiment 3b: 60% validity, with error bars representing the standard error of each mean.
valid and invalid retro-cues at larger set sizes is due to individual differences, with baseline capacity determining the point beyond which both the valid retro-cue benefit and the invalid retro-cue detriment occur.

4.3.2.2 Response time

The Cue type X Validity X Set size for RTs from all correct trials (group means in Figure 4.2b) showed shorter RTs overall following retro-cues relative to simu-cues, $F(1,23) = 32.15$, $MSE = 41584.82$, $p < 0.0001$, shorter RTs for valid trials compared to invalid, $F(1,23) = 122.37$, $MSE = 22224.02$, $p < 0.0001$, and overall increases in RT with increasing load, $F(4,92) = 20.25$, $MSE = 4790.37$, $p < 0.0001$. Cue type interacted significantly with Validity, $F(1,23) = 59.01$, $MSE = 7803.63$, $p < 0.0001$, indicating that the difference between valid and invalid retro-cues was much greater than that between valid and invalid simu-cues. Neither Cue type nor Validity interacted with Set size, $F(4,92) = 0.43$, $MSE = 4289.62$, $p > 0.75$, and $F(4,92) = 1.42$, $MSE = 3829.60$, $p > 0.20$, respectively, nor was the Cue type X Validity X Set size interaction significant, $F(4,92) = 1.35$, $MSE = 3637.11$, $p > 0.25$, showing that increasing memory load had a constant effect on RTs across cueing conditions.

4.3.3 Comparing retro-cue effects across validity rates: 100% vs. 80% vs. 60%

The pattern of results for Exps 3a-b are virtually identical, showing that capacity was increased for retro-cues relative to simu-cues only when they were valid; when retro-cues were invalid, such that the probed location was different from what had been cued, the estimated capacity was diminished relative to simu-cue trials. Meanwhile, RTs were selectively shorter for valid retro-cue trials overall. The pattern of effects did not differ according to the validity rate: valid and invalid retro-cues appeared to elicit comparable benefits and deficits, respectively, regardless of the proportion of trials in which the probed location had been accurately cued. To better compare
the patterns for 80% and 60% validity tasks, the data from both experiments were submitted to common analyses, with Task as a between-groups factor. Furthermore, in order to assess how the magnitude of the retro-cue benefit might have been influenced by the presence of invalid trials, the data for valid trials from Exps 3a-b were compared to the data for set sizes two through six from Exp 2a.

4.3.3.1 Capacity

All K-estimates from the 80% and 60% validity tasks were submitted to a four-way ANOVA (Task X Cue type X Validity X Set size). The between-groups measure of Task did not have an overall effect on capacity, $F(1,46) = 0.05, MSE = 5.93, p > 0.80$. Nor did Cue type, $F(1,46) = 0.97, MSE = 0.54, p > 0.30$, nor the interaction between Task and Cue type, $F(1,46) = 0.01, p > 0.85$, confirming that the lack of an overall effect of Cue type was comparable for the separate experiments. The effect of Validity was significant across the two experiments, $F(1,46) = 50.62, MSE = 0.86, p < 0.0001$, with no interaction with Task, $F(1,46) = 0.42, p > 0.50$. Similarly, Set size was overall significant, $F(4,184) = 64.21, MSE = 0.74, p < 0.0001$, with no interaction with Task, $F(4,184) = 0.61, p > 0.65$. As in the separate analyses for 80% and 60% validity, the Cue type X Validity interaction was significant, $F(1,46) = 61.27, MSE = 0.52, p < 0.0001$, and this did not differ by Task, $F(1,46) = 0.61, p > 0.40$. The envelope that valid and invalid retro-cues formed around simu-cues was similar irrespective of the predictive accuracy of the cues. The interaction between Cue type and Set size again was not significant, $F(4,184) = 0.71, MSE = 0.40, p > 0.55$, and did not interact with Task, $F(4,184) = 0.92, p > 0.45$. The interaction between Validity and Set size, which was significant in both experiments, was also significant across the two, $F(4,184) = 9.30, MSE = 0.39, p < 0.0001$, and again there was no interaction with Task, $F(4,184) = 0.45, p > 0.75$. Finally, the Cue type X Validity X Set size interaction that characterized the divergence of valid and invalid retro-cues around simu-cues maintained its
significance across both experiments, $F(4,184) = 7.44$, $MSE = 0.37$, $p < 0.0001$, and continued to be modulated by baseline capacity, $F(4,180) = 3.27$, $MSE = 0.35$, $p < 0.05$, with no interaction with Task, $F(4,184) = 1.10$, $p > 0.35$. In other words, not only were the 80% and 60% validity tasks virtually identical, they were also statistically identical with respect to capacity estimates. To reiterate, while Validity had no impact on $K$-estimates for simu-cue trials, it did affect capacity for retro-cue trials once baseline capacity had been exceeded, such that valid retro-cues resulted in greater capacity estimates and invalid retro-cues led to diminished estimates.

While estimates for the 80% and 60% validity tasks were indistinguishable, the magnitude of the benefit conferred by the retro-cues may have been influenced by the presence of invalid trials. To test this, the data from these two experiments were compared to those from Exp 2a, which had 100% valid cues. The $K$-estimates for valid trials only were submitted to a Task X Cue type X Set size ANOVA (see Figure 4.3a for group means across tasks), which again verified that the between-groups factor of Task had no overall effect on capacity, $F(2,69) = 0.02$, $MSE = 3.00$, $p > 0.95$. While both Cue type and Set size had significant effects on capacity, $F(1,69) = 109.70$, $MSE = 0.31$, $p < 0.0001$ and $F(1,69) = 145.09$, $MSE = 0.41$, $p < 0.0001$, neither interacted with Task, $F(1,69) = 0.79$, $p > 0.45$ and $F(1,69) = 0.27$, $p > 0.95$, respectively. Likewise, the Cue type X Set size interaction remained significant across all three experiments, $F(4,276) = 14.57$, $MSE = 0.19$, $p < 0.0001$, with no interaction with Task, $F(8,276) = 0.85$, $p > 0.55$, and was again modulated by baseline capacity, $F(4,272) = 1.44$, $MSE = 0.19$, $p > 0.20$. That is, regardless of the proportion of trials that violated the expectation that the cued location would be probed, the boost in estimated capacity when the cue did predict the probed location was comparable.
Figure 4.3 (a) Capacity estimates compared across Experiments 2a, 3a, and 3b to examine effects of validity rates, with and the retro-cue benefit shown added to the baseline capacity estimate, and (b) RT slopes compared across Experiments 2a, 3a, and 3b.
4.3.3.2 Response time

RTs from all correct trials for both Exps 3a-b were submitted to a Task X Cue type X Validity X Set size ANOVA, which indicated that the 80% and 60% validity tasks were overall comparable with respect to RT, $F(1,46) = 0.003$, $MSE = 576127.99$, $p > 0.95$. As in the separate experiments, Cue type had a significant effect overall, $F(1,46) = 75.41$, $MSE = 27438.15$, $p < 0.0001$, which did not interact with Task, $F(1,46) = 1.41$, $p > 0.20$. The main effect of Validity was also significant, $F(1,46) = 275.96$, $MSE = 23476.31$, $p < 0.0001$, and did not interact with Task, $F(1,46) = 1.93$, $p > 0.15$. The main effect of Set size, $F(4,184) = 43.64$, $MSE = 7126.11$, $p < 0.0001$, did have a marginal interaction with Task, $F(4,184) = 2.24$, $p < 0.07$, such that the 80% validity task had a larger range of RTs from the lowest set size to the largest relative to 60% validity (i.e., the slope for increasing RT with increasing load was larger overall in the former). The Cue type X Validity interaction was again significant, $F(1,46) = 118.31$, $MSE = 10718.38$, $p < 0.0001$, with proportionately greater boost in RT for valid relative to invalid trials for retro-cues compared to simu-cues, and this did not interact significantly with Task, $F(1,46) = 2.58$, $p > 0.10$. Cue type and Set size did not interact with one another, $F(4,184) = 0.48$, $MSE = 5536.75$, $p > 0.75$, or with Task, $F(4,184) = 0.89$, $p > 0.45$. Validity and Set size did again interact significantly with one another, $F(4,184) = 5.12$, $MSE = 4881.03$, $p < 0.001$, and with Task, $F(4,184) = 4.12$, $p < 0.005$, with a greater range in RTs specifically following invalid cues in the 80% validity task. Finally, while the Cue type X Validity X Set size interaction was significant across the two experiments, $F(4,184) = 2.63$, $MSE = 5348.58$, $p < 0.05$, it did not interact with Task, $F(4,184) = 1.46$, $p > 0.20$.

The RTs across predictive accuracies were compared using only valid trials and the data from Exp 2a in a Task X Cue type X Set size ANOVA, which again confirmed that Task did not exhibit an overall effect on RTs, $F(2,69) = 0.92$, $MSE = 206741.34$, $p > 0.40$. However, the
significant effect of Cue type across the three experiments, $F(1,69) = 508.04$, $MSE = 11582.05$, $p < 0.0001$, did interact with Task, $F(2,69) = 3.64$, $p < 0.05$, such that retro-cue trials had shortest RTs when the cues had 100% predictive accuracy, and slightly longer RTs for 80% relative to 60% validity, which were still shorter than RTs for simu-cues. The significant main effect of Set size, $F(4,276) = 41.97$, $MSE = 3699.52$, $p < 0.0001$, did not interact with Task, $F(8,276) = 0.65$, $p > 0.70$. The significant Cue type X Set size interaction, $F(4,276) = 7.38$, $MSE = 3457.86$, $p < 0.0001$, however, did interact with Task, $F(8,276) = 2.40$, $p < 0.05$, indicating that the Task X Cue type interaction was driven by a proportionately smaller slope of increasing RTs with increasing load for retro-cue relative to simu-cue trials specifically in the 100% validity task (see Figure 4.3b for slopes calculated by taking the difference in RT between set sizes 6 and 2 and dividing by 4). That is, when the cue indicated the location of the impending probe with 100% probability, retro-cues showed the shortest RTs relative to simu-cues across set sizes, while disrupting the predictive accuracy of the cues decreased the difference in response slopes following retro-cues and simu-cues.

4.4 Discussion
Experiments 3a-b showed that the retro-cue benefit was specific to the cued representation in VSTM and that this performance advantage came at the expense of the remaining items being stored, supporting the view that VSTM resources were flexibly reallocated via retrospective selection. When the retro-cue location accurately predicted the probe location, increased capacity estimates and shortened RTs were observed relative to simu-cue trials, as before. When the probe ultimately appeared at a location other than that indicated by the retro-cue, however, the estimated capacity decreased relative to simu-cue trials without dropping to chance levels, indicating that non-cued representations had lower representational strength (i.e., reduced VSTM weights) without being completely eliminated from memory. As in Experiments 2a-c, the retro-
cue benefit was influenced by baseline capacity such that retro-cues were only effective once the set size exceeded an individual’s baseline capacity. Again, the timing of the change-detection decision was influenced by having a retro-cue, but only on valid trials. When the probe followed an invalid retro-cue, RTs were no longer facilitated, as the non-cued location still had to be selected before the full probe comparison could be accomplished. Because the relative weightings on non-cued locations were lowered to dedicate more weight on the cued location, the chances of selecting the correct location for comparison were diminished and thus the capacity estimates on invalid retro-cue trials were lower.

While these experiments show that the non-cued representations on retro-cue trials were weaker relative to cued representations, they do not provide conclusive evidence that only one location can be selected at a time or that weights on representations remain in a malleable state until the probe comparison is complete. While it is unlikely that the retro-cue led to complete and irreversible selection of one location given that non-cued representations were still available albeit weaker, the next chapter provides stronger evidence that the reallocation of weights due to retro-cues is flexible until the full probe appears by presenting multiple retro-cues presented either simultaneously or sequentially.
5 Study 4: Using multiple location cues to examine cue specificity and reallocation of resources

5.1 Introduction

The previous chapters showed that retrospectively cueing a location during a VSTM delay biased the representation for that location such that the subsequent probe comparison time was shorter and performance more accurate. This benefit for a cued representation was accompanied by a detriment to the remaining, non-cued representations, indicating that relative representational strength was altered due to *retro-cue* selection such that weights were reallocated from all non-cued locations to strengthen the cued location. This is comparable to the visual selection that occurs when multiple perceptual inputs compete for visual processing resources, in which case select inputs may be biased for further processing at the expense of other inputs (Desimone & Duncan, 1995). A variety of studies have shown that such selection for visual input can occur at more than one location at a time (e.g., Awh & Pashler, 2000; Kramer & Hahn, 1995; Müller, Malinowski, Gruber, & Hillyard, 2003; McMains & Somers, 2004). While there is agreement that attention can be split for perceptual information, it is unclear whether attentional resources can be split between two separate representations in VSTM. The first experiment in this chapter (*Exp 4a: Simultaneous*) biased VSTM representations of two locations at a time, with RT and accuracy benefits for either of the cued locations. A corollary question is whether VSTM resources remained in a malleable state even after *retro-cue* selection, such that weights could be shifted again to another location representation. While other authors have suggested that a *retro-cue* may protect the cued item from degradation (Matsukura, Luck, & Vecera, 2007), capacity estimates following *invalid retro-cues* (*Study 2a-b*) indicate that non-cued representations are still maintained to some degree. Thus the second experiment in this chapter (*Exp 4b: Sequential*)
presented two retro-cues in sequence to determine whether VSTM resources could shift from one cued location to another, which would demonstrate that VSTM representations remain in a malleable state throughout the delay until probe comparison is complete.

5.2 Methods

5.2.1 Participants

Independent samples of 24 and 30 participants were tested for Exp 4a-b, respectively. As with all tasks, they were recruited from the University of Toronto community and tested under procedures approved by the Ethics Review Board, had normal or corrected-to-normal vision and were not color-blind. Basic demographic information, including age, sex, and handedness, is reported for all experiments in the Appendix.

5.2.2 Apparatus

The experimental setup was identical to that of Exp 2a-c and 3a-b.

5.2.3 Stimuli and procedure

To discourage participants from making eye movements, trials in Study 4 included a black fixation cross subtending 2° presented at the center of the display for the duration of each trial; thus the background color was medium grey rather than black. Memory and probe stimuli were the colored circles from Exp 2a, and the cue was composed of four small white triangles positioned at the corners of an imaginary square 1° in diameter, which was presented at the location of the probe item (see Figure 5.1). In Exp 4a: Simultaneous, stimuli were presented in any of 24 locations spaced along a 4x6 grid, with rows 4.5° or 1.5° above or below the horizontal meridian and columns 2°, 4.5° or 7° to the left or the right of the vertical meridian. In Exp 4b: Sequential, the display was simplified with stimuli presented around the fixation cross in any of eight potential locations spaced along an imaginary square with 3.5° diameter.
Figure 5.1 Illustration of event sequences for (a) one-cue and (b) two-cue trials in Experiment 4a: Simultaneous.
In Exp 4a: Simultaneous, the fixation cross was presented alone for 500 ms before a memory array with two, four, or six different colored circles appeared for 1000 ms, followed by a delay of 1000 ms. After this initial delay, participants saw either no cue but a continued display with fixation only (no-cue trials), a single cue at the location of one of the items from the memory array (one-cue trials), or cues at the locations of two items from the memory array (two-cue trials), in all cases lasting for 100 ms. Following another delay lasting 400 ms, a single probe appeared until a response was made, which terminated the trial and initiated an ITI of 500 ms (see Figure 5.1). As in the previous experiments, the probe item was equally likely to be a match or a non-match to the item that had appeared previously at the probe location, with non-matching items randomly selected from the remaining items that were not present in the memory array. Participants were instructed to respond as accurately as possible and were told that any cue(s) would indicate the location that was about to be tested. Although two-cue trials had lower certainty than one-cue trials since either of the two cued locations could be probed with equal probability, participants were advised to use the cues to help focus their memory.

In Exp 4b: Sequential, the fixation cross again was presented alone for 500 ms before a memory array with anywhere from two to six different colored circles appeared for 1000 ms, followed by a delay of at least 1000 ms. At this point, one of three different trial sequences ensued, which were randomly intermixed and controlled for various timing intervals between the presentation of the memory array and probe. On no-cue trials, a single probe was presented after a delay interval of 1000 to 2000 ms. On one-cue trials, a cue appeared for 100 ms during a delay interval of 1000 to 2000 ms before the probe appeared. On two-cue trials, the first cue appeared for 100 ms immediately after the initial 1000 ms delay, followed by a second delay of 400 ms before the second cue appeared for 100 ms, and finally followed by a third and final delay of 400 ms before
the probe appeared. Some two-cue trials showed a cue in the same location twice (same two-cue trials); some showed the two cues in two different locations (different two-cue trials), in which case either the first or the second could be probed (first two-cue and second two-cue trials, respectively). In all cases, the probe remained on the screen until a response was made, which initiated the 500 ms ITI, and the color of the probe was equally likely to be a match or non-match to the color that appeared in that location in the memory array (see Figure 5.2). Participants were told that trials would have different numbers of cues and that the cues, when they appeared, would always predict where the probe would appear. As in Exp 4a, different two-cue trials had lower certainty than one-cue and same two-cue trials, and participants were warned that the probe could appear at the location of either the first or the second cue.

5.3 Results

5.3.1 Exp 4a: Simultaneous

5.3.1.1 Capacity

$K$-estimates were generated for each condition (see Figure 5.3a for group means), and baseline capacity for each individual was estimated as the best $K$ achieved for any set size in the no-cue condition. The average baseline capacity across individuals ranged from 1.80 to 5.07 (mean ± standard deviation $= 3.24 ± 0.16$). All but seven participants showed an overall greater capacity for one-cue trials relative to no-cue trials (average improvement in $K$ of $0.29 ± 0.13$), but fewer than half showed any improvement on two-cue trials (average difference in $K$ of $-0.18 ± 0.11$). $K$-estimates were submitted to a repeated measures ANOVA with two factors, Cue type (no-cue vs. one-cue vs. two-cue) X Set size (2 vs. 4 vs. 6). The effect of Cue type was significant, $F(2,46) = 14.25$, $MSE = 0.19$, $p < 0.0001$, with greater capacity estimates on one-cue trials relative to both no-cue and two-cue trials, between which there was no difference. The effect of Set size was also significant, $F(2,46) = 44.46$, $MSE = 0.68$, $p < 0.0001$, such that storage estimates
Figure 5.2 Illustration of event sequence for trials with two retro-cues, in this case a first two-cue trial, in Experiment 4b: Sequential.
Figure 5.3 Group mean (a) capacity and (b) response time across set sizes in Experiment 4a: Simultaneous, with error bars representing the standard error of each mean.
increased from 2 to 4 items in the memory array with no difference in estimates for 4 and 6 items. There was a significant interaction between the Cue type and Set size, \( F(4,92) = 2.85, MSE = 0.15, p < 0.05 \), that was modulated when baseline capacity was entered as a covariate, \( F(4,88) = 1.82, p > 0.1 \), showing once again that the point at which a retro-cue elicited any advantage was determined by baseline capacity.

5.3.1.2 Response time
RTs from correct trials were also submitted to a Cue type X Set size ANOVA (see Figure 5.3b for group means). Cue type exhibited a significant overall effect on RT, \( F(2,46) = 308.19, MSE = 1548.71, p < 0.0001 \), with significantly shorter RTs on one-cue trials. Interestingly, RTs on two-cue trials, while not as short as on one-cue trials, were still significantly shorter than those on no-cue trials. Set size also exhibited a significant effect on RT, \( F(2,46) = 72.00, MSE = 1171.02, p < 0.0001 \), with significantly longer RTs for every increase in Set size. Furthermore, the interaction between Cue type and Set size was significant, \( F(4,92) = 8.58, MSE = 749.59, p < 0.0001 \), with a greater slope of increasing RT with increasing Set size for no-cue trials relative to both one-cue and two-cue trials, whose RT slopes were similar.

5.3.2 Exp 4b: Sequential
5.3.2.1 Capacity
Baseline capacity was estimated again as the best \( K \)-estimate on no-cue trials (see Figure 5.4a for group means). The average baseline capacity ranged from 1.75 to 6.00 (mean ± standard deviation = 3.76 ± 0.18), with all but six of the 30 participants showing an improvement in \( K \) for one-cue trials relative to no-cue trials (group average improvement of 0.34 ± 0.09). Capacity was analyzed with a Cue type (no-cue, one-cue, same two-cue, first two-cue, second two-cue) X Set size ANOVA, which showed a main effect of Cue type, \( F(4,116) = 14.92, MSE = 0.30, p < \)
Figure 5.4 Group mean (a) capacity and (b) response time across set sizes in Experiment 4b: Sequential, with error bars representing the standard error of each mean.
0.0001, such that no-cue and first two-cue trials had overall lower estimates relative to one-cue, same two-cue, and second two-cue trials, which did not differ from one another. The effect of Set size was significant, $F(4,116) = 75.06, MSE = 1.06, p < 0.0001$, with estimates on average across conditions reaching a plateau by five items. The Cue type X Set size interaction did not reach significance, $F(16,464) = 1.27, MSE = 0.30, p > 0.2$, and entering baseline capacity as a covariate had no effect on the interaction term, $F(16,448) = 0.88, p > 0.5$. Overall, these results indicate that VSTM weights were reallocated a second time to bias the representation of the second two-cue location. Post-hoc $t$-tests between no-cue and first two-cue trials at each Set size confirm that capacity estimates were no different when there was no cue relative to when the probed representation had been cued initially but then weights were shifted away by a subsequent cue, all $ps > 0.1$. Likewise $t$-tests between same and second two-cue trials show comparable capacity estimates for the most recently cued item, even when weights had been initially allocated toward a different representation, all $ps > 0.1$.

5.3.2.2 Response time

The Cue type X Set size ANOVA for RTs from correct trials (group means in Figure 5.4b) showed a significant effect of Cue type, $F(4,116) = 173.02, MSE = 5005.73, p < 0.0001$. RTs to same two-cue trials were overall shortest, followed by one-cue trials, then by both first and second two-cue trials, which did not differ from one another, and finally RTs to no-cue trials were longest. The significant effect of Set size, $F(4,116) = 14.74, MSE = 2907.90, p < 0.0001$, showed that RT increased with increasing load. Cue type interacted significantly with Set size, $F(16,464) = 3.44, MSE = 1104.14, p < 0.0001$, such that RT slopes with increasing load were larger for no-cue and first two-cue trials relative to all others.
5.4 Discussion

Together, Exp 4a-b demonstrated that retro-cue selection was most effective when one location was biased within VSTM at a time but that this reallocation of resources remained flexible, such that weights could be shifted to one representation by an initial retro-cue and then shifted again to a different representation by a subsequent retro-cue. In Exp 4a, there was no benefit to capacity estimates when the two retro-cues at two different locations were presented simultaneously. Nonetheless, RTs following two simultaneous retro-cues were shortened to an intermediate level between no-cue and two-cue RTs, suggesting that the two retro-cues promoted some degree of resource reallocation that was insufficient to provide the level of selection necessary to improve capacity. That is, more than one location representation could be biased in VSTM, allowing selection of either location at probe comparison to proceed more efficiently, but the lack of spatial specificity, or the lack of certainty concerning which of the strengthened representations ultimately would be selected, prevented the location aspect of the comparison to proceed in advance. Thus the location comparison was still necessary at the time that the full probe appeared, and the estimated storage capacity looked no different than when no advanced cue information was available, despite differences in RT.

This interpretation was further supported by the results of Exp 4b, in which VSTM resources were shifted multiple times. In that case, when two locations were retro-cued sequentially and either could be probed, ostensibly it would have been beneficial to keep both representations active. The RT results suggest that this was the case, as participants showed intermediate RTs when uncertain which of two locations would be probed subsequently, as was observed for Exp 4a. However, in Exp 4b, estimates were improved for the second of the two cued locations, while capacity for the first of the cued locations was no better but also no worse than when there were no cues at all. A single retro-cue was beneficial to probe comparison RT and accuracy when no
further retro-cues were presented, indicating that participants were attending to the first of the retro-cues. Likewise, the first two-cue trials did not show diminished capacity estimates relative to no-cue trials, as observed following invalid retro-cues in Exps 3a-b, further showing that these representations did not have impoverished weights. But the relative increase in estimated capacity for second two-cue trials showed that weighting in VSTM remained malleable even after an initial cue, such that a subsequent cue could promote selection of a different location and thus generate a redistribution of weights to that location. Thus, VSTM resources remained flexible such that multiple locations had strengthened representations, but selection could only be completed under conditions of spatial specificity when the probe location was known with certainty, and in the case of sequential cues the most recently cued representation was most readily available for comparison to the probe. The question of spatial specificity was further addressed by the experiment in the next chapter, in which spatial retro-cues were compared against spatially non-informative cues to examine the potential that retro-cues function as a general alerting mechanism. Finally, the preceding studies have shown that features and locations are represented separately in VSTM and that resources within one system can be flexibly allocated, with Exp 4b demonstrating that these representations are malleable until compared to the probe; the final study tests whether these separate systems interact online or are integrated only at probe comparison.
6 Study 5: Verifying the specificity of retro-cues

6.1 Introduction

The foregoing chapters presented the argument that retro-cues operated by promoting selection in one VSTM system, via a flexible reallocation of resources, such that probe comparison RT was shortened and capacity estimates increased when retrospective cues were spatially specific. A potential alternative interpretation for the shorter RTs following location retro-cues, however, is that of a general alerting mechanism whereby the advanced appearance of the retro-cue served as a readiness signal that warned participants that the probe was about to appear and prepared them to make a response, independent of orienting to any specific location (Fernandez-Duque & Posner, 1997). If this were the case, then the retro-cue benefit should not have been driven by spatial specificity: Any advanced cue would serve as a readiness signal and generate shortened RTs and improved estimates. To test this specifically, Experiment 5 compared location retro-cues to non-spatial ready-cues that appeared with the same timing as retro-cues but did not predict the location of the upcoming probe.

6.2 Methods

6.2.1 Participants

Twenty-one participants were tested for Exp 5, all recruited from the University of Toronto community and tested under procedures approved by the Ethics Review Board. As with all the studies, they had normal or corrected-to-normal vision and were not color-blind. Basic demographic information is reported for all experiments in the Appendix.

6.2.2 Apparatus

The experimental setup was identical to that of Exps 2a-c, 3a-b, and 4a-b.
6.2.3 Stimuli and procedure

The stimuli and display were identical to those of Exp 4b, and the trial timing was similar to that of Exp 4a. The key exception was that trials in Exp 5 could have no-cue, a retro-cue presented at the location of the impending probe, or a ready-cue presented around the fixation cross at the center of the display (see Figure 6.1). All three cue types were equally likely for each set size of one to six items, and the probe color was equally likely to be a match or a non-match to the item presented at the probed location in the memory array.

6.3 Results

6.3.1 Capacity

$K$-estimates were generated for each condition in each set size (see Figure 6.2a for group means), and baseline capacity for each individual was estimated as the best $K$ achieved for any set size in the no-cue condition. The average baseline capacity across individuals ranged from 2.14 to 5.60 (mean ± standard deviation = 3.29 ± 0.21). All but four participants showed an overall greater capacity for retro-cue trials relative to no-cue trials (average improvement in $K$ of 0.55 ± 0.15), and only seven showed any improvement on ready-cue trials (average difference in $K$ of 0.10 ± 0.13). $K$-estimates were submitted to a repeated measures ANOVA with two factors, Cue type (no-cue vs. retro-cue vs. ready-cue) X Set size (one through six). The effect of Cue type was significant, $F(2,40) = 13.05, MSE = 0.30, p < 0.0001$, with greater capacity estimates on retro-cue trials relative to both no-cue and ready-cue trials, between which there was no difference. The effect of Set size was significant, $F(5,100) = 44.89, MSE = 0.89, p < 0.0001$, with estimates reaching a plateau by four items. The interaction between the Cue type and Set
Figure 6.1 Illustration of event sequences for (a) spatially informative retro-cue and (b) spatially non-informative ready-cue trials in Experiment 5.
Figure 6.2 Group mean (a) capacity and (b) response time across set sizes in Experiment 5, with error bars representing the standard error of each mean.
size was also significant, $F(10,200) = 2.63$, $MSE = 0.21$, $p < 0.005$, which was modulated when baseline capacity was entered as a covariate, $F(10,190) = 1.01$, $p > 0.4$, again demonstrating that the *retro-cue* confers a constant advantage on $K$-estimates once baseline capacity has been exceeded.

### 6.3.2 Response time

RTs from correct trials were also submitted to a Cue type X Set size ANOVA (see Figure 6.2b for group means). Cue type has a significant effect on RT overall, $F(2,40) = 148.50$, $MSE = 4615.16$, $p < 0.0001$, with significantly shorter RTs on *retro-cue* trials. While RTs on *ready-cue* trials were shorter than those for *no-cue* trials, they were significantly longer than those for *retro-cue* trials. Reaction times were shorter for *ready-cue* trials relative to *no-cue* trials by an average of 35.53 ms ($p < 0.005$), but *retro-cue* trials were markedly shorter than both *no-cue* and *ready-cue* trials by roughly 141.75 and 106.21 ms, respectively (both $ps < 0.0001$). Set size also had a significant effect on RT, $F(5,100) = 25.12$, $MSE = 8189.45$, $p < 0.0001$, with positive RT slopes as load increased, and the interaction between Cue type and Set size was significant, $F(10,200) = 5.07$, $MSE = 2316.30$, $p < 0.0001$, with a flatter slope of increasing RT with increasing Set size for *retro-cue* trials relative to both *no-cue* and *ready-cue* trials. Post-hoc $t$-tests verify that capacity estimates for *ready-cue* trials are significantly different from those for *retro-cue* trials for Set size of three and above, all $ps < 0.05$, and no different from *no-cue* trials at all Set sizes, all $ps > 0.2$.

### 6.4 Discussion

The four-fold increase in the RT benefit for *retro-cue* relative to *ready-cue* trials, combined with the increased capacity only for *retro-cues* trials, indicated that performance benefits following *retro-cues* were due primarily to the spatial specificity of the location cue. Fernandez-Duque and
Posner (1997) demonstrated a spatially broad alerting mechanism that is independent of orienting to a specific location during visual attention tasks, with distinct effects of spatial cues that surpassed the effects of non-spatial warning signals. A similar dissociation may be made for the present change-detection tasks. While any advanced cue may have some general alerting function, seen by slightly shorter RTs on ready-cue trials relative to no-cue trials, the results observed in all these behavioural studies cannot be explained by a non-spatial readiness signal. Directing attention to a particular location in VSTM reallocated resources to that specific location representation, and this advanced selection facilitated probe comparison, resulting in shorter RTs and improved capacity estimates.

Earlier in Chapter 3, judging whether a probe matched or not was characterized as a two-step process, wherein first the location of the item to be compared was selected to then allow comparison of the features. The retro-cue was effective in that it facilitated early selection of the probe location, thereby shortening the time to match the feature location and, critically for set sizes that exceeded baseline capacity, reducing the comparison load to boost capacity estimates. In Experiment 5, a non-spatial ready-cue did shorten RTs somewhat, but not to the degree of a spatially specific retro-cue and without the boost in estimated capacity that accompanied the retro-cue. Thus the general alerting function of a cue during the delay period may have served to enhance general response preparation, but this ultimately did not impact the specific response criteria involved in probe comparison. Taken together with the preceding chapters, the data best support a two-system model of VTSM storage in which resources can be biased toward certain representations to augment capacity estimates when those representations are specifically selected. The final study in the next chapter tests whether such selection within one system has an online influence on storage within the other system or whether location VSTM and feature VSTM remain independent until compared to the probe.
7 Study 6: Testing whether location retro-cues affect feature resolution

7.1 Introduction

The studies so far have demonstrated that location *retro-cues* promoted advanced selection of a given location by redistributing VSTM weights toward that representation at the expense of all other location representations, that this early facilitation was effective only when the *retro-cue* was spatially specific and indicated the probe location with certainty, and that VSTM representations remained malleable (i.e., weights could continue to be reallocated) until probe comparison. Given the task demands that a location be selected prior to completing a feature comparison, one intuitive explanation is that location selection and feature comparison were carried out consecutively and independently. An alternative account that has not yet been ruled out, however, is that VSTM representations for locations and for features interacted online, such that biasing VSTM weights to a particular location representation also generated a shift in weights for feature representations. The current study turns to this remaining question concerning whether the two systems interacted online. The first chapter introduced the idea of a fixed slots model with variable resolution that depended on the complexity of memory items and hence the similarity of probes to items in memory (Awh et al., 2007). If a feature-based VSTM system were integrated online with a location-based VSTM system, such that resolution were intrinsically associated to slots, then such a theory would predict that selecting a slot via a location *retro-cue* would enhance the feature resolution at that slot, similar to the effects of spatial attention on feature resolution in perceptual tasks (Yeshurun & Carrasco, 1998). If, on the other hand, locations and features were stored independently, such that changes to one system
did not influence storage within the other system, then a location retro-cue should not improve feature resolution.

To assess the issue of enhancing resolution by selecting a location via a retro-cue, the paradigm introduced by Awh and colleagues (2007; Barton et al., 2009) was modified to include retro-cues and simu-cues. Items in the memory array were selected from different categories of complex stimuli. When a probe was a non-match, it could be selected from within the same category (same-category change), in which case there was a high probability of comparison error due to similarity between the feature held in memory and the probe feature. Alternatively, a non-matching probe could be selected from a different category (cross-category change), thus reducing the likelihood of comparison errors because of greater feature discriminability between the memory and probe items. Awh and colleagues used two categories with different levels of complexity, Chinese characters and shaded cubes. The key question for the experiment in this chapter was whether or not the two separate VSTM systems interact online. Thus the main variable of interest was not how capacity estimates differ for varying levels of complexity but rather whether location retro-cues altered feature resolution, as indexed by differing estimates of capacity for same-category versus cross-category changes. For this purpose, the complex stimuli were of the same degree of feature complexity (both drawn from the unfamiliar shapes used in Exp 2b), with the two categories differentiated by a highly discriminable feature (i.e., color). To be more specific, a model in which location and feature VSTM representations interact online would predict that spatial selection increases feature resolution, resulting in the same capacity boost either for retro-cues or for cross-category changes. Thus an interaction between the two would be observed, with retro-cues having an effect for same-category but not cross-category changes (or that cross-category changes have an effect for simu-cue but not retro-cue trials). Alternatively, a model in which representations are maintained independently until probe
comparison would predict that spatial selection during maintenance has no influence on feature resolution, resulting in additive effects of \textit{retro-cues} and \textit{cross-category} changes.

7.2 \hspace{1em} \textbf{Methods}

7.2.1 \hspace{1em} \textbf{Participants}

Thirty participants were tested for \textit{Exp 6}, all recruited from the University of Toronto community and tested under procedures approved by the Ethics Review Board. As with all the studies, they had normal or corrected-to-normal vision and were not color-blind. Basic demographic information is reported for all experiments in the Appendix.

7.2.2 \hspace{1em} \textbf{Apparatus}

The experimental setup was identical to that of the other behavioural studies.

7.2.3 \hspace{1em} \textbf{Stimuli and procedure}

Stimuli were presented against a medium grey background, with a black, central fixation cross (subtending 1") presented centrally throughout each trial. Stimuli were seven different colored circles (1.31" diameter, filled with red, yellow, green, light blue, dark blue, and purple) and twelve unfamiliar shapes (1.4" wide, drawn from the same pool of shapes as those in \textit{Exp2b}), six of which were colored bright green and the other six bright blue (the specific stimuli assigned to either color were counterbalanced across participants). Cues were similar to those in \textit{Exps 4a-b} and \textit{Exp 5}, composed of four triangles at the corners of an imaginary square subtending 1.5, and were presented at the probe location (see Appendix 2 for examples of all stimuli). Stimuli could be presented in any of eight locations evenly spaced along an imaginary circle centered on the display with a diameter of 5°.
Trials with colored circles were all presented first to establish capacity baselines using stimuli in which feature load was roughly commensurate with location load (see Chapter 3; also c.f. Awh et al., 2007). Each trial started with a fixation period for 500 ms, after which a memory array with two to five different colored circles was presented for 1000 ms. After a delay of 1000 ms, either a cue appeared simultaneously with a probe circle presented at the same location (simu-cue trials), or the cue appeared alone for 100 ms, followed by an additional delay of 400 ms before the probe circle appeared (retro-cue trials). In either case, the probe remained visible until a response was made and was equally likely to be a match or non-match to the circle that had appeared at that location in the memory array. The response initiated a 500 ms ITI. Cue types and set sizes were fully counterbalanced and randomly intermixed.

The trial timing for unfamiliar shapes, presented in the second part of the experimental session, was similar to that for the simple colored circles (see Figure 7.1). Memory arrays were comprised of anywhere from two to five of the shapes, randomly drawn from the green and blue shapes with the constraint that one color did not outnumber the other by more than one shape (e.g. a set size of five had to have three green and two blue shapes or vice versa). Retro-cue and simu-cue trials were again balanced across the four set sizes, and the single probe shape was equally likely to be a match or non-match to the shape previously presented at that location. Non-matching probes, however, could be drawn either from the remaining same-colored shapes (same-category change) or from the remaining other-colored shapes (cross-category change). For example, on a trial where a green shape changed at probe, a same-category change was another green shape that had not appeared in the memory array (see Figure 7.1a), and a cross-category change was a blue shape that had not appeared in the memory array (see Figure 7.1b). This was equally likely to occur for retro-cue and simu-cue trials, with cue type and change type counterbalanced across all set sizes.
Figure 7.1 Illustration of event sequences for (a) same-category and (b) cross-category change trials in Experiment 6.
As with the other experiments, participants received full instructions and completed a brief practice session for each part of the session (i.e., separate instructions and practice for simple colors and for unfamiliar shapes). Participants were able to discriminate all stimuli when they were previewed simultaneously during task instructions. They were instructed to remember all items presented during the memory array, were told that the cue would always indicate which memory item was being tested, and thus were advised that the cue could help them when it appeared in advance of the probe. In the case of the unfamiliar shapes, however, they were not explicitly informed that non-matching probes could be either same-category or cross-category changes.

7.3 Results

7.3.1 Capacity

$K$-estimates were computed separately for simple colors and for unfamiliar shapes; in the case of the complex shapes, separate capacities were estimated for same-category and cross-category changes using the same set of matching trials. Two measures of baseline capacity were generated for each participant: simple capacity was estimated as the highest $K$ achieved on simu-cue trials using simple colored circles, and complex capacity was estimated as the highest $K$ for simu-cue trials with same-category changes. The average simple baseline capacity across individuals ranged from 2.40 to 4.50 (mean ± standard deviation = 3.32 ± 0.09), and the average complex baseline capacity ranged from 1.52 to 3.85 (2.46 ± 0.12). The two baseline measures were moderately but significantly correlated across individuals, $r = 0.42$, $p < 0.02$ (see Figure 7.2 for all correlations). In the case of simple colors, 20 of the 30 participants showed larger estimates for retro-cue trials relative to simu-cue trials (average improvement in $K$ across all participants of 0.41 ± 0.11; 0.77 ± 0.12 when restricted to those who showed an improvement). Capacity for
Figure 7.2 Correlations between individual capacity estimates for (a) colors on simu-cue trials and shapes on simu-cue, same-category change trials, (b) colors on simu-cue trials and shapes on simu-cue, cross-category change trials, and (c) shapes on simu-cue, same-category change trials and shapes on simu-cue, cross-category change trials in Experiment 6, with fits for the regression lines shown at the top of each scatter plot.
complex shapes also improved following a *retro-cue* for 20 of the 30 participants (average improvement in $K$ collapsed across *same-* and *cross-category* changes of $0.18 \pm 0.08; 0.39 \pm 0.06$ for only those who showed an improvement; see Figure 7.3a for group means). In addition to the correlation between baseline capacities for simple shapes and for *same-category* complex shapes, baseline capacity for simple shapes was correlated significantly with capacity estimates for *cross-category* complex stimuli, $r = 0.54, p < 0.002$; *same-* and *cross-category* complex capacities also showed a strong correlation, $r = 0.86, p < 0.0001$ (see Figure 7.2).

A Cue type X Set size ANOVA for color $K$-estimates showed significant effects of both Cue type and Set size, $F(1,29) = 17.10, \text{MSE} = 0.36, p < 0.0001$, and $F(3,87) = 62.72, \text{MSE} = 0.30, p < 0.0001$, respectively. As with the other behavioural studies, the interaction between Cue type and Set size was significant, $F(3,87) = 5.69, \text{MSE} = 0.40, p < 0.001$, and interacted significantly with simple baseline capacity, $F(3,84) = 3.55, \text{MSE} = 0.37, p < 0.02$. The $K$-estimates for complex shapes were submitted to a three-way ANOVA to examine the effects of Cue type (*retro-* vs. *simu-cue*), Change type (*same-* vs. *cross-category*), and Set size (2 through 5). All three factors had significant main effects on estimated capacity. $K$-estimates were overall higher for *retro-cue* trials than for *simu-cue* trials, $F(1,29) = 27.87, \text{MSE} = 0.22, p < 0.0001$, were higher for *cross-category* changes than for *same-category* changes, $F(1,29) = 85.78, \text{MSE} = 0.18, p < 0.0001$, and increased with increasing Set size, $F(3,87) = 24.00, \text{MSE} = 0.60, p < 0.0001$.

Both Cue type and Change type had some interaction with Set size, $F(3,87) = 2.36, \text{MSE} = 0.22, p < 0.08$, and $F(3,87) = 20.50, \text{MSE} = 0.06, p < 0.0001$, respectively. These interactions were modulated when baseline capacity was entered to the analysis $F(3,84) = 1.22, \text{MSE} = 0.22, p > 0.30$, and $F(3,84) = 0.03, \text{MSE} = 0.06, p > 0.30$, respectively. Critically, the effects of Cue type and Change type did not interact significantly with one another, $F(1,29) = 2.52, \text{MSE} = 0.10, p >
Figure 7.3 Group mean (a) capacity and (b) response time across set sizes in Experiment 6, with error bars representing the standard error of each mean.
0.10, indicating that the two had independent effects on estimated capacity. Likewise, the three-way interaction term was not significant, $F(3,87) = 0.83, \text{MSE} = 0.05, p > 0.40$. Thus while the effects of both Cue type and Change type were influenced by the point at which baseline capacity was exceeded, these effects were additive and there was no point at which the magnitude of one was dependent on the other. This was confirmed by the lack of a significant correlation between the retro-cue benefit and the cross-category benefit across individuals, $r = 0.30, p > 0.10$.

7.3.2 Response time

RTs from all correct trials for the simple colors were submitted to a Cue type X Set size ANOVA (see Figure 7.3b for group means), which showed significant effects of Cue type and of Set size, $F(1,29) = 170.94, \text{MSE} = 8831.26, p < 0.0001$, and $F(3,87) = 21.20, \text{MSE} = 5367.14, p < 0.0001$, respectively. The interaction between the two factors was also marginally significant, $F(3,87) = 2.41, \text{MSE} = 3605.02, p < 0.07$. Thus, as observed in the foregoing behavioural studies, RTs were shorter following retro-cues, and the slope of increasing RTs as Set size increased was shallower for retro-cue trials. The same pattern regarding retro-cues was upheld for complex shapes. The three-way ANOVA on RTs exhibited a main effect of Cue type, $F(1,29) = 495.41, \text{MSE} = 6433.58, p < 0.0001$, and a smaller albeit significant effect of Change type, $F(1,29) = 17.88, \text{MSE} = 689.47, p < 0.0001$, as well as an effect of Set size, $F(3,87) = 77.70, \text{MSE} = 3783.28, p < 0.0001$. That is, RTs were shorter overall following retro-cues, shorter after cross-category changes, and longer for larger Set sizes. Again, Cue type and Change type did not interact significantly, $F(1,29) = 1.41, \text{MSE} = 306.88, p > 0.20$, and in this case Change type did not interact with Set size, $F(3,87) = 1.59, \text{MSE} = 737.27, p > 0.10$, nor did the three-way interaction achieve significance, $F(3,87) = 0.75, \text{MSE} = 433.03, p > 0.5$. Cue type and Set size,
on the other hand, did have a significant interaction effect on RT, $F(3,87) = 6.13$, $MSE = 1742.96$, $p < 0.001$, demonstrating a shallower RT slope with increasing load for retro-cue relative to simu-cue trials, regardless of Change type, congruent with the results from the previous chapters.

7.4 Discussion

The key findings from Exp 6 are that cross-category changes and retro-cues both exhibited strong yet independent (i.e., additive) effects on estimated capacity for complex stimuli: While both increased the measures of estimated storage capacity, the former did so by reducing probe comparison errors due to low feature resolution and the latter by biasing VSTM resources to the probe location, with no interactions between the two. That is, retro-cues operated online but only on representations within location VSTM, and cross-category changes acted offline only on representations in feature VSTM. Furthermore, the comparison RT was influenced independently by retro-cues and cross-category changes, with much larger effects for the former. Thus, selection of a particular location representation in VSTM facilitates probe comparison at that location, but only with respect to the location aspect of the comparison, as it does not affect the resolution of the features that correspond to that location in VSTM. This provides further support for a two-system model of VSTM, but one in which feature and location representations were stored independently rather than interactively, such that changes to one system had no online influence on the other system.

Awh et al. (2007) argued for a two-system model of VSTM storage, but one in which one system corresponds to an absolute number of items being maintained (i.e., fixed slots) and the other system is responsible for the resolution of those items’ representations. A primary contribution to this interpretation was the strong correlation between the estimated capacity for simple stimuli
that held a one-to-one correspondence between features and locations (i.e., colored circles) and complex stimuli tested under conditions in which the impact of item resolution was minimized (i.e., cross-category changes). That is, under reduced likelihood of committing comparison errors due to difficulty in discriminating complex stimuli, capacity seemed unaffected by feature load, supporting the argument that a given number of locations were maintained in VSTM regardless of the complexity of the items appearing at those locations. The results from the current experiment corroborate this argument, particularly given the significant correlation between baseline capacities for simple colors and for complex shapes. The interpretation supported here diverges from that of Awh et al. (2007), however, in that the current results demonstrate that location VSTM resources could be reallocated flexibly to a particular location representation. Moreover, location selection does not change feature resolution, showing that feature representations are stored independently of location VSTM.

Taken together with the findings discussed in the preceding chapters, these data collectively support a model in which locations and features are maintained in separate VSTM stores in which the likelihood of accurately judging whether the probe matches any given representation is determined by the amount of resources dedicated to maintaining that representation; these resources can be shifted via attentional selection generated by a retrospective cue, facilitating the comparison process. The current proposal of a two-system account of VSTM is consistent with models advanced by proponents of fixed slots (e.g., Awh et al., 2007), while the characterization of functional capacity as the relative weighting among competing items is more congruent with supporters of the flexible resources model (e.g., Alvarez & Cavanagh, 2004). The next, final chapter returns to the distinctions made by these competing theories in light of all the experimental data.
8 General discussion and conclusions

8.1 The nature of representation in VSTM

The work in this dissertation addressed the nature of VSTM representations with respect both to what information is maintained and to how that information is maintained during the retention interval of change-detection tasks. This was achieved using both ERP and several behavioural manipulations of a retrospective cueing paradigm, which all converged on a model in which features and locations are stored flexibly and independently, such that shifts in resources within one system affect the other only at the point of probe comparison.

8.2 Summary and integration of findings

The initial study (Exp 1) characterized VSTM load by examining whether the contralateral delay activity (CDA) reflected the number of locations or the number of features to be remembered during a lateralized change-detection task. The behavioural and ERP results indicated that VSTM load was comprised of both features and locations; critically, the effects of locations and features on CDA amplitude had distinct topographical distributions, in line with previous fMRI studies showing different brain areas responsible for individuation and identification (e.g., Xu & Chun, 2006). Having established that locations and features were represented in separate VSTM systems, the subsequent behavioural studies used the retro-cue paradigm to investigate how representations were stored within these systems and whether the two interacted online. The first of these (Exps 2a-c) demonstrated the basic effects of retro-cues, wherein providing attentional selection for one aspect of VSTM storage led to shorter response times when the probe appeared and increased capacity estimates once the set size exceeded an individual’s baseline capacity. The retro-cue benefit was reciprocal between locations and features, in that advanced selection
within either system generated comparable effects. Whereas the initial step on these change-detection tasks would be to narrow the comparison to one relevant item, determined by either a specific location or a particular feature, a retro-cue within either one of these systems allowed this initial step of probe comparison to proceed even before a full probe was provided. In other words, the selection aspect of probe comparison proceeded in advance by shifting VSTM resources to the relevant location or feature representation via a retro-cue, thus facilitating the full comparison that ensued once the probe appeared.

The next study (Exps 3a-b) showed that VSTM resources were flexibly reallocated by the retro-cue such that the relative weighting between representations was shifted, as increases in performance for the cued representation were accompanied by detriments to performance for the remaining, non-cued items. While the magnitude of the retro-cue benefit was not altered by the predictive validity of cues across tasks, the spatial selectivity of the retro-cue was necessary for the boost in estimated capacity. As seen when two retro-cues were presented (Exp 4a), only the probe comparison RT was impacted when more than one location was selected. That is, multiple location representations were biased by multiple cues, effectively narrowing the search space in advance of the probe appearance, but the uncertainty over which location would be probed did not allow that aspect of the probe comparison to proceed in advance, thus not affecting the ultimate comparison load and failing to boost capacity estimates. This pattern in comparison RT was seen again when multiple locations were cued sequentially (Exp 4b), in which case the most recently cued location did show increased capacity estimates, showing the malleability of resources in VSTM and again demonstrating that spatial specificity was critical. The following study (Exp 5) further confirmed that spatial specificity was requisite for the boost in estimated capacity by demonstrating that a general alerting signal could account for neither the scale of the RT effects following retro-cues nor for the capacity effects.
The research culminated in the final manipulation of the *retro-cue* paradigm, designed to examine whether a location cue affected feature resolution (*Exp 6*), which would indicate whether location and feature VSTM interacted online or were integrated only at probe comparison. This was done by introducing different change types for non-matching probes in a task using complex stimuli; increased capacity estimates for *cross-category* relative to *same-category* changes on non-matching probes reflected a decrease in the probability of comparison errors by making the probe more discriminable from the VSTM representation. If location and feature representations interacted online, then selecting a specific location via a *retro-cue* would have increased feature resolution; an interaction between *retro-cues* and *cross-category* changes would have been observed. This was not the case, as the two effects were additive, showing that location *retro-cues* acted only on the location VSTM representations, while *cross-category* changes reflected only feature VSTM representations. Thus not only were location and feature representation maintained in separate VSTM systems, with flexible allocation of resources within each system, but the two did not interact online.

### 8.3 Relation of current results to previous theories

The experimental findings support a two-system model of VSTM in which location and feature representations are maintained in independent VSTM stores, with flexible allocation of resources that can be shifted during maintenance via attentional selection. There are two main, competing theories that must be addressed in light of the data presented in this dissertation, which essentially supports a particular hybrid of the two. The first of these positions is that of fixed slots. The most viable of the fixed slots models is the current formulation from Awh, Vogel and colleagues (Awh et al., 2007; Barton et al., 2009), a two-system model in which capacity is determined by an absolute number of objects that can be maintained, with resolution for those slots determined by the complexity of the features that constitute the objects. This theory already
grants that the functional capacity of VSTM is ultimately bounded by the lower limit of the two systems, which is usually determined by feature complexity. However, this still firmly classifies as a fixed slots approach because it holds that only a given number of items can be represented in the first place. Any items presented in excess of this fundamental limit would not be represented in VSTM at all. By contrast, flexible resource models such as that promoted by Alvarez and Cavanagh (2004) contend that VSTM resources are distributed over as many items as are presented (see also Bays & Husain, 2008). This does not mean, however, that there is no observable limit to the number of objects that can be maintained effectively, as eventually the representational strength of any one item as set sizes increase will be below the threshold necessary for accurate comparison to a probe.

Each of these theories presents some compelling elements given the present data, but neither alone can fully account for the results. First, the current fixed slots model does have an advantage in that it specifies two systems for VSTM representations, one that ultimately corresponds to locations and the other to features. As has been noted in several chapters, there is now ample evidence that locations and features are maintained in different VSTM systems, not only from previous research such as that of Xu and Chun (2006), but also from present results using ERP (Exp 1) and behavioural manipulations (Exp 2a-c and 6). However, some aspects of two-system model promoted by Awh et al. (2007) cannot account for some of the novel findings presented here. One key discrepancy is that the fixed slots model does not allow reciprocity between location storage and feature resolution on overall change-detection performance; rather, feature resolution is applied to slots in a more hierarchical fashion, such that slots in a location-based system are first occupied by objects, and then resources in a feature-based system code for the features present in those objects. However, Exp 2a-c show that the effects of attentional selection on change-detection performance are comparable in either system and Exp 6 shows that
feature storage is independent of location storage. Thus VSTM maintenance is not inherently hierarchical, violating a tenet of this fixed slots model. More importantly, the idea that any part of VSTM is fixed is problematic, as Exps 3a-c and 4b show that representational weights are reallocated such that cued items benefit at the cost of non-cued items and that these weights can be shifted multiple times. A potential argument is that the retro-cue guarantees that the cued item occupies a slot. The data do not support such an argument. For most individuals, two simple objects are well within capacity limits, so guaranteeing selection of two items should lead to maximal change-detection performance; presenting two retro-cues in Exp 4a, however, did not generate any capacity benefit. Therefore items are not simply moving in and out of slots but are represented with variable weights that are subject to flexible reallocation.

Given the objections to a fixed slots approach, a flexible resources model is favored. The main problem with the current flexible allocation theory is that it makes no distinction between systems for different aspects of VSTM load. On the contrary, any aspect of a VSTM representation is drawn from a common pool of resources, with visual load defined as a tradeoff between number of objects and complexity of those objects. This contradicts findings that show separate response profiles for different brain areas, including the results of Exp 1 that demonstrate how feature and location aspects of visual stimuli are maintained separately and of Exp 6 that show functional independence of these two systems. Thus the studies presented here support the idea that VSTM resources are allocated flexibly, with success on a change-detection task determined by whether sufficient VSTM weights were assigned to the tested item. The proposed model, however, presses beyond the notion of a unified pool of resources to specify distinct pools for location and feature aspects of VSTM, with flexible allocation of resources within each pool.
In sum, the studies in this dissertation address a standing controversy in the literature regarding the nature of VSTM representations, with evidence for independent systems for feature and location maintenance and for flexible allocation of resources within each system. The proposed model incorporates separate systems for the storage of locations and of feature information, similar to previous theories (Awh et al., 2007; Xu & Chun, 2006) but with the novel stipulation that representations in these systems are maintained independently. Furthermore, the strength of representations within these systems is determined by the flexible allocation of resources, comparable to other models (Alvarez & Cavanagh, 2004; Bays & Hussain, 2008) but with the specification of different resource pools for different aspects of visual load.

8.4 Future directions
The change-detection tasks used here presented single probes that necessarily matched a memory item with respect to either location or feature. The manipulation of retrospective cues then allowed these two aspects of VSTM maintenance to be disambiguated, in line with the ERP evidence for distinct systems for locations and features. The final experiment then demonstrated that these two systems do not interact online, with attentional selection for one attribute affecting only the representations within that system. That is, locations and features are represented in separate VSTM systems. These studies make no further distinction for multiple systems depending on specific feature categories. Yet it is conceivable that multiple systems exist, one for each stimulus dimension, in line with theories claiming that working memory emerges from sustained activity in the various brain areas that initially code for visual attributes (Magnussen & Greenlee, 1999; Postle, 2006). If this were the case, then the retrospective cueing technique may prove fruitful in dissociating these systems. For example, stimuli could have both colors and shapes; while the probe would necessarily match an item from the memory array on one dimension (e.g., color), participants would indicate whether it matched with respect to other
dimension (e.g., shape). If these two feature dimensions are stored in independent VSTM systems, then using a retrospective color cue should facilitate probe comparisons RT for all set sizes but have an effect on capacity estimates only once baseline capacity for shapes has been exceeded. The studies presented here provide some precedent for using individual differences in this way to help dissociate VSTM systems.

The model that emerges from the present studies may also make some interesting predictions regarding the response of different brain areas to VSTM storage of locations and features. Based on neuroimaging data during change-detection tasks, a popular view is that the iIPS is involved in object individuation while the sIPS and LOC are involved in object identification (Xu, 2008). In other words, one system separates items based on locations and the other classifies items based on features. The current conception of how these systems operate follows the hierarchical account described above where objects must be individuated first and subsequently identified, akin to designating slots and then filling them in with feature information. However, the results presented here dispute such a framework by showing that selection within feature storage can occur antecedent to judgment of a probe location (see Exp 2c). That is, a sort of feature individuation may be observed prior to location identification. Xu’s (2008) study most closely resembled Exp 2b, where the task ultimately depended on judging whether the probe matched an item from the memory array with respect to feature. The argument was that encoding of location was incidental but compulsory, as shown by the increasing activity in iIPS with increasing number of objects (Xu, 2008). A similar task could be implemented with the key difference that responses should be made with respect to location rather than feature. If location individuation is a mandatory first step, then that process should be sufficient to complete a probe comparison based on location and there should be no modulation of activity in sIPS depending on feature load. On the other hand, if activation corresponding to feature load is observed in sIPS even
when the task does not depend on feature identity, this would provide further evidence that representations in the two VSTM systems are consolidated in tandem.

Two lateralized ERP components, namely the N2pc and the Ptc, may help further delineate any potential interactions and directionality between feature and location VSTM representations. The N2pc reflects spatial differentiation of stimuli (Eimer, 1996), and Kuo et al. (2009) demonstrated that N2pc responses are comparable for searches through perceptual stimuli and for searches through representations held in VSTM. Recently, Hilimire, Mounts, Parks, and Corballis (2009) documented the Ptc, which was elicited when close proximity of stimuli in a visual display necessitated additional discrimination of items based on features. An equivalent response may be observed for searching through items maintained in VSTM, where N2pc and Ptc amplitudes following a probe are inversely related to one another depending on spatial overlap and feature resolution, respectively. If this is the case, and if the current model for dissociable systems holds, then the two components should be modulated independently by location versus feature retrospective cues. That is, a location cue during the VSTM delay should attenuate the N2pc following the probe but have no effect on the Ptc; vice versa, a feature cue during the VSTM delay should attenuate the Ptc but have no effect on the N2pc. Such restricted effects would provide further evidence that resources for locations and features are allocated independently during VSTM maintenance.

Despite the demonstration of reciprocity between location storage and feature storage, there is an apparent asymmetry between the two systems regarding the resolution of stored representations. The distinction between simple and complex features is well supported, but it is not clear what the analogue would be in terms of locations, particularly given the 1:1 correspondence between simple stimuli and locations. The proposed model purports that the difference in capacity
estimates for simple and complex stimuli arises from flexible allocation of VSTM weights in the feature system, with successful change detection when the resolution of the representation provides a sufficient level of discriminability. The model furthermore posits a comparable flexibility of resource allocation within the location system. A specific test of this claim could involve a retrospective cueing task that probes for location, similar to Exp 2c, but one in which the number of test locations given the overall display size necessitates much greater resolution of location information. That is, doubling the number of potential locations within the same area as Exp 2c may make discrimination among locations more difficult, resulting in estimates for location capacity that are more comparable to those for complex features. Such a finding would further support the assertion that representational resources within location VSTM are subject to the same flexible allocation as those within feature VSTM.

8.5 Conclusion

Collectively, the studies presented in this dissertation show that location information and feature information are maintained in behaviourally and anatomically dissociable VSTM systems. Representations within each system are subject to attentional selection that facilitates probe comparison and ultimately enhances capacity estimates when the memory load exceeds an individual’s usual limit. The results presented in this dissertation converge nicely with the existing neuroimaging literature regarding dissociable storage areas within VSTM, yet push further to better understand the distribution and malleability of weights within each system, with no online interactions between separate systems. Thus the proposed model of VSTM incorporates independent stores for locations and features, with flexible allocation of resources within each system.
References


## Appendix 1 Basic demographic information, including age, sex, and handedness, for all experiments.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>N</th>
<th>Age range in years (Mean ± std.dev.)</th>
<th>Males:Females</th>
<th>RH:LH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>19 – 31 (24.11 ± 4.50)</td>
<td>5:10</td>
<td>15:0</td>
</tr>
<tr>
<td>2a</td>
<td>24</td>
<td>19 – 33 (21.96 ± 3.09)</td>
<td>12:12</td>
<td>21:3</td>
</tr>
<tr>
<td>2b</td>
<td>24</td>
<td>18 - 27 (20.81 ± 2.32)</td>
<td>8:16</td>
<td>21:3</td>
</tr>
<tr>
<td>2c</td>
<td>24</td>
<td>18 - 28 (21.79 ± 2.55)</td>
<td>5:19</td>
<td>23:1</td>
</tr>
<tr>
<td>3a</td>
<td>24</td>
<td>18 - 22 (20.51 ± 1.06)</td>
<td>6:18</td>
<td>22:2</td>
</tr>
<tr>
<td>3b</td>
<td>24</td>
<td>19 - 34 (21.27 ± 3.03)</td>
<td>2:22</td>
<td>24:0</td>
</tr>
<tr>
<td>4a</td>
<td>24</td>
<td>18 - 25 (20.46 ± 1.51)</td>
<td>4:20</td>
<td>21:3</td>
</tr>
<tr>
<td>4b</td>
<td>30</td>
<td>18 - 30 (20.75 ± 2.23)</td>
<td>7:23</td>
<td>30:0</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>18 - 32 (21.22 ± 3.53)</td>
<td>8:13</td>
<td>20:1</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>19 - 36 (22.01 ± 3.73)</td>
<td>13:17</td>
<td>23:7</td>
</tr>
</tbody>
</table>
Appendix 2 Stimuli used in the experiments. (a) All simple shapes, colors, and orientations used in Experiment 1, with (b) samples of multi-feature stimuli created by combining simple features. (c) Simple colored circles used in Experiments 2a, 3a, 3b, 4a, 4b, 5, and 6. (d) Achromatic shapes used in Experiments 2b and 2c (first two columns only). (e) Colored shapes used in Experiment 6; half the participants saw these stimuli, and for the other half the colors were reversed.
Appendix 2. continued

(b)
Appendix 2. continued

(c)

<table>
<thead>
<tr>
<th>Color</th>
<th>RGB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown</td>
<td>115,69,56</td>
</tr>
<tr>
<td>Red</td>
<td>218,37,29</td>
</tr>
<tr>
<td>Orange</td>
<td>231,120,23</td>
</tr>
<tr>
<td>Yellow</td>
<td>248,195,0</td>
</tr>
<tr>
<td>Green</td>
<td>0,146,63</td>
</tr>
<tr>
<td>Blue</td>
<td>0,124,195</td>
</tr>
<tr>
<td>Purple</td>
<td>40,22,111</td>
</tr>
<tr>
<td>Magenta</td>
<td>144,30,120</td>
</tr>
<tr>
<td>White</td>
<td>220,220,220</td>
</tr>
</tbody>
</table>
Appendix 2. continued

(d)
Appendix 2. continued

(e)