

DISSERTATION

THE IMPACT OF LONG-TERM VISUAL REPRESENTATIONS ON  
CONSOLIDATION IN VISUAL WORKING MEMORY

Submitted by

Lisa Durrance Blalock

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2010

COLORADO STATE UNIVERSITY

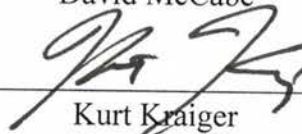
June 25, 2010

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY LISA DURRANCE BLALOCK ENTITLED THE IMPACT OF LONG-TERM VISUAL REPRESENTATIONS ON CONSOLIDATION IN VISUAL WORKING MEMORY BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work



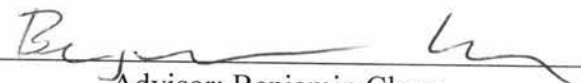
David McCabe



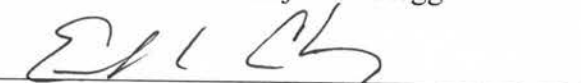
Kurt Kraiger



Charles A.P. Smith



Advisor: Benjamin Clegg



Department Chair: Ernest Chavez

## ABSTRACT OF DISSERTATION

### THE IMPACT OF LONG-TERM VISUAL REPRESENTATIONS ON CONSOLIDATION IN VISUAL WORKING MEMORY

Visual working memory (VWM) is the cognitive mechanism that encodes, maintains, manipulates, and retrieves visual and spatial information over the short-term. VWM is distinct from the more temporary and perceptual sensory store which has a high capacity but can only maintain information for a few hundred milliseconds (Sperling, 1960). In order to prevent visual representations in VWM from being overwritten by subsequent visual information, some process must stabilize the activated representations so that they are less susceptible to interference. This process has been termed short-term consolidation (Jolicoeur & Dell'Acqua, 1998; Vogel, Woodman, & Luck, 2006).

The current dissertation examines this short-term consolidation process, specifically whether or not long-term visual representations facilitate the stabilization of visual representations in working memory. Currently, WM models vary greatly in how they account for the relationship between long-term memory (LTM) and working memory (WM). The current experiments contrasted three models of WM in terms of how they could account for early effects of LTM in VWM processing: the multiple-component model (e.g., Baddeley & Logie, 1999), the embedded-processes model (e.g., Cowan, 1999), and the newer workspace model (e.g., Logie 2003).

Four experiments were conducted that compared performance between familiar and unfamiliar stimuli in a backward masking paradigm that provided information on the time course of consolidation as well as how the process varied depending on the number of to-be-remembered items in the memory array. In the backward masking paradigm, a set of visual masks are presented at varying intervals following a memory array (17, 134, 250, 367, and 484 ms) before the test array is shown and participants make a same/different judgement (i.e., did one of the items change between study and test). Varying the presentation of visual masks provides insight into the consolidation process: masks shown soon after the memory array will decrease performance because the activated representations in VWM have not been stabilized to reduce interference from subsequent visual information. Additionally, the set size of the visual array was manipulated (1, 2, 3, or 4 items) to examine how the impact of the visual masks varied depending on how many items there were in the array.

In Experiment 1, simple, colored squares were compared to complex, random polygons as an extreme comparison between stimuli that vary in terms of familiarity. This experiment showed an advantage for squares over polygons in both accuracy and capacity. Additionally, familiar squares were consolidated faster into VWM than random polygons. Experiments 2a and 2b equated for the type of stimulus by training on a subset of random polygons and comparing those trained items with novel items. In Experiment 2a, participants trained using a four alternative forced choice task (in which participants view an item then pick that item out of a set of four) and an advantage was demonstrated for trained items in both accuracy and capacity as well as faster consolidation for trained

items. Additionally, there was no interaction between set size and delay for trained items, suggesting that participants were able to consolidate the array as a single unit instead of requiring more time for more items. In Experiment 2b, participants trained using a change detection task and no differences were demonstrated between trained and untrained items, including no interaction between set size and delay for either stimulus type, demonstrating transfer of change detection training to novel stimuli and evidence that participants created a unitary configuration of the items with increased training. To examine the impact of creating a configuration of items, Experiment 4 used an incidental learning paradigm in which the same four shapes were shown in the same spatial locations on every third trial and those repeating items were compared with novel items. With this repeating context, familiar items were remembered better than novel items and the impact of masks did not vary depending on the number of items for familiar items only. This supports the implication from Experiments 2a and 2b that spatial relations between items is part of the consolidation process.

Overall, these results support both the Unitary and Workspace Models of WM. Additionally, there are three primary conclusions from the current set of experiments. First, long-term memory plays an early role in VWM by facilitating the process involved in stabilizing WM representations. Second, short-term consolidation is a multidimensional process that strengthens both visual and spatial relational information to create more stable representations of perceptual input. Finally, training of visual stimuli can transfer to novel stimuli depending on the type of training task used and how similar it is to the criterion task. Taken together, these conclusions provide strong

evidence that WM and LTM are highly interactive and provide unique insights as to how these processes work together to facilitate everyday, complex cognition.

Lisa Durrance Blalock  
Department of Psychology  
Colorado State University  
Fort Collins, CO 80523  
Summer 2010

## TABLE OF CONTENTS

CHAPTER 1 - INTRODUCTION.....	1
Working Memory.....	1
Multiple-component/Gateway model.....	2
Embedded Processes/Unitary model.....	5
Workspace Model.....	7
Summary.....	11
Properties of Visual Working Memory.....	13
Measuring Visual Working Memory.....	14
Capacity and Representation in Visual Working Memory.....	15
Consolidation in Visual Working Memory.....	20
Summary.....	25
Interactions Between VWM & LTM.....	26
Explicit Training of Visual Stimuli.....	27
Implicit Training of Visual Stimuli.....	28
Inherently Familiar versus Unfamiliar Stimuli.....	29
Summary.....	31
Current Experiments.....	31
CHAPTER 2 - EXPERIMENT 1.....	34
Method.....	35
Participants and Design.....	35

Stimuli.....	35
Procedure.....	37
Results.....	38
Accuracy .....	39
Capacity .....	45
Discussion .....	46
CHAPTER 3 - EXPERIMENT 2A.....	50
Method.....	52
Participants and Design.....	52
Stimuli.....	52
Procedure.....	52
Results.....	54
4AFC & Recognition Task Performance .....	54
Accuracy .....	55
Capacity .....	60
Discussion .....	61
CHAPTER 4 - EXPERIMENT 2B.....	65
Method.....	66
Participants.....	66
Training.....	67
Results.....	67
Change Detection Training & Recognition Performance .....	67
Accuracy .....	69
Conditionalized Accuracy Analysis .....	73

Capacity .....	77
Discussion .....	77
CHAPTER 5 - EXPERIMENT 3 .....	82
Method .....	85
Participants and Design .....	85
Stimuli .....	85
Procedure .....	86
Results .....	87
Recognition Test Accuracy .....	87
Accuracy .....	87
Capacity .....	92
Discussion .....	92
CHAPTER 6 - GENERAL DISCUSSION .....	96
What does this tell us about the relationship between WM and LTM? .....	98
Model Overview & Fit With Current Data .....	99
An Alternative View: Cue-based Models of WM .....	106
Summary .....	108
What does this tell us about short-term consolidation? .....	108
What is being consolidated? .....	108
How is information being consolidated? .....	111
Summary .....	116
Concluding Remarks .....	116
CHAPTER 7 - REFERENCES .....	118

## CHAPTER 1 - INTRODUCTION

Over the last decade, there has been a surge in research examining the encoding, maintenance, organization, and retrieval of visual information over the short-term, a process termed visual working memory (VWM, or perhaps better defined as working memory for visual information). VWM has proved to be crucial for a wide range of cognitive tasks from facilitating gaze corrections in eye movements that assist with goal directed visual search (Hollingworth, Richard, & Luck, 2008) to complex dynamic tasks requiring updating of visual and spatial information (e.g., Wickens, Vincow, & Yeh, 2005). The current body of literature on VWM has focused primarily on the capacity of VWM and the nature of visual representations in VWM (e.g., Jiang, Olson, & Chun, 2000; Luck & Vogel, 1997). However, recently research has begun to examine the processes underlying creating, maintaining, and accessing visual representations, as well as how VWM interacts with visual representations in long-term memory (Luck, 2008). The current paper will examine how VWM representations are created following perceptual input and how that varies depending on the complexity and familiarity of the stimuli.

### **Working Memory**

At the most general level, working memory (WM) can be defined as a temporary storage structure that maintains goal relevant information for any given task (Baddeley,

2007; Miyake & Shah, 1999). WM is critical for complex cognition and underlies literally any everyday cognitive task from visualizing how a picture will look in your living room to comprehending a complex statistical analysis. Because of WM's crucial role in human cognition, it has been a primary focus in the cognitive psychology literature resulting in a wide range of theories that vary everything from what controls the information in WM to how many WM systems there are (see Miyake & Shah, 1999 for an overview of these different theories). The following review will focus on three main theories of WM, the multiple-component model, the embedded-processes model, and the workspace model with an emphasis on how the models account for the relationship between WM and representations in long-term memory (LTM).

### **Multiple-component/Gateway model**

One of the most influential WM models was originally proposed by Baddeley and Hitch (1974). In their paper, Baddeley and Hitch (1974) challenged the accepted model of the time, the modal model (Atkinson & Shiffrin, 1971), which argued that short-term memory was a unitary construct that both maintained and manipulated information (Baddeley, 2007). Baddeley and Hitch (1974) demonstrated through a series of experiments that this modal view was overly simplistic. Rather, they argued that the manipulation of information within WM is done by a separate control system (termed the central executive) and that information is maintained in WM via two separate passive storage mechanisms for verbal (termed the phonological loop) and visual (termed the visuospatial sketchpad) information.

While these three basic components are still in place, Baddeley has modified the multiple-component model some over the last few decades (e.g., Baddeley, 2000; Baddeley, 2003; Baddeley, 2007; Baddeley & Logie, 1999) by further specifying the storage subcomponents and by including a new component that binds together information from different modalities. In terms of the storage subcomponents, they have now been further subdivided in to two systems: an active rehearsal mechanism (the inner scribe for refreshing visual information and inner speech mechanism for rehearsing verbal information) and a passive storage mechanism (the visual cache for visual information and the phonological store for verbal information; Baddeley & Logie, 1999). Baddeley (2000) also proposed the addition of a new component to the multiple component model, the episodic buffer, which acts as a limited-capacity, temporary storage system that integrates information from different modalities and from LTM into a combined memory episode.

While not explicit in the multiple-component model, the model does work on the implicit assumption that WM is a gateway between perception and LTM, similar to the modal model (Atkinson & Shiffrin, 1971). As shown in Figure 1 (panel A), the individual storage subsystems interact with long-term representations after directed by the central executive, suggesting that as information is perceived in the environment, it first enters the functionally and structurally distinct WM system before accessing any long-term representations. Because of this, this model will be referred to as the gateway model for the remainder of the paper, using the terminology from van der Meulen, Logie, and Della Sala (2009).

Despite being a widely accepted model, it has been criticized. In particular, while Baddeley and Hitch (1974) cited neuropsychological evidence demonstrating dissociations between WM and LTM, several researchers have argued for a unitary model (e.g., Cowan, 1988, 1999; Nairne, 2002). These unitary models have critiqued the gateway model because it cannot easily account for LTM effects within WM. For example, research has shown that high-frequency and concrete words are remembered better than low-frequency and abstract words (Hulme, Maughan, & Brown, 1991; Walker & Hulme, 1999). While Baddeley recently added the Episodic Buffer component to the Gateway Model to account for some of these effects, there is currently little empirical support for that component.

Additionally, the assumption of domain specific storage has been criticized. While behavioral and neuropsychological double dissociations have been demonstrated supporting a division between verbal and visuospatial (see Baddeley & Logie, 1999; Baddeley, 2007 for reviews), others have suggested that these dissociations do not demonstrate separate memory structures, but rather different mental codes that are better processed for processing domain specific information (Cowan, 1999). That is, while visual information is represented differently than verbal information in the brain, this does not necessitate separate systems.

From a theoretical standpoint, making domain specific distinctions can be problematic. Specifically, while Baddeley has argued for verbal and visual components, there is no reason to not also have separate buffers for kinesthetic or olfactory processing. At that point, there could potentially be hundreds of WM buffers for maintaining specific

types of information, essentially making the division between them meaningless (see also Tulving, 2007). The advantage of such a separation, however, is that it makes very clear predictions about what will influence WM performance. Thus, while the Gateway Model has been tremendously influential in directing WM research, it is not without criticisms and it has difficulties in constraining the number of buffers as well as accounting for LTM effects in WM.

### **Embedded Processes/Unitary model**

An alternative model to the Multiple-Component Model is the Embedded-Processes Model of WM. There are two key ways in which these two models differ. First, the Embedded-Processes account is more unitary, arguing that WM is a portion of activated LTM that is within the focus of attention (Cowan, 1988, 1993, 1999). Second, there are no subcomponents for different types of information. Instead, Cowan has argued that different codes exist for different types of information (i.e., verbal, visual, spatial, etc.) but that these codes all conform to similar activation patterns, temporal dynamics, and decay functions (Cowan, 1999). Thus, they should be viewed together as a whole.

The main structures within Cowan's model are LTM, the portion of LTM that is currently activated (i.e., activated memory), and the focus of attention, which is the portion of activated memory that is currently in awareness (Cowan, 1993, 1999). Additionally, there is a central executive component, which influences the focus of attention via effortful processes that direct the control of attention (Cowan, 1988, 1999). However, Cowan also suggests that automatic attention processes influence what enters the focus of attention. Specifically, he argues that when a novel stimulus appears,

attention will automatically be oriented to it and it will go directly into the focus of attention (Cowan, 1988, 1999). However, if a stimulus is continuously presented without any changes, (and no voluntary, top-down control to maintain attention), habituation will occur and those habituated stimuli will go into activated memory where they can be brought into the focus of attention by the central executive if they are needed (Cowan, 1988, 1999). Thus, WM emerges as a result from both automatic and effortful attentional processes that determine what enters the focus of attention. In that sense, Cowan's model incorporates a much larger (or at the very least, a more explicit) role of attention in WM processes.

As evident from the above discussion and Figure 1 (panel B), the embedded-processes model makes no structural distinction between WM and LTM. Rather, WM is just a different state of activation than information in LTM. Because of this, the embedded-processes model will be called the unitary model in this paper (again, similar to van der Meulen et al., 2009). The advantage of a unitary approach is that it can account for the long-term memory effects outlined above (e.g., word frequency effect). Additionally, taking a domain general approach to working memory (i.e., not having separate verbal and visual buffers) gets around the theoretical difficulties of having too many subsystems.

However, the Unitary Model is not without its own difficulties. Primarily, the Unitary Model has a difficult time accounting for neuropsychological dissociations in brain damaged patients, a key reason why the Gateway Model (and the Workspace Model outlined below) proposed separate memory systems. Additionally, the Unitary Model has

a difficult time explaining how novel items are learned (i.e., how does information go from WM to LTM?). If information does not already reside in LTM or activated LTM, it is not clear how this information would operate within the Unitary Model. This is likely because the majority of work examining the Unitary Model has used verbal tasks (though see Morey & Cowan, 2005 for an exception), where it is difficult to find completely novel stimuli. At a more theoretical level, the Unitary Model does not make as specific predictions as the Gateway Model about what influences short-term memory, making it a difficult model to test empirically. One exception to this is the specificity in how the central executive component determines the contents of the focus of attention, as well as how the central executive is involved with activating information in LTM. The relationship between the central executive and LTM is much more explicit than in either the Gateway Model outlined above or the Workspace Model outlined below.

### **Workspace Model**

Recently, Logie and colleagues (Logie, 2003; Logie & van der Meulen, 2009; van der Meulen, et al., 2009) have proposed the workspace model of WM in which perceptual information directly accesses LTM and then this activated long-term knowledge goes into WM which acts as a workspace that maintains and manipulates the activated representations. The WM system in the workspace model is a multiple-component system with separate stores for visual and verbal information and a central executive component that manipulates and coordinates the memory systems. Thus, this view is very similar to Baddeley's multiple-component model (indeed, the WM component of the workspace model is identical to Baddeley's model), but the crucial difference is that unlike

Baddeley's model, the WM system in the workspace model does not deal directly with perceptual or sensory information. Instead, "it deals with objects and shapes that have been identified by the processes of perception and that draw on our knowledge base of past experience" (Logie, 2003, p. 42). Adding this component allows for the Workspace model to account for LTM effects in WM as well as the behavioral and neurological dissociations found between WM and LTM as well as between verbal and visual memory.

Support for the workspace model comes primarily from the visual domain, particularly research examining representational unilateral spatial neglect in brain-damaged patients and dual-task interference research comparing mental imagery processes (i.e., long-term visual memory) and VWM processes. Patients with representational unilateral spatial neglect, for example, show impairments in one half of their visual mental representation without a similar impairment in visual perception, visual attention, or long-term visual knowledge (Beschlin, Cocchini, Della Sala, & Logie, 1997; Della Sala & Logie, 2002; Logie, Della Sala, Beschlin, & Denis, 2005; van der Meulen, et al., 2009). That is, in imagining a common object or scene, these patients will only draw or report one half of that memory representation, even if that representation was formed prior to any damage, but these patients are still able to make sense of their perceptual world (i.e., follow a familiar route; Beschlin, et al., 1997).

van der Meulen et al. (2009; see also Logie, 2003) argue that this result is incompatible with a gateway model of WM because if WM served as a gateway between perception and LTM (as Baddeley's model does), then these patients with impaired VWM should not have access to LTM representations that allow for understanding of the

perceptual environment (i.e., recognize a cup). Similarly, the unitary model has a difficult time accounting for these data as it proposes that WM is activated LTM representations, thus any impairment in WM should be associated with an impairment in accessing representations in LTM. Thus, these neurological data provide support for a more direct relationship between perception and LTM.

This neuropsychological data is supported by empirical work using dual-task methodology. Primarily this work has examined the influence of irrelevant visual information (in which a series of images are presented during a retention interval) and/or dynamic visual noise (similar to television static) on various type of visual processes. For example, van der Meulen et al. (2009) examined interference effects using irrelevant visual information on both VWM and visual LTM. In the task, participants either had to remember the visual properties of presented letters (i.e., case and order) or they heard the letters and had to generate mental images of those letters and determine if those letters met a particular criteria (i.e., horizontal lines, curves, etc.). In addition to these primary tasks, two interference tasks were used to examine how they impacted performance: irrelevant visual pictures presented during the retention interval or a spatial tapping task in which participants tapped out an hour glass figure on a 3 x 3 grid.

van der Meulen et al. (2009) were testing three main hypotheses about the relationship between WM and LTM based on the WM models outlined above. The gateway model would predict that irrelevant visual information would severely impair the ability to maintain visual information in WM since perceptual input accesses WM directly but it would not have any clear predictions about how irrelevant visual pictures

would influence imagery from LTM or how spatial tapping would interfere with either primary task. The unitary model would predict that more attentionally demanding tasks would cause interference. Thus, pictures may interfere with imagery more than memory, but given the complexity of tapping out a pattern in a non-visible space that task should greatly interfere with all tasks.

Finally, the workspace model would predict specific interference effects. Irrelevant visual information should not interfere with the memory task since retaining visual information in VWM does not necessarily require access to LTM. Irrelevant pictures should cause interference in retrieving information from LTM to generate images, however, since perceiving irrelevant pictures would directly lead to LTM activation, a process that would interfere with activating the relevant LTM representations for the imagery task. Conversely, the opposite pattern would be predicted for spatial tapping: the memory task should be impaired since the maintenance of the tapping task requires temporary memory for the current position as well as planning the next move while the imagery task should be unaffected. Their results supported the workspace model by showing that items in VWM can be maintained “independently of perceptually driven activation of LTM” (van der Meulen, 2009, p. 1575).

Based on the neuropsychological and empirical evidence, Logie (2003) argues that the notion of there being a direct link between information in VWM and our visual perceptions of the world is erroneous. Rather, he argues that perception is directly linked with long-term representations with WM being a structurally separate system that contains that activated information (van der Meulen, et al., 2009).

Thus, as Figure 1 (panel C) suggests, the workspace model is somewhat of a hybrid of the gateway and unitary models of WM, with WM and LTM being separate structures (consistent with the gateway model), but the contents of WM being determined by perceptually activated long-term representations (consistent with the unitary model). This is a key advantage of this model since it can overcome the shortcomings from the Gateway and Unitary Models. However, the Workspace Model does make domain-specific distinctions between verbal and visual leading to similar problems with multiple memory systems as the Gateway Model. Again, however, breaking up WM into these systems allows the model to make clear, specific predictions about what will influence WM performance, particularly how LTM will influence performance.

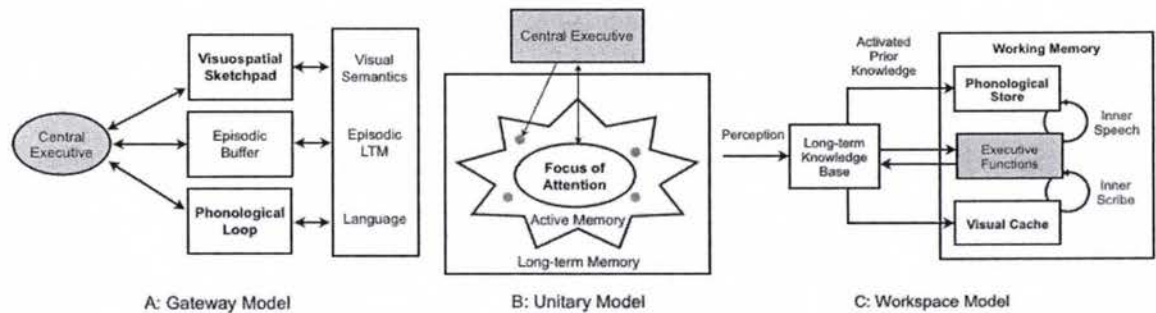


Figure 1. The three WM models discussed in text: (A) the multiple-component model (adapted from Baddeley, 2000); (B) the unitary model, grey dots represent activated portions of LTM that are not in the focus of attention (adapted from Cowan, 1993); (C) the workspace model (adapted from Logie & van der Meulen, 2009). Bolded text indicates component of model in which consolidation is presumed to occur.

## Summary

As the above review highlights, models of WM vary greatly in how they conceptualize WM, particularly the interaction between WM and LTM (see Figure 1). Each of the models has pros and cons in terms of understanding the WM/LTM relationship. However, WM has been studied in relative isolation. That is, in the WM

literature little research has examined the extent to which LTM and WM interact with each other to facilitate cognition. A key focus for the current set of experiments is examining these interactions in the visual domain, particularly the relationship between WM and LTM, as well as how LTM influences early processing in VWM.

Note that while the above review focused on three WM models that exemplify the leading ideas in the WM literature, there are several other approaches to WM that were not discussed. For example, cue based models of WM (for example, the Feature Model; Nairne 1990, 2002) argue that memory is unitary and that short-term and long-term memory processes are cue based. That is, there is no activated information residing in memory, instead Nairne (2002) argues essentially that pointers to information reside in memory that retrieve information based on similarity. However, these models are almost entirely based on verbal memory processes, making it difficult to make clear predictions about visual memory processes.

Additionally, other work in WM has focused on the central executive component of WM and how it controls the content of WM. For example, Engle, Kane, and colleagues (e.g., Engle & Kane, 2004; Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2000; Kane, et al., 2004) have proposed the two-factor theory of WM capacity (i.e., the capacity of the domain general central executive). This theory argues that the central executive is responsible for (1) controlling attention in order to maintain goal relevant information in an active state, and (2) maintaining and retrieving information the face of interference or distraction. More specifically, a key function of the central executive is to resist proactive interference,

which occurs when previously learned information interferes with learning new information (Kane & Engle, 2000). The ability to resist the build up of proactive interference has been shown to be a key function of the central executive, and this ability is highly predictive of higher-order cognition (e.g., fluid intelligence; see Engle et al., 1999).

While these additional models bring up important issues in the WM literature, they do not speak directly to the key questions of the current set of experiments, which is how LTM influences VWM at early processing stages.

### **Properties of Visual Working Memory**

As briefly defined above, VWM encompasses all processes involved in encoding, maintaining, manipulating, and retrieving visual and spatial information. This short-term store is separate from the more transient sensory memory, which only retains information for very brief intervals and is more perceptual in nature (Sperling, 1960). The visual domain provides an excellent medium for examining interactions between WM and LTM because it is easier to find stimuli that have no long-term associations compared to verbal material and because it is easier to rule out other potential confounding cognitive processes such as perceptual processes or strategy use. Additionally, there is currently very little work examining the relationship between long-term visual representations and VWM representations and given that in real-world situations visual representations will usually have some association with long-term memory, it is important to understand this link to better understand the functional importance of VWM.

## **Measuring Visual Working Memory**

Because visual representations cannot be directly measured, they can be difficult to assess. When measuring VWM, researchers “must be careful not to use techniques that alter or destroy the memories as they are being measured and must distinguish between the memory representations and the processes used to create and use them” (Luck, 2008, p. 45). To satisfy these constraints, the vast majority of VWM research uses the change-detection paradigm (Rensink, 2002; Simons, 2000; Simons & Rensink, 2005) in which participants determine whether or not an aspect of a visual array or scene has changed between successive presentations. In the most common version of the change-detection procedure, a brief visual array is presented (called the memory or sample array), followed by a short retention interval. Following the retention interval, a second test array is presented in which participants determine if the test array is the same as the study array. Items in the test array could be replaced by new items (e.g., Luck & Vogel, 1997) or they could vary in spatial location (e.g., Jiang, et al., 2000).

Luck (2008) cites three main advantages for the change-detection task as a measure of VWM. First, the non-memory processes required for the task are relatively minor, ensuring that performance is primarily influenced by the VWM system. For example, verbal memory processes can be easily eliminated by implementing a concurrent verbal task (or articulatory suppression) that requires participants to say a word or syllable aloud during each trial. Similarly, perceptual processes can be controlled for by using simple, highly-discriminable stimuli. Second, the change detection task does not require any transformation of the information in order to report the stored items. In

recall tasks for example, the act of recalling the information could alter the representations of that information (Luck, 2008). The change detection task merely asks participants to compare their memory representation with a presented set of stimuli, thus leaving the representation in tact.

Finally, the change detection procedure is similar to how VWM is used in real-world vision. For example, work by Hollingworth and colleagues (e.g., Hollingworth, 2004; Hollingworth, 2008; Hollingworth & Luck, 2009; Hollingworth, et al., 2008) has shown an important role of VWM in maintaining visual information across eye movements when processing real-world scenes. VWM likely stores information as visual information is disrupted (e.g., blinking) and that representation is then compared to the external world to see if things have changed. By using the more simplified change detection task, the basic processes underlying VWM can be examined while still maintaining relevance to real-world visual processing.

### **Capacity and Representation in Visual Working Memory**

The primary focus of research on VWM thus far has been on capacity and nearly all the research conducted thus far indicates that VWM is severely limited in capacity. Luck and Vogel (1997; see also Vogel, Woodman, & Luck, 2001) took an initial look at capacity in VWM. Using a color change detection task (in which the difference between the study and test arrays was a change in color of one of the items), they examined performance as a function of the set size of the array and the number of features per item. They showed a significant drop in performance with set sizes larger than four, suggesting that VWM can maintain about four objects. Additionally, performance was equivalent

when remembering a single feature of an item and when remembering the conjunction of features. For example, if the stimulus is a smaller square embedded within a larger square of a different color, VWM performance is the same if you have to remember either square independently or if you have to remember the combination of the two squares. Based on this, Luck and Vogel (1997) concluded that items in VWM are stored as integrated objects, not individual features, and they suggested that so long as features were integrated into objects that a seemingly infinite number of features could be stored.

Alvarez and Cavanagh (2004), however, challenged the notion that stimulus complexity (i.e., the number of features per item) did not influence VWM capacity. In their experiment, Alvarez and Cavanagh (2004) calculated capacity of VWM based on change detection performance across a wide range of stimuli that varied in visual complexity. They determined complexity by determining the visual search slope (i.e., the number of ms per item required to locate the target object), with more complex objects requiring longer search times. Alvarez and Cavanagh (2004) showed capacity estimates dropped as stimulus complexity increased. They argued that capacity in VWM is not necessarily fixed (as implied in Luck & Vogel, 1997) but rather capacity can vary to some extent up until an absolute capacity limit of four to five items.

There are some problems with Alvarez and Cavanagh's (2004) study however, and more recent research suggests that the differences between stimulus types found in their study could be due to other factors. For example, Awh, Barton, and Vogel (2007) replicated Alvarez and Cavanagh (2004) and demonstrated that the effects were largely explained by stimulus similarity. That is, change detection performance could be limited

by errors in comparing the memory representation with the test array. Indeed, when Awh and colleagues (2007) varied stimulus complexity but kept similarity low, they showed no capacity differences between the complex objects and the simple objects with the capacity for both at about four items. Thus, the consensus in the VWM literature that the capacity of VWM is around four integrated objects and this limit is relatively stable across varying stimuli.

Why would VWM be so limited in capacity? Vogel et al. (2001; also see Luck, 2008) proposed a neural mechanism that both explains how objects in VWM are bound together and why the limit for these objects is so low. They suggest a synchronized firing mechanism in which neurons that represent various features of an object fire synchronously forming a cell assembly, a view that is largely based on Hebb's (1949) seminal work. Items in memory remain distinct from one another based on asynchronous firing between neurons in different cell assemblies. For example, if a vertical blue bar and a horizontal yellow bar are being stored in VWM, one cell assembly would exist for the blue bar (with neurons tuned to blue colors and vertical orientations) and one would exist for the yellow bar (with neurons tuned to yellow colors and horizontal orientations), and the firing rate within each of these cell assemblies would be different so that they did not overlap. With this type of mechanism, blue and vertical are linked via synchronous firing but that assembly is distinguished from yellow and horizontal via asynchronous firing (Vogel et al., 2001).

While this mechanism allows for binding of features into objects and keeping those objects distinct, it lends itself susceptible to interference as the number of objects

(or cell assemblies) increases. Specifically, as more objects are bound simultaneously, the neurons within the cell assemblies might fire at the same time by chance, a phenomena that Vogel et al. (2001) called accidental synchronization, thereby reducing the distinctiveness of each object. As the number of objects increases, so does the probability of accidental synchronization, suggesting that the reason VWM capacity is limited to a few objects is to ensure that these representations remain distinct and reduce interference between items in WM (Vogel et al., 2001). Additionally, this type of neural mechanism explains why objects in VWM can seemingly have an infinite number of features. That is, the limitation is in the number of cell assemblies and not in the number of neurons within those cell assemblies (Vogel et al., 2001).

Unfortunately, because previous research suggested that VWM was relatively unaffected by stimulus complexity the vast majority of VWM research has used simple, highly distinguishable shapes, limiting the generalizability of VWM research to more complex tasks. Additionally, some more recent research suggests that complexity does impact VWM, putting this capacity consensus in question (e.g., Gao, et al., 2008; Luria, Sessa, Gotler, Jolicoeur, & Dell'Acqua, 2009; Makovski & Jiang, 2008). For example, measuring brain activity using event related potentials during a change detection task Luria et al. (2009) showed greater brain activity with complex shapes compared to simple shapes suggesting that more complex shapes were harder to maintain in VWM. They also showed lower accuracy with complex shapes, going against the claim that VWM is unaffected by stimulus complexity. This research highlights the necessity in generalizing

current ideas on VWM to other types of stimuli to further develop our understanding of how VWM works.

It has been established that the capacity of VWM is about four integrated objects. But how are those objects represented in VWM? Jiang et al. (2000) addressed this issue using a standard change-detection task in which a single item was cued at test. In their task, however, both the presence or absence of the distractor probes (i.e., non-cued items), as well as their identity (i.e., color) and their location was varied during the probe image. The primary question was whether items in VWM are stored independently or whether relational information stored in memory. If items are stored independently, then memory for an item within an array should not be affected by changes in surrounding items. Conversely, if relational information is stored, memory for an item would be affected by surrounding items. Since memory was impaired when the configuration changed between the memory image and the probe image, these findings supported the relational information hypothesis.

Jiang et al. (2000) drew three main conclusions about the nature of organization in VSWM: (1) VSWM is organized based on a global spatial configuration where items are represented in relation to each other (this explains why memory is impaired when this configuration is distorted, or absent); (2) The formation of the configuration is flexible in which both top-down and bottom-up processes work together to select important aspects needed to form the configuration; (3) Information is organized in a hierarchical manner – thus, items in an array can be stored in relation to each other to form a global spatial configuration, but also information about the individual items must be stored. Storing items in a global configuration seems to be a fundamental aspect of VWM processing and

whether or not this configuration can be created will greatly impact memory for a set of items (e.g., Blalock & Clegg, 2010). It is currently unclear what, if any, role LTM plays in the creation of these global configurations.

### **Consolidation in Visual Working Memory**

When perceiving information in the visual world, sensory information leads to perceptual representations in the brain that are fragile. For example, in Sperling's (1960) partial-report experiment, a 3 x 3 letter matrix was presented briefly and then followed by a cue (a high, middle, or low pitched tone) to indicate which row of the matrix should be recalled. Sperling (1960) showed that when the cue was presented shortly after the matrix, recall accuracy was quite high, suggesting that all items in the matrix were stored and accessible since participants did not know ahead of time which row would be cued. When the cue was delayed, however, performance dropped dramatically, with performance bottoming out around 500-800 ms. Additionally, when a mask was presented following the letter matrix, performance was markedly worse than a no mask condition. These results suggest that perceptual information is initially stored in a high-capacity sensory store, but information in this store only remains for a very brief interval before decaying (Gegenfurtner & Sperling, 1993; Sperling, 1960).

Thus, in order to retain visual information over longer periods of time (i.e., longer than a few hundred milliseconds) and to protect it from subsequent interfering visual information (e.g., a mask) some mechanism must stabilize visual representations so that they can persist after the visual information is no longer available in sensory memory (or in the environment) and so that they can remain active when new sensory input is

presented (Luck, 2008). Stated differently, some process is required to take information within sensory memory and transfer it to working memory (Gegenfurtner & Sperling, 1993). This process of stabilizing active WM representations so that they are less susceptible to interference from subsequent visual information is called short-term consolidation (Jolicoeur & Dell'Acqua, 1998; Vogel, Woodman, & Luck, 2006). Consolidation is a loaded term that can have multiple different meanings and in this case may not accurately reflect the process being described, particularly in terms of how it is used in the neuroscience literature (i.e., consolidating disparate pieces of information into a single representation; see Vogel et al., 2006). However, to maintain consistency in terminology with the prior literature the current set of experiments will use the phrase short-term consolidation to refer to stabilizing WM representations.

Research suggests that this consolidation process is very limited in capacity and resource demanding (e.g., Chun & Potter, 1995; Dell'Acqua & Jolicoeur, 2000; Jolicoeur & Dell'Acqua, 1998; Raymond, Shapiro, & Arnell, 1992; Ward, Duncan, & Shapiro, 1996). For example, Jolicoeur and Dell'Acqua (1998) used a dual-task paradigm that combined a VWM task (remembering visually presented characters) and an auditory response task. In this task, a visual array of alphanumeric characters was presented which was followed by a low or high pitched tone. Participants had to make an immediate, speeded response to the tone and then recall the characters. Jolicoeur and Dell'Acqua (1998) showed that when the delay between the visual display and the tone was short (i.e., 300-550 ms), reaction time to respond to the tone was slow, but reaction times got

faster as the delay increased. Additionally, this effect was larger when participants had to remember more characters.

Jolicoeur and Dell'Acqua (1998) argued that the slower response times to the tone following the presentation of the characters is due to interference between consolidating the visual information and performing the auditory task and that this interference was due to both consolidation and response selection requiring central resource mechanisms (Vogel et al., 2006). Additionally, because the interference was larger with larger set sizes of the visual array, Jolicoeur and Dell'Acqua's data suggest that the consolidation process is limited in capacity, requiring more time as the memory demand increases.

**Attentional blink paradigm.** Much of the work on short-term consolidation comes from the attentional blink literature. In a typical attentional blink task, a stream of stimuli are rapidly presented and participants have to respond to two target stimuli within that rapid stream of stimuli (also called a rapid serial visual presentation, or RSVP task). For example, a stream of numbers will be presented rapidly (~10 items per second) and within that stream two letters will be presented and participants then have to report those letters at the end of the trial (e.g., Di Lollo, Kawahara, Ghorashi, & Enns, 2005). In these tasks, accuracy is usually very high for the first target stimulus but is reduced for the second when the second target is presented less than 500 ms after the first target (Raymond, et al., 1992). Accuracy for reporting the second target is greatly improved when instructions indicate to ignore the first target, suggesting that it is not simply a perceptual interference problem but rather an attention problem. Thus, when attention is

focused on processing the first target, the ability to not only report but detect a second target is greatly diminished, showing a temporary lapse, or blink, in attentional resources.

While several theories of why the attentional blink occurs exist (e.g., Di Lollo, et al., 2005), one of the leading theories argues that the drop in performance for the second target is due to a delay in selecting the second target for consolidation while the first target is undergoing consolidation (Chun & Potter, 1995; Nieuwenstein, Chun, van der Lubbe, & Hooge, 2005; Shapiro, Arnell, & Raymond, 1997; Vogel et al., 2006). Thus, the assumption of this view is that consolidating the first target is fairly time consuming (~500 ms) and during that time a second target cannot be consolidated.

**Time course of consolidation.** While the above work supports a distinct consolidation process, the nature of the tasks used (e.g., dual-task and attentional blink paradigms) do not allow for accurate estimates of how much time is required to consolidate information. That is, for both dual-task and attentional blink paradigms, other cognitive processes such as perception, response selection, or task switching are involved in the task. Thus, any time estimates of consolidation based on these tasks will be inflated due to contamination from other processes.

Vogel and colleagues (2006) addressed this problem and isolated the time course of consolidation in VWM. Using a color change detection task, they presented a mask array at variable intervals (stimulus onset asynchronies [SOAs] varying from 117 ms to 583 ms) following the initial memory array. The idea was that the mask array would only impair change detection performance if it was presented while information from the memory array was still being consolidated (i.e., at shorter SOAs). They also varied the

size of the memory array from one to four items to examine the effects of load on consolidation.

They showed that with a set size of one, the mask array did not lead to any drop in performance. However, as the set size increased from two to four, the mask array led to a drop in change detection accuracy at the shorter SOAs and this effect was larger for larger set sizes. This supports previous research arguing that the consolidation process is limited in capacity, requiring more time as the number to-be-consolidated items increases. Additionally, Vogel et al. (2006) determined that approximately 50 ms is required to consolidate an item into VWM. Thus, for a visual array of four items about 200-300 ms is required to create a stable representation in VWM.

Vogel et al.'s (2006) rate of consolidation is notably faster than earlier work using dual-task and attentional blink paradigms, suggesting that that prior research was in fact inflating the time required to consolidate information into WM, likely due to contamination with other cognitive processes. However, a rate of 50 ms per item is slightly slower than other work based on a similar masking paradigm. Using a partial report paradigm (as used by Sperling, 1960) combined with a masking procedure similar to Vogel et al. (2006), Gegenfurtner and Sperling (1993) examined consolidation processes in verbal WM. In their task, a matrix of letters was presented followed by both a cue and a visual mask, both of which were independently manipulated and presented at varying intervals following the offset of the visual letter array. Their data showed a consolidation rate of 20-30 ms per letter.

Vogel et al. (2006) suggest that the discrepancy between their study and Gegenfurtner and Sperling (1993) could be due to differences between visual and verbal WM or due to the fact that the participants in Gegenfurtner and Sperling's (1993) study were highly trained in the task (over 1,000 trials completed prior to the experiment). This over-learning of the stimuli in Gegenfurtner and Sperling (1993) could lead to an increase in consolidation speed. This idea of how familiarity with the stimuli affects consolidation will be a primary focus of the current set of experiments.

How does this consolidation process fit within the context of the WM models outlined above? None of the current WM models directly address consolidation and where it operates, however given that the models discuss in detail the activation of information within the WM store, it is safe to make some assumptions about where the consolidation process would occur. Likely, consolidation occurs in the part of the WM model in which representations are maintained. For the gateway and workspace models, that would be in the domain-specific storage buffers. In the unitary model that would be in the focus of attention.

## **Summary**

The above review highlights three key aspects of VWM: (1) that VWM is very limited in capacity; (2) that items in VWM are stored relative to each other in a global spatial configuration; and (3) that a resource limited consolidation process stabilizes active WM representations. However, despite the recent increase in research focused on understanding short-term visual processing, there are still many unresolved issues. First, what knowledge we do have about VWM processing has been developed using very

simple stimuli despite the fact that recent research suggests that VWM processing varies as a function of stimulus complexity. Second, while research has started to examine how VWM representations are stabilized via the short-term consolidation process, this research has not addressed what is being consolidated or how this process may operate. Finally, the relationship between long-term visual representations and VWM has been largely ignored. The current set of experiments will examine these gaps, in particular the relationship between VWM and LTM

### **Interactions Between VWM & LTM**

Luck (2008) notes that currently no theories of the relationship between visual LTM and VWM exist, specifically regarding the link between object representations in LTM and VWM representations. Essentially, he argues that there are two possible relationships, either long-term visual representations assist VWM processing or that visual long-term representations are not used in VWM.

While Luck (2008) ties these hypotheses to the object perception and representation literature (e.g., Biederman, 1987; Pessig & Tarr, 2007), they map on nicely with the theories of WM outlined above in terms of how they model the relationship between WM and LTM. The unitary and workspace models would predict that visual LTM should facilitate VWM processing since it argues that WM is activated portions of LTM. The Gateway Model, on the other hand, predicts that visual LTM representations should play a reduced role in VWM processing since representations in VWM are based solely on the visual properties of those objects without going through LTM first.

Currently, the small amount of research examining how VWM and visual LTM interact has focused on how familiarity with stimuli influences change detection performance, particularly if familiar objects can increase VWM capacity. A few different techniques have been used to examine this question including explicit training, implicit training, and comparing already known stimuli with novel stimuli. The general hypothesis for all of this research has been that if VWM is activated LTM representations (or if LTM representations are activated prior to information in WM as in the workspace model) that performance and/or capacity should increase. Currently however, the results are mixed.

### **Explicit Training of Visual Stimuli**

Currently, the only experiment that explicitly trained a set of stimuli was Chen, Eng, and Jiang (2006). While other research has examined the impact of training with VWM tasks on general ability (e.g., Buschkuhl, et al., 2008; Wright, Thompson, Ganis, Newcombe, & Kosslyn, 2008), only Chen and colleagues (2006) examined the impact of training stimuli (and thereby creating a long-term representation) on VWM performance.

In their experiment, participants first trained on a subset of eight polygons by completing 320 change detection trials that used only the eight trained shapes. Following a basic recognition task in which participants determined which of two shapes they studied previously (to ensure that the shapes were learned), participants then completed the change detection task again with both trained and novel polygons. In this testing phase, the arrays could consist of either all novel, all trained, or a mixture of novel and

trained shapes. Finally, a second recognition test was administered to examine the retention of the trained shapes following the primary change detection task.

Across several experiments Chen et al. (2006) found no improvements in VWM performance for trained shapes compared to novel shapes. With more practice, participants improved their performance on the change detection task, but this improvement was not specific to the trained shapes. Chen et al. (2006) suggest that familiarity of stimuli may only play a role with extreme degrees of familiarity which is developed over long periods of time (e.g., faces). However, they also note that while training did not impact overall performance, training could have led to improvements more subtle than can be captured by overall accuracy in change detection (e.g., improved encoding). Thus, while Chen et al. (2006) did not show any improvements, it remains an open question if training can facilitate early VWM processes such as consolidation.

### **Implicit Training of Visual Stimuli**

Some work has attempted to examine the link between long-term representations and VWM using more implicit techniques. With implicit learning, information is repeated over the course of the experiment without participants' awareness. Typically, performance is better for this repeated information than on other novel trials, even though participants typically do not realize this and they are unable to articulate any coherent pattern to the trials (e.g., Chun & Jiang, 1998, 1999, 2003).

For example, Olson and Jiang (2004) examined performance over time when either spatial information or both object and spatial information was repeated every third trial. Using a typical change detection paradigm in which one item was cued at test, they

showed no benefit for repeated information across trials for either the spatial or object dimensions. Olson and Jiang (2004) suggest that the resolution of representations in visual LTM is not as high, making it less useful to rely upon than VWM representations. However, more recent research on long-term memory for objects suggests that not only is the capacity for visual LTM very high, the representations are quite detailed (Brady, Konkle, Alvarez, & Olivia, 2008), leaving open the question as to why long-term implicit representations do not improve VWM performance.

Olson, Jiang, and Moore (2005) extended upon Olson and Jiang (2004) by using a similar procedure (repeating spatial information across trials) but comparing repeating locations of non-tested items to repeating locations of tested items. That is, on a repetition trial, the tested item would either have been tested before or would vary randomly between the repeating items. They showed that VWM performance was only enhanced when tested items were repeated. Combined with Olson and Jiang (2004), this research suggests that perhaps one function of VWM is to control what information is retained across trials and information that has been tested before will be retained and learned whereas other information may not.

### **Inherently Familiar versus Unfamiliar Stimuli**

Unlike the work by Chen et al. (2006) who trained on a set of shapes, some visual stimuli already have long-term representations that were formed prior to the experiment and are inherently familiar to participants. For example, Pashler (1988) compared change detection performance between normal alphanumeric characters (which have high degrees of familiarity) and inverted ones (which are less familiar). Similar to Chen et al.

(2006), however, he found no significant differences between the two stimuli, though Pashler (1988) notes that given the extreme degree of familiarity with letters, even inverted letters could be familiar enough to negate any advantages.

More recently, work on face recognition has supported the notion that familiarity with stimuli leads to improvements in VWM. Buttle and Raymond (2003), for example, compared change detection between famous and non-famous faces and found change detection was significantly better when a famous face changed to another famous face than when a non-famous face changed to another non-famous face. This effect was eliminated when the faces were inverted, supporting their argument that super-familiarity of stimuli lead to enhanced VWM performance (see Jackson & Raymond, 2008 for a similar effect). In a similar experiment, Curby and Gauthier (2009) compared change detection performance for upright versus inverted faces as well for cars with car novices and car experts. In their task, they used a backward masking paradigm similar to Vogel et al. (2006) to examine the time course of perceptual expertise. That is, if an object is more familiar (i.e., upright faces or car experts recognizing cars) can that object be accessed and responded to faster? To examine this, they varied how long objects were presented before a mask was shown on the screen (essentially, they manipulated encoding time). They found that performance for upright faces and cars with car experts was significantly faster (33 - 70 ms) compared to inverted faces and cars with car novices. This data support the idea that “experience influences the availability of information early, in processing” (Curby & Gauthier, 2009, p. 1). Note, however, that this expert visual

knowledge did not transfer to other similar stimuli such as inverted faces. This suggests that visual knowledge is very domain specific.

## **Summary**

As the above review highlights, currently the status of the relationship between long-term visual representations and VWM is ambiguous. Lab based training tasks, such as those used by Chen et al. (2006), show no benefit for familiar items whereas more real-world tasks that rely on inherent familiarity show a benefit for familiar items. One reason for this discrepancy could be a lack of sensitive measures of VWM in the lab. That is, examining overall accuracy or capacity may not reveal early benefits of LTM like those shown in Curby and Gauthier (2009).

In trying to better understand WM and LTM and how they interact, a logical place to start would be to examine the impact of stimulus familiarity on early consolidation processes in VWM. Using this approach allows for the strongest test of the different theories of the VWM/LTM interaction: if LTM representations are first accessed prior to any WM processes (or if WM is simply activated portions of LTM), then theoretically we should see facilitation at very early processing stages. The backward masking consolidation paradigm used by Vogel et al. (2006) allows for a clean examination of these ideas and should provide the ideal medium to discern any differences between familiar and unfamiliar stimuli.

## **Current Experiments**

The goal of the current set of experiments is to examine the influence that long-term visual representations have on consolidation in VWM. While prior studies have

examined the impact of long-term representations on overall performance or capacity in VWM with mixed results, the current experiments argue that the influence of long-term representations in VWM may occur at a very early processing stage, allowing visual information to be stabilized in VWM faster and be more resistant to retroactive interference from masks compared to novel visual stimuli.

To address this question, the current experiments will adopt the backward-masking consolidation paradigm used by Vogel et al. (2006) and compare VWM performance across mask delays for familiar and unfamiliar shapes. Experiment 1 will compare unfamiliar, complex polygons with familiar, colored squares. This experiment is designed to examine any possible differences in consolidation in the most extreme case to ensure that a difference does in fact exist. Experiments 2a and 2b will build upon Experiment 1 by eliminating any confounds due to using two types of stimuli by training participants on a subset of abstract polygon shapes and comparing performance on those trained shapes to untrained shapes. For both of these experiments, it is expected that unfamiliar shapes will take longer to consolidate from a transient perceptual trace to a more stable working memory trace.

Experiment 3 will take a slightly different approach to examining differences between familiar (or trained) shapes and unfamiliar shapes. The goal of this experiment is to see if long-term representations can be developed incidentally (i.e., without conscious awareness or intentional effort from the participant) and if so will that long-term knowledge influence consolidation into VWM. Again, if long-term visual representations

influence consolidation into VWM, consolidation should be faster for the trained shapes compared to novel shapes.

Overall, if the predicted pattern of results are found (i.e., that long-term representations influence short-term consolidation) they would support the workspace model of WM (e.g., van der Meulen, et al., 2009) and the unitary models of WM (e.g., Cowan, 2001). These models predict that information from long-term memory is activated directly from perceptual traces and is then transferred to WM for further processing and manipulation (i.e., the workspace model) or becomes the contents of the focus of attention (i.e., unitary model). If, however, no influence from long-term visual representations is found for short-term consolidation processes, the gateway model of WM would be supported (e.g., Baddeley & Hitch, 1974). The gateway model would not predict long-term representations to influence short-term consolidation as any perceptual information has to first go through WM before long-term representations can be activated (van der Meulen, et al., 2009). Thus, the current experiments will further our current understanding of how long-term memory and working memory interact, particularly at early processing stages.

## CHAPTER 2 - EXPERIMENT 1

As stated above, the goal of Experiment 1 is to examine the influence that long-term visual representations have on consolidation in VWM. Are items that have representations in long-term memory consolidated more quickly than novel items with no long-term associations? Experiment 1 compared differences in short-term consolidation between two different types of stimuli: unfamiliar, complex polygons and familiar, colored squares. To examine the consolidation process, visual masks were presented at varying intervals following the memory array as in Vogel et al. (2006). Examining performance across these delay intervals provides insight into the time course of consolidation: early masks should interfere with performance more than later masks if consolidation occurs. Additionally, the set size of the memory array was manipulated to examine how the consolidation process varies as a function of the load of the trial. Vogel et al. (2006) argued that short-term consolidation was a resource limited process: as the number of to-be-remembered items increases, the more time is need to consolidate the array into VWM. Comparing familiar and unfamiliar shapes across set sizes also provides a comparison in terms of which stimulus type is more resource demanding.

While there are a number of differences between the two stimuli types, the goal of Experiment 1 was to test for any differences in the most extreme case possible. The logic being that if there are no differences in this case, then long-term memory (or at least

differences between different types of stimuli) is not likely playing a role during consolidation into VWM. Again, if long-term visual representations can speed consolidation, then we should see faster consolidation for the familiar square stimuli than for the unfamiliar polygons, which would support the workspace and unitary models of WM.

## **Method**

### **Participants and Design**

Thirty-six introductory psychology students participated in this experiment for optional, partial course credit. Participants were required to have normal or corrected vision (including normal color vision). A 2 (stimulus type: squares vs. polygons) x 4 (set size: 1, 2, 3, & 4 items) x 5 (mask delay: 12, 134, 250, 367, & 484 ms) within-subjects design was used.

### **Stimuli**

Both familiar and unfamiliar stimuli (described below) were shown on a computer monitor against a grey background for familiar stimuli and a white background for unfamiliar stimuli. The items were randomly assigned to one of 16 positions within a  $9.80^\circ \times 7.30^\circ$  region in the center of the screen.

The familiar stimuli were identical to those used in Vogel et al. (2006). Seven colored squares (subtending approximately  $0.65^\circ \times 0.65^\circ$  of visual angle), white, red, blue, green, black, violet and yellow, were used in this condition. Depending on the set size condition, one to four items were randomly selected (without replacement) for presentation on each trial. For the familiar stimuli mask, four of the possible colored

squares were abutted to form a 2 x 2 checkerboard pattern. The colors for each mask were randomly selected with replacement and the masks were shown at the locations of each item from the memory array (see Figure 2, panel A).

Attneave shapes (complex polygons; Attneave, 1957; Attneave & Arnoult, 1956) were used for the unfamiliar stimuli (see Figure 1, panel B). Attneave shapes with four to seven sides were used with an approximate visual angle from  $1.0^{\circ}$ - $2.0^{\circ}$ . The shapes were selected randomly from a set of 400 total shapes. The mask for the unfamiliar stimuli consisted of 4 x 4 matrices with half of the cells filled in black and half filled in white (see Figure 2, panel B).

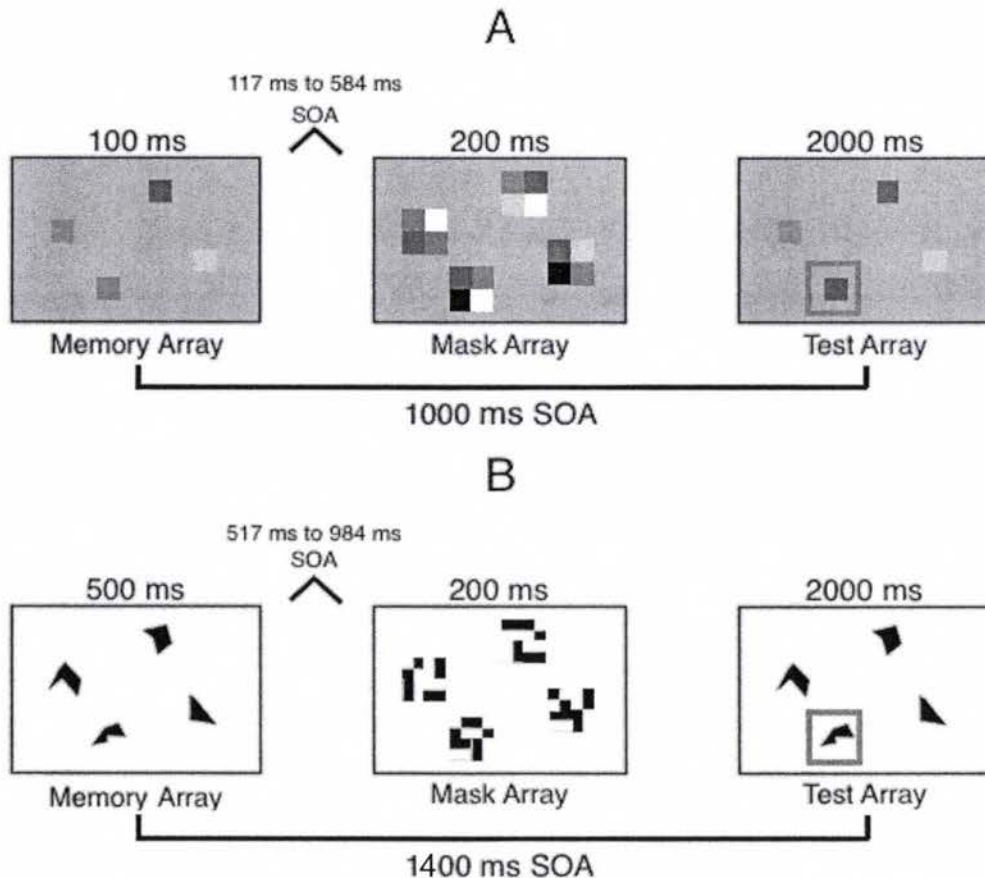


Figure 2. Trial procedures for familiar (A) and unfamiliar trials (B). SOA = Stimulus Onset Asynchrony. Not drawn to scale.

## Procedure

Since the unfamiliar stimuli were inherently more complex than the familiar stimuli, there were some slight differences in timing across the familiar and unfamiliar trials. That is, more study time was allotted to the unfamiliar stimuli to ensure that any differences between the stimuli was not due to difficulty perceiving the shapes in the unfamiliar condition. Familiar and unfamiliar stimuli were presented in separate blocks of 200 trials each (four blocks total in the experiment, for a total of 800 trials, 400 familiar and 400 unfamiliar) and the order of blocks was randomized across participants. This leaves 20 trials per stimulus type x mask delay x set size combination. To avoid any comparison errors in comparing the memory array with the test array for the more homogenous attneave shapes, a single item in the array was cued by a magenta box at test and participants responded if that item changed between study and test. Half of the trials were same trials and half were different trials.

For the familiar trials, the memory array was presented for 100 ms, which was followed by a blank screen until the mask array onset. The delay between the memory and mask arrays randomly varied between 17 ms, 134 ms, 250 ms, 367 ms, or 484 ms on each trial. Thus, the SOAs (adding 100 ms for the memory array) for the familiar stimuli are 117 ms, 234 ms, 350 ms, 467 ms, and 584 ms. The mask array was followed by a blank screen (the duration varied depending on when the mask array is presented) until the test array was presented which remained on the screen until a response was made or 2,000 ms passed. The test array was identical to the study array with the exception that on half of the trials, one item was replaced by a new item (selected randomly from the

remaining colors). Participants responded same (mapped to the “z” key on a standard keyboard) if the two arrays did not differ and different (mapped to the “/” key on a standard keyboard) if they did.

Additionally, in order to rule out any verbal based strategies, participants completed a concurrent articulatory suppression task. For this task, two digits were presented at the start of each trial and participants repeatedly said these two digits aloud throughout the course of each trial. For example, if the digits 2 and 9 were presented, participants said “two-nine, two-nine...” until the end of the trial. This method was used by Vogel et al. (2006) and has been shown to be effective at discouraging any verbal recoding of stimuli (Baddeley, 1986, 2007; Besner, Davies, & Daniels, 1981).

The procedure was similar for unfamiliar trials, with the following exceptions. The array of shapes were presented for 500 ms instead of 100 ms to ensure proper encoding of the more complex stimuli. This makes the SOAs for this condition 517 ms, 634 ms, 750 ms, 867 ms, and 984 ms (500 ms memory array plus the five different delays). A new item also replaced one of the studied items at test on half of the trials and this item was different from the original four shapes but chosen randomly from the remaining set of shapes. No articulatory suppression task was used for the unfamiliar trials.

## Results

Five participants were dropped from the analyses because they either did not finish in the allotted 1.5 hr session ( $n = 2$ ), they opted to not finish due to illness ( $n = 1$ ), they did not follow instructions ( $n = 1$ ), or their performance was two standard deviations

below the average overall performance ( $n = 1$ ). Thus, the following analyses are based on 31 participants. Some accuracy analyses did not meet the sphericity assumption for repeated-measures analysis of variance (ANOVA), so the degrees of freedom for those analyses were adjusted using the Huynh-Feldt correction. Pairwise comparisons were conducted using the least significant differences method. All analyses are significant at the  $p < .05$  level unless otherwise noted.

### **Accuracy**

To examine if the consolidation process differs between the two stimuli types, a 2 (stimulus type: squares vs. polygons) x 4 (set size: 1, 2, 3, & 4 items) x 5 (mask delay: 17, 134, 250, 367, & 484 ms) repeated measures ANOVA was conducted on the accuracy data. This analysis revealed significant main effects of stimulus type [ $F(1, 30) = 52.65$ ,  $MSE = .05$ ,  $\eta_p^2 = .64$ ], set size [ $F(1.59, 47.59) = 114.64$ ,  $MSE = .04$ ,  $\eta_p^2 = .79$ ], and delay [ $F(4, 120) = 11.59$ ,  $MSE = .01$ ,  $\eta_p^2 = .28$ ]. Additionally, interactions were shown between stimulus type and set size [ $F(2.03, 60.85) = 18.44$ ,  $MSE = .01$ ,  $\eta_p^2 = .38$ ], stimulus type and delay [ $F(4, 120) = 3.04$ ,  $MSE = .004$ ,  $\eta_p^2 = .09$ ], set size and delay [ $F(10.09, 302.69) = 2.08$ ,  $MSE = .01$ ,  $\eta_p^2 = .07$ ], and critically the 3-way interaction between stimulus type, set size, and delay was significant [ $F(12, 360) = 2.98$ ,  $MSE = .01$ ,  $\eta_p^2 = .09$ ]. These results, particularly the significant 3-way interaction, support the hypotheses for Experiment 1 and suggest that consolidation differed between the two stimulus types (see Figure 3). Specifically, visual masks disrupted VWM performance at different time intervals across the two stimulus types and this disruption varied depending on the set

size of the visual array, supporting Vogel et al.'s (2006) conclusion that consolidation is resource limited.

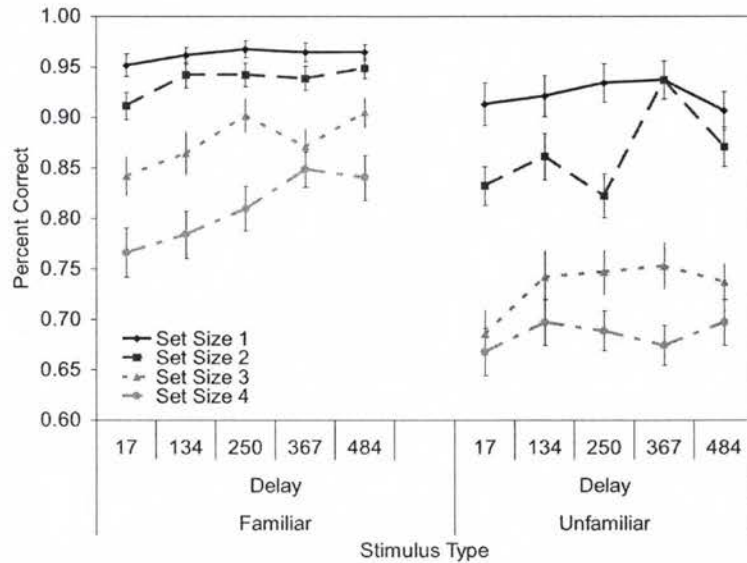


Figure 3. Accuracy data for Experiment 1. Error bars represent one standard error of the mean.

To further examine the main effects of stimulus type, set size, and delay, pairwise comparisons were calculated. For stimulus type, these comparisons showed that accuracy was higher for familiar than unfamiliar trials, suggesting that the unfamiliar shapes were more difficult to remember than the familiar stimuli. For set size, accuracy for set size 1 was significantly higher than all other set sizes, accuracy for set size 2 was significantly higher than set size 3, accuracy for set size 4 was significantly lower than all other set sizes. This pattern replicates Vogel et al. (2006) and suggests that as the number of to-be-remembered items increased, task difficulty increased. The significant set size x delay interaction also supports Vogel et al.'s (2006) conclusion that the consolidation process is resource limited, with more time needed as more items need to be consolidated.

To examine the main effect of delay, performance at the shortest and longest delays were compared first to see if accuracy increased across the delay conditions before the more specific pairwise comparisons were conducted between all the conditions. A paired-samples t-test between the 17 ms delay condition ( $M = .82$ ;  $SD = .07$ ) and the 484 ms delay condition ( $M = .86$ ;  $SD = .05$ ) showed that accuracy significantly increased with additional time between the memory array and the mask array [ $t(30) = -5.41$ ]. This suggests that early masks interfered more with remembering the shapes than later masks, and the interactions with set size and stimulus type suggest that the magnitude of this effect varied depending on how many items were in the array and how familiar the items were. Additional pairwise comparisons showed that not only was accuracy in the 17 ms delay condition significantly lower than the 484 ms condition, it was also lower than all other delay conditions. Accuracy in the 367 ms condition was significantly higher than accuracy in the 134 ms and 250 ms conditions but was equivalent to the 484 condition. This suggests that up to 367 ms were needed to consolidate information into VWM, though this process varied depending on set size, with less time required for a single item and more time required for four items, and by stimulus type with more time needed to consolidate unfamiliar stimuli.

To further examine the 3-way interaction, separate 4 (set size) x 5 (delay) repeated measures ANOVAs were conducted for each stimulus type. The familiar analysis yielded a significant set size x delay interaction [ $F(8.72, 261.62) = 2.08$ ,  $MSE = .01$ ,  $\eta_p^2 = .07$ ]<sup>1</sup>. This result replicates Vogel et al. (2006) and suggests that a certain amount of time is

---

<sup>1</sup> This analysis also yielded significant main effects of set size [ $F(1.74, 52.22) = 68.39$ ,  $MSE = .02$ ,  $\eta_p^2 = .70$ ] and delay [ $F(4, 120) = 10.78$ ,  $MSE = .004$ ,  $\eta_p^2 = .26$ ].

needed to consolidate information into VWM and this process takes longer with more visual stimuli. This interaction was examined by comparing the difference in accuracy between the shortest and longest delay intervals (17 ms and 484 ms) for Set Sizes 1 and 4 using paired-samples t-tests (using a Bonferroni correction to adjust the experimentwise error). This provides the strongest test for how the effects of delay vary depending on the number of items. For Set Size 1, there was no difference in accuracy between the 17 ms ( $M = .95$ ,  $SD = .06$ ) and 484 ms ( $M = .97$ ,  $SD = .04$ ) delay conditions [ $t(30) = -1.05$ ,  $p = .30$ ]. In contrast, there was a significant difference at Set Size 4 between the 17 ms ( $M = .77$ ,  $SD = .13$ ) and 484 ms ( $M = .84$ ,  $SD = .12$ ) delay conditions [ $t(30) = -3.49$ ]. This pattern demonstrates that at larger set sizes, more time was required to create a stable VWM representation as the number of items in the memory array increased (see Figure 5).

To get a sense of the time course of the consolidation process (i.e., a rough estimate of when consolidation occurred), pairwise comparisons were conducted on the delay main effect. These comparisons showed that accuracy after a 17 ms mask delay was significantly lower than all other delay conditions and accuracy following a 134 ms mask delay was significantly lower than delays of 250 ms, 367 ms, and 484 ms (see Figure 4). This pattern suggests that after 134 ms, performance was roughly equivalent, but that varied depending on how many items there were in the memory array (i.e., performance was equivalent across all delay conditions with a set size of 1, but differences across delay conditions started with set size of 2).

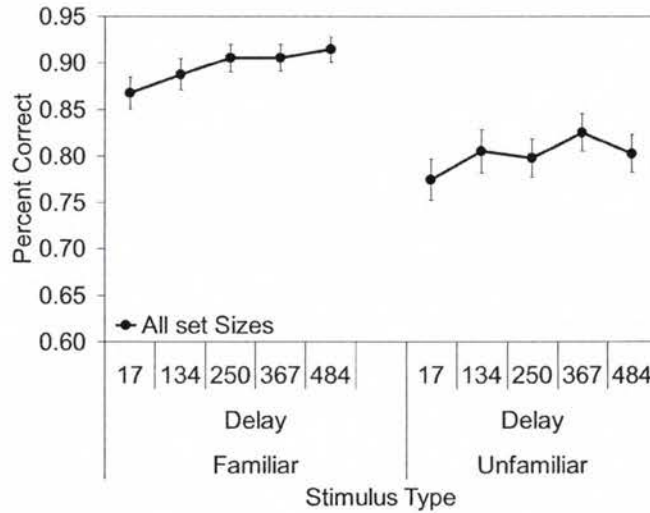


Figure 4. Experiment 1 accuracy across the five delay conditions collapsed across set size conditions for familiar and unfamiliar stimuli.

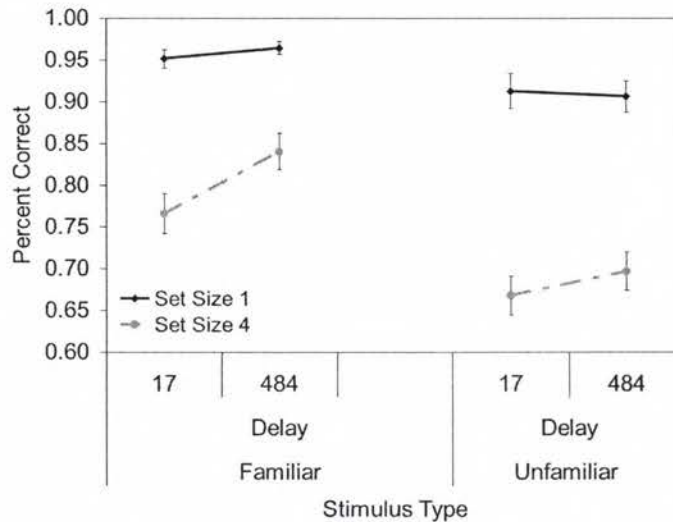


Figure 5. Experiment 1 accuracy for the shortest and longest delay conditions (17 ms and 484 ms) with Set Size 1 and Set Size 4 for familiar and unfamiliar items.

The unfamiliar analysis also yielded a significant set size x delay interaction [ $F(9.78, 293.33) = 2.75, MSE = .01, \eta_p^2 = .08$ ]<sup>2</sup>. This generally replicates the effects seen for familiar items, and suggests that unfamiliar shapes are also sensitive to visual masks when they are shown too soon after the memory array and this effect varies depending on set size. Paralleling the Familiar analysis, paired-samples t-tests (using the Bonferroni

<sup>2</sup> This analysis also yielded significant main effects for set size [ $F(1.71, 51.32) = 92.91, MSE = .04, \eta_p^2 = .76$ ] and delay [ $F(9.78, 293.33) = 2.75, MSE = .01, \eta_p^2 = .08$ ].

correction) were conducted between the 17 ms and 484 ms delay conditions for Set Sizes 1 and 4. For Set Size 1, there was no difference between the 17 ms ( $M = .91, SD = .12$ ) and the 484 ms ( $M = .91, SD = .10$ ) delay conditions [ $t(30) = .43, p = .67$ ] suggesting that with one item masks did not interfere with accuracy in any delay condition. In contrast to the Familiar analysis, at Set Size 4 there was also no significant difference between the 17 ms ( $M = .67, SD = .13$ ) and 484 ms ( $M = .70, SD = .13$ ) delay conditions [ $t(30) = -1.09, p = .28$ ]. This lack of a significant increase is likely due to the fact that accuracy drops in the 484 ms condition compared to the 367 ms condition. Thus, the lack of a difference between 17 ms and 484 ms may be due to the fact that information had decayed at 484 ms.

The pattern of results across the delay conditions was also examined for Unfamiliar items to get a sense of the time course of the consolidation process. Pairwise comparisons showed that performance following a 17 ms delay was significantly lower than all other delay conditions and performance following a 367 ms delay was significantly higher than following a 250 ms or a 484 ms delay and marginally higher than a 134 ms delay ( $p = .06$ ). This pattern suggests that backward masking interfered with consolidation up to 250 ms, which is longer than the 134 ms required for familiar stimuli. Again, the drop in performance at a 484 ms delay suggests that the consolidation process may be susceptible to decay with more complex stimuli.

Taken together, the accuracy analyses support the hypotheses of Experiment 1: short-term consolidation varied between familiar and unfamiliar objects, with the more

complex, unfamiliar polygons requiring more time to consolidate than simple, familiar squares.

### **Capacity**

The number of items,  $K$ , that are available in WM for a change detection task was calculated to determine if the capacity estimate varied depending on stimulus type (as used by Vogel et al., 2006; see also Cowan, 2001). Essentially, the equation “assumes that if an observer can hold  $K$  items in memory from an array of  $S$  items, then the item that changed should be one of the items being held in memory on  $K/S$  trials, leading to correct performance on  $K/S$  of the trials on which an item changed” (Vogel et al., 2006, p. 1441). Accounting for guessing (i.e., accounts for the false alarm rate,  $F$ , and the correct rejection rate,  $C$ ), the available items in WM is  $K = C * [S * (H - F) / (1 - F)]$  or  $K = S * (H + C - 1)$ . This estimate was calculated for each stimulus type condition at set size four. Given the current debate in the VWM literature about how stimulus complexity impacts capacity, this estimate was compared between Familiar and Unfamiliar stimuli in a paired-samples t-test. This test showed that capacity for Familiar items ( $M = 2.49$ ;  $SD = .73$ ) was significantly higher than capacity for Unfamiliar items [ $M = 1.51$ ;  $SD = .65$ ;  $t(30) = 7.27$ ]. This demonstrates that participants were not able to hold as many Unfamiliar items in VWM as Familiar items, suggesting that capacity of VWM varies depending on the stimuli and supporting criticisms by Alvarez and Cavanagh (2004) that a fixed capacity is overly simplistic.

Before examining the slope, paired-samples t-tests were conducted between the 17 ms and 484 ms delay conditions at Set Size 4 for both Familiar and Unfamiliar items.

Conducting this test determines if there are differences in accuracy between the shortest and longest delays. If no differences in accuracy exist, then plotting capacity against delay is not informative. For Familiar items, this test showed a significant increase in capacity between the 17 ms ( $M = 2.16$ ;  $SD = 1.08$ ) and 484 ms ( $M = 2.78$ ;  $SD = .84$ ) conditions [ $t(30) = -3.63$ ]. However, there was no difference between 17 ms ( $M = 1.36$ ;  $SD = 1.05$ ) and 484 ms ( $M = 1.70$ ;  $SD = .96$ ) delay conditions for Unfamiliar items [ $t(30) = -1.58$ ,  $p = .12$ ].

### Discussion

Can long-term visual representations facilitate consolidation into VWM? The goal for Experiment 1 was to answer this question, specifically to examine how the process of short-term consolidation differed between two different types of stimuli that varied in terms of how familiar they were (i.e., colors vs. abstract shapes). It was expected that differences in consolidation would be found, and the three-way interaction between stimulus type, set size, and delay supports that prediction. More specifically, consolidation was slower for the unfamiliar polygons than for the more familiar colored squares, particularly as set size increases.

These results have two key implications. First, they provide preliminary evidence that long-term representations play a key role at very early information processing stages by facilitating the consolidation of representations within VWM. This is largely inconsistent with the gateway model of WM, and provides strong support for WM models that argue for a closer relationship between LTM and WM (e.g., the Workspace and Unitary models). Specifically, the current experiment suggests that visual information

can interact directly with LTM at very early processing stages. It would be difficult for the gateway model, where long-term memory does not influence processing until information reaches WM, to account for the early effects seen here.

Second, in demonstrating differences between different types of stimuli, the current experiment provides a crucial extension of Vogel et al. (2006) and the results suggest that many of the current frameworks and models of VWM processing that are based primarily on data using highly discriminable stimuli may not generalize to other, more complex stimuli. This has important implications not just for the theoretical understanding of how visual information is processed over the short term, but also for how this understanding is applied to real-world situations. If effects found with simple squares do not generalize to more complex, but still relatively simple shapes, then how will they apply to complex, dynamic environments? These data suggest that the current frameworks for VWM are incomplete in accounting for more complex visual processing.

Additionally, previous research has indicated that the capacity of VWM is about four items regardless of complexity (e.g., Awh et al., 2007). The current experiment, however, not only showed lower overall capacity estimates for the unfamiliar polygons, it also suggests that complex items require more time to get into VWM. Since individual items were cued at test, the current experiment calls into question Awh et al.'s (2007) fixed capacity conclusion. That is, Awh et al. (2007) argued that prior research showing lower capacity estimates with more complex shapes was due to comparison errors in comparing the test array with the initial memory array stored in VWM. When individual items are cued, this is no longer a problem because participants only have to make a

single comparison regardless of the size of the array. Thus, the current experiment suggests that more complex shapes require more time to get into VWM and overall people can hold fewer of those items in VWM.

There are a couple of alternative explanations to consider as to why there are differences in consolidation between familiar and unfamiliar shapes. First, during the familiar trials, participants completed an articulatory suppression task to prevent any verbal recoding of the visual stimuli. While this is the same method that Vogel et al. (2006) used, this method could have imposed an additional WM load on the familiar shapes, which may have led to the significant set size x delay interaction found. That is, performing the task under load could impact consolidation into VWM, particularly when the VWM task was more difficult (i.e., at set size 4). Work by Moray and Cowan (2005) provides support for this possibility. In their experiments, they showed interference between visual and verbal tasks when the visual task was difficult and when the verbal task required overt rehearsal (i.e., saying the numbers out loud) of seven random digits. It is possible that a similar interference effect is at work here. If so, it would suggest a role for general attentional resources (i.e., the central executive) in the consolidation process. Future work will have to assess the impact that the type of suppression task has on consolidation.

A second alternative explanation to the current set of results is that there are differences in the level of interference build up across the experiment between the familiar and unfamiliar stimuli. That is, the same seven colored squares were used for the familiar trials whereas different shapes were used for the unfamiliar trials, though the

unfamiliar shapes were more similar to each other compared to the familiar shapes which were highly discriminable. It is possible that the build-up of proactive interference (in which previously learned information interferes with learning new information) varied between the stimuli, leading to different effects of the visual masks. Differences in the level of proactive interference could also lead to differences in the WM load of the task between the stimulus types, leading to differences in consolidation. Work by Engle and colleagues (e.g., Engle & Kane, 2004; Kane, et al., 2001; Kane & Engle, 2000; Kane, et al., 2004) has demonstrated that the level of proactive interference plays a key role in WM tasks, particularly VWM tasks. Again, a key follow-up to the current experiment would be to examine how WM load impacts the consolidation process.

While the significant differences between the stimuli make it difficult to make strong conclusions about the nature of long-term memory in VWM consolidation, the current experiment provides preliminary evidence for an early role of long-term memory in stabilizing active representations in WM. Additionally, the current experiment challenges some of the current ideas about VWM, particularly how VWM is impacted by the complexity of the shapes. The backward-masking consolidation paradigm provides a fine-grained analysis of VWM performance across specific intervals and allows for insight about very early processing of visual stimuli into VWM.

## CHAPTER 3 - EXPERIMENT 2A

Experiment 1 showed preliminary evidence of an early role for visual long-term representations in facilitating short-term consolidation of information into VWM: familiar items were consolidated faster into VWM than unfamiliar items. So far, the data support Logie's workspace model of WM, because they suggest that there is a more direct relationship between perception and LTM. However, difficulties in comparing two types of stimuli leave open the key question for this set of experiments: does visual LTM influence early processing stages by facilitating short-term consolidation?

The goal for Experiments 2a and 2b is to build upon results from Experiment 1 and more directly examine the influence of long-term visual representations on short-term consolidation by training a subset of stimuli and comparing performance between trained and novel stimuli. The experiments will use two different training methods. Experiment 2a will train using a four alternative forced choice recognition paradigm and 2b will train using a more standard change detection paradigm. Again, the backward masking paradigm will be used that varies mask delay (to examine how performance varies based on how soon the masks interfere with consolidation) and set size (to examine how the process varies depending on the load of the task).

By using this cleaner methodology, it is easier to make strong conclusions regarding the relationship between long-term visual representations and consolidation in

VWM. Additionally, Experiments 2a and 2b may help to explain why there have been mixed results in the literature (e.g., Curby & Gauthier, 2009; Buttle & Raymond, 2003; Chen et al., 2006) as to the effects of LTM on VWM performance. That is, work that has explicitly trained a set of stimuli and examined VWM performance has not shown any performance improvements for trained over novel shapes (Chen et al., 2006), but work comparing inherently familiar visual stimuli (e.g., famous faces) to unfamiliar visual stimuli (e.g., novel faces) has shown better VWM performance for familiar items. The current experiment attempts to find differences between explicitly trained shapes and novel shapes, while Experiment 1 was more of a comparison between inherently familiar stimuli (i.e., colored squares) and unfamiliar stimuli (i.e., random polygons).

There are two key reasons why prior work on explicit training of stimuli has not shown effects. First, it may be that prior work has not used precise enough measures of VWM to see any benefits of training, leading to null results in some cases. The backward-masking paradigm gets around this issue by comparing performance levels across various time frames and set sizes, leading to more specific insights about processing in VWM. Second, currently work in this area has not contrasted different training methods in the same task with most studies using their own variant of a training procedure. Experiments 2a and 2b should provide insights as to how critical the type of training is to VWM performance and if some training tasks are better suited than others.

If long-term representations influence short-term consolidation, then we should see a similar pattern as Experiment 1. More specifically, assuming that having long-term visual representations improves consolidation, we should see faster consolidation for

trained shapes than for novel shapes and this consolidation pattern should vary depending on the set size of the array.

## **Method**

Except where noted, the methodology for Experiment 2a closely followed the methods used in Experiment 1.

### **Participants and Design**

One-hundred two introductory psychology students participated for optional, partial course credit. A 2 (stimulus type: trained vs. novel) x 4 (set size: 1, 2, 3, & 4 items) x 5 (mask delay: 12, 134, 250, 367, & 484 ms) within-subjects design will be used.

### **Stimuli**

The stimuli used in the unfamiliar condition for Experiment 1 were used in Experiment 2a.

### **Procedure**

Experiment 2a will involve three main tasks which will be completed in a single session lasting approximately one and a half hours: a training task, a recognition task, and the consolidation task.

**Training.** A subset of 12 polygons were randomly selected for training. Participants performed a four alternative forced choice task (4AFC) during the training session. In this task, one of the 12 polygons were presented for 1 s on the screen followed by a mask presented at 484 ms following the offset of the shape. Participants then selected the studied shape out of four possible shapes. The distractor shapes consisted of three other shapes selected at random from the subset of 12 selected for training. Each of

the four shapes presented at test were labeled from 1 to 4 from left to right on the screen. Participants responded by pressing 1, 2, 3, or 4 on the keyboard, which ever number corresponded to the correct item they studied. Participants received visual feedback on their performance during the training portion of the experiment only. Participants completed 240 training trials (20 4AFC trials for each shape).

**Intermediate Recognition Test.** Following the training task, a two alternative forced choice (2AFC) recognition test was administered to ensure that participants created long-term representations of the stimuli. For this task, two polygons were presented on the screen (labeled A and B), and participants had to select which of the two shapes they studied previously in the experiment. The twelve distractor shapes used in this task were not be used in any other portion of the experiment.

**Consolidation Task.** The consolidation task was the same as in Experiment 1 except that trained and novel shapes were compared instead of squares and polygons. For trained trials, all the shapes used in that trial will be from the subset of 12 trained polygons. The novel trials will be randomly selected from a set of 376 completely new shapes that have not been presented in any other stage of the study.

**Final Recognition Task.** To ensure that participants still have long-term representations of the trained shapes, another 2AFC recognition task was administered following the consolidation task in which participants again were asked to identify the shape that they studied previously in the experiment.

## Results

Of the 102 participants, 10 were dropped due to low performance (overall accuracy 2 standard deviations below the mean), 4 were dropped for not following directions (e.g., randomly pressing buttons, texting during experiment, etc.), and 17 were dropped because they did not finish in the allotted 1.5 hour timeframe. This resulted in 71 participants for the analyses. Some accuracy analyses that included the set size variable did not meet the sphericity assumption for repeated-measures analysis of variance (ANOVA), so the degrees of freedom for those analyses were adjusted using the Huynh-Feldt correction. Pairwise comparisons were conducted using the least significant differences method. All analyses are significant at the  $p < .05$  level unless otherwise noted.

### 4AFC & Recognition Task Performance

Accuracy for the 4AFC task was high ( $M = .96$ ,  $SD = .02$ ). When broken down into four 60 trial blocks, performance remained constant (Block 1:  $M = .96$ ,  $SD = .03$ ; Block 2:  $M = .97$ ,  $SD = .03$ ; Block 3:  $M = .96$ ,  $SD = .03$ ; Block 4:  $M = .96$ ,  $SD = .04$ ). A one-way repeated measures ANOVA conducted on the 4AFC performance data across blocks showed no significant difference over time [ $F < 1$ ]. This suggests that the task was relatively easy, and learning occurred very quickly (i.e., within a single block).

Overall recognition accuracy was moderately high for both the intermediate recognition test [ $M = .85$ ,  $SD = .31$ ] and the final recognition test [ $M = .83$ ,  $SD = .28$ ]. Both of these are significantly above chance [Intermediate:  $t(70) = 9.64$ ; Final:  $t(50) = 8.60$ ; note there is some missing data for the second recognition test due to experimenter

error in misnaming files]. A paired-samples t-test showed no significant difference in performance between the intermediate recognition test and the final recognition test [ $t(50) = .08$ ]. This suggests that participants were able to form long-term representations of the shapes and these representations were robust since there was no significant decline across the experiment.

### Accuracy

Because recognition accuracy was significantly above chance, the accuracy analysis was not conditionalized based on recognition performance<sup>3</sup>. To examine differences in consolidation between trained and novel stimuli, a 2 (stimulus type: familiar vs. unfamiliar) x 4 (set size: 1, 2, 3, & 4 items) x 5 (mask delay: 17, 134, 250, 367, & 484 ms) repeated measures ANOVA was conducted. This analysis revealed significant main effects for stimulus type [ $F(1, 70) = 4.20, MSE = .03, \eta_p^2 = .06$ ], set size [ $F(2.45, 171.18) = 715.80, MSE = .01, \eta_p^2 = .91$ ], and delay [ $F(4, 280) = 8.19, MSE = .01, \eta_p^2 = .11$ ]. Additionally, there was a significant stimulus type x delay interaction [ $F(4, 280) = 3.38, MSE = .01, \eta_p^2 = .05$ ], a marginally significant set size x delay interaction [ $F(11.02, 771.13) = 1.68, p = .07, MSE = .01, \eta_p^2 = .02$ ], and a significant stimulus type x set size x delay interaction [ $F(10.44, 730.72) = 1.95, MSE = .01, \eta_p^2 = .03$ ]. The stimulus type x set size interaction did not reach significance [ $F(2.61, 182.33) = 1.28, p = .28, MSE = .01, \eta_p^2 = .02$ ]. These results suggest that consolidation differed between trained and novel shapes. Similar to Experiment 1, visual masks disrupted VWM performance at

---

<sup>3</sup> A conditionalized analysis was conducted (recognition accuracy had to be greater than 90%) and did not differ significantly from the unconditionalized analysis, so only the unconditionalized analysis is included.

different time intervals across the two stimulus types and this disruption varied depending on the set size of the visual array (see Figure 6).

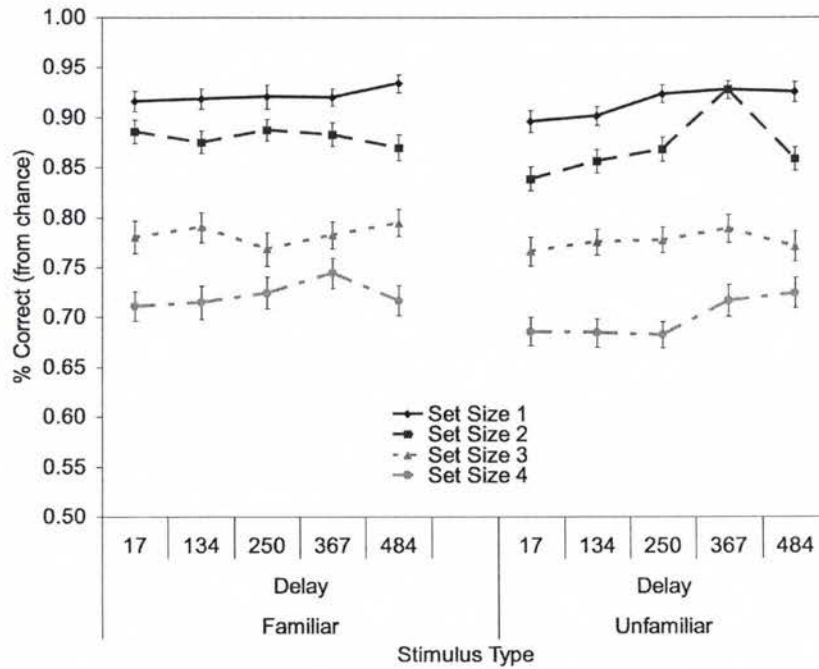


Figure 6. Accuracy data for Experiment 2a. Error bars represent one standard error of the mean.

Pairwise comparisons were conducted to further examine the main effects for stimulus type, set size, and delay. For stimulus type, the comparisons showed that accuracy was higher for familiar than unfamiliar stimuli. This positive result of training is unique compared to prior work (e.g., Chen et al., 2006), suggesting that long-term representations of visual stimuli can be created over a short period of time and used more efficiently than completely novel shapes. In terms of set size, accuracy for set size 1 was significantly higher than all other set sizes, accuracy for set size 2 was significantly higher than set size 3, accuracy for set and size 4 was significantly lower than all other set sizes. This effect mirrors the pattern seen in Experiment 1.

Paralleling the analyses in Experiment 1, to examine the main effect for delay, first performance at the shortest and longest delay conditions were compared to see if accuracy increased across the delay conditions before the more specific pairwise comparisons were conducted between all delay conditions. A paired-samples t-test between the 17 ms delay condition ( $M = .81$ ;  $SD = .07$ ) and the 484 ms condition ( $M = .82$ ;  $SD = .07$ ) showed that accuracy significantly increased with additional time between the memory array and the mask array [ $t(70) = -2.97$ ]. This suggests that early masks interfered more with remembering the shapes than later masks, and the interactions with set size and stimulus type suggest that the magnitude of this effect varied depending on how many items were in the array and how familiar the items were. Additional pairwise comparisons showed that performance following 17 ms and 134 ms delays were significantly lower than 367 ms and 484 ms delays. Performance following a 250 ms delay was significantly less than a 367 ms delay. Performance after a 367 ms delay was significantly higher than performance following all other delays. Interestingly, this suggests that performance was lower at 484 ms compared to 367 ms, an effect that will be addressed in the discussion.

To further examine the 3-way interaction, separate 4 (set size) x 5 (delay) repeated measures ANOVAs were conducted for each stimulus type. The familiar analysis did not yield a significant set size x delay interaction [ $F(10.34, 723.56) = 1.06, p = .39, MSE = .01, \eta_p^2 = .02$ ]<sup>4</sup>. This result is in contrast to the results from Experiment 1 and Vogel et al. (2006) and suggests that trained items were not susceptible to interference from the visual

---

<sup>4</sup> There was also a significant main effect for set size [ $F(2.69, 188.14) = 315.40, MSE = .01, \eta_p^2 = .82$ ]. The main effect for delay did not reach significance [ $F < 1$ ].

masks at any point indicating a very early role for long-term memory in short-term consolidation. This will be revisited in the discussion. To mirror the analyses conducted below for the unfamiliar items, accuracy between the shortest and longest delay intervals for Set Size 1 and Set Size 4 were compared using a paired-samples t-test (using a Bonferroni correction to adjust the experimentwise error, resulting in a  $p < .025$  criterion). This provides the strongest test for how set size impacts consolidation. For Set Size 1, there was no difference in accuracy between the 17 ms ( $M = .92$ ;  $SD = .08$ ) and 484 ms ( $M = .93$ ;  $SD = .08$ ) delay conditions [ $t(70) = -1.99$ ,  $p = .05$ ]. Similarly, there was no difference between the 17 ms ( $M = .71$ ;  $SD = .12$ ) and 484 ms ( $M = .72$ ;  $SD = .13$ ) delay conditions for set size 4 [ $t(70) = -.31$ ,  $p = .76$ ]. This pattern suggests that participants did not require more time to stabilize the VWM representations as the number of items increased. No more pairwise comparisons were conducted on delay since there was no effect of delay for Familiar items (see Figure 7).

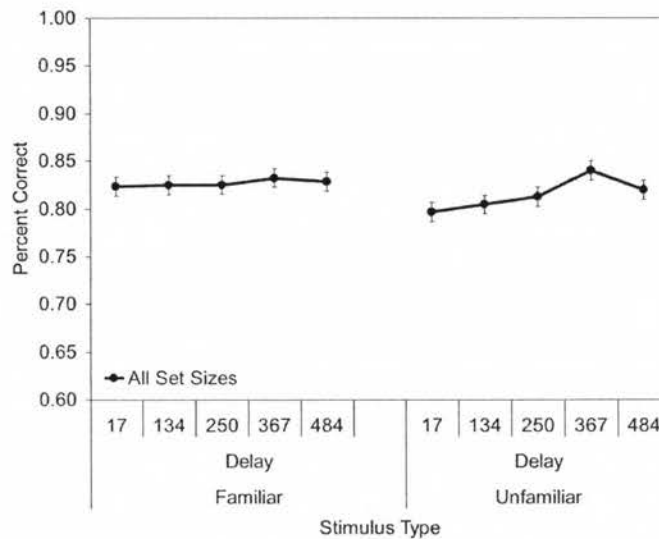


Figure 7. Experiment 2a accuracy across the five delay conditions collapsed across set size conditions for familiar and unfamiliar stimuli. Error bars represent one standard error of the mean.

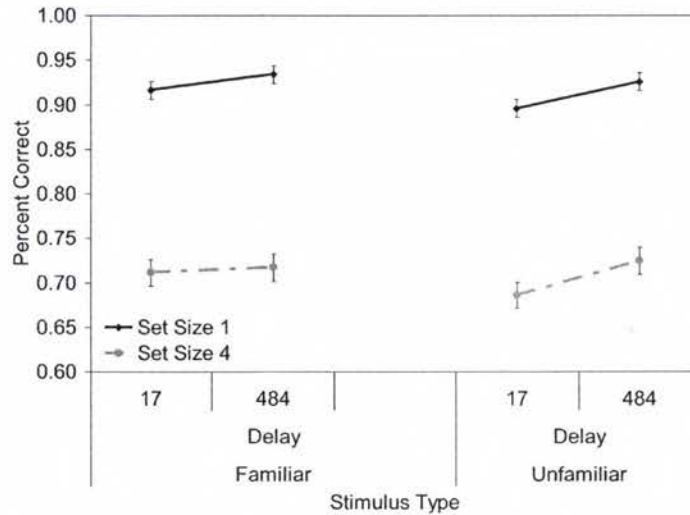


Figure 8. Experiment 2a accuracy for the shortest and longest delay conditions (17 ms and 484 ms) with Set Size 1 and Set Size 4 for familiar and unfamiliar items. Error bars represent one standard error of the mean.

The unfamiliar analysis yielded a significant set size x delay interaction [ $F(10.82, 757.67) = 2.63, MSE = .01, \eta_p^2 = .04$ ]<sup>5</sup> suggesting that unfamiliar items are sensitive to visual masks when they are shown too soon following the memory array and that effect varies depending on how many items are in the array. To further examine this interaction, paired-samples t-tests (using the Bonferroni correction to adjust the criterion to  $p < .025$ ) were conducted to compare performance between the 17 ms and 484 ms conditions for set sizes 1 and 4. For Set Size 1, there was a significant increase in accuracy between the 17 ms ( $M = .90; SD = .09$ ) and 484 ms ( $M = .93; SD = .08$ ) delay conditions [ $t(70) = -2.77$ ]. Similarly, there was a significant increase in accuracy between the 17 ms ( $M = .69; SD = .12$ ) and 484 ms ( $M = .73; SD = .13$ ) delay conditions for set size 4 [ $t(70) = -2.37$ ]. Given that the difference between the 484 ms delay condition and the 17 ms delay

<sup>5</sup> This analysis also yielded a significant main effect for set size [ $F(2.69, 188.55) = 421.63, MSE = .01, \eta_p^2 = .86$ ] and delay [ $F(4, 280) = 10.61, MSE = .01, \eta_p^2 = .13$ ]

condition was larger at Set Size 4, this suggests that masks led to more interference with more items in the memory array (see Figure 8).

Additionally, pairwise comparisons for the delay effect showed that masks interfered with performance up to 367 ms: performance following a 17 ms delay was significantly lower than performance following 250 ms, 367 ms, and 484 ms delays; performance following a 134 ms delay was significantly lower than performance following a 367 ms delay or a 484 ms delay; and performance following a 367 ms delay was significantly higher than all other delay conditions (see Figure 7). This pattern is clearly different from the familiar items since there was no evidence for mask interference with Familiar items, but these data mirror the pattern found in Experiment 1 for unfamiliar shapes. Additionally, there is a drop in performance at a 484 ms delay similar to the pattern found in Experiment 1, again suggesting that the consolidation process may be susceptible to decay with unfamiliar stimuli.

### **Capacity**

The number of items,  $K$ , that are available in WM for a change detection task was also calculated to compare capacity between familiar and unfamiliar items as in Experiment 1. This estimate was calculated for both stimulus type conditions at set size four across all delay conditions. This provides a comparison as to how capacity varies depending on how familiar the items are when complexity is equivalent between the familiar and unfamiliar conditions. A paired-samples t-test between familiar stimuli ( $M = 1.79$ ;  $SD = .71$ ) and unfamiliar stimuli ( $M = 1.60$ ;  $SD = .69$ ) showed that capacity was significantly higher for trained, familiar shapes than for novel, unfamiliar shapes [ $t(70) =$

2.41]. Thus, even with an equal level of stimulus complexity, training participants on a set of complex polygons increases the capacity of VWM for those items compared to novel items.

### **Discussion**

Experiment 2a more directly examined the impact that visual long-term representations have on short-term consolidation into VWM compared to Experiment 1. The results provide strong evidence that visual long-term representations facilitate the consolidation process: familiar items were unaffected by visual masks even when those masks were presented 517 ms following the memory array whereas performance did not peak for unfamiliar items until the masks were shown 867 ms following the memory array (similar to the unfamiliar shapes in Experiment 1). These results provide strong support for the workspace and unitary models of WM, as they can most easily account for the early effects of long-term memory in VWM processing.

Interestingly, there was an overall main effect for trained versus novel stimuli, a result that is in contrast with Chen et al. (2006). There are several potential reasons for this difference. First, the current experiment used a different training task than Chen et al. (2006) which may have led to stronger effects of training. Chen et al. (2006) used a change detection procedure while the current study used a 4AFC recognition task. In the 4AFC procedure, participants study only one shape at a time, compared to four in the change detection procedure, and this increased exposure to the individual shapes may have led to stronger long-term representations of those items. Experiment 2b should

provide some insights as to how much the training task matters since it will replicate the current experiment using a training procedure similar to Chen et al. (2006).

Second, in their studies, Chen et al. (2006) did not use any visual masks following the presentation of the memory array. It is possible that increased familiarity only facilitates VWM processing in the face of interfering information. As noted in Experiment 1, one leading theory of the central executive component of WM argues that the key function of the central executive is to resist the build-up of interference (e.g., Engle & Kane, 2004; Kane & Engle, 2000); thus, if training only improves consolidation when interfering masks are used, it would suggest a role for general cognitive resources in short-term consolidation. Finally, power was a problem with the experiments in Chen et al. (2006) with only 12 subjects in each of their experiments. It is possible that training effects did exist in their studies, but Chen et al. (2006) were unable to detect them.

Overall, there was an interesting trend towards performance following the longest delay (484 ms) being lower than performance following the second longest delay (367 ms), particularly for unfamiliar shapes. This suggests that there may be some sort of critical period for complex unfamiliar shapes: new information cannot be presented too soon to avoid interference during consolidation, but testing memory long following the memory array leads to decay of the visual details which reduces performance. Given that very little work has examined VWM using shapes other than colored squares, there is no previous research that could explain this effect. Future work should examine the persistence of complex visual information in the face of interfering information. This effect is, however, consistent with Cowan's (1993; 1999) argument that activated

representations in WM are vulnerable to decay and without reactivation of WM traces information will be lost. This will be revisited in the General Discussion.

There are a couple alternative interpretations to these results, however. First, in Experiment 1 participants completed a verbal suppression task during the familiar trials that was relatively demanding (e.g., repeating two random numbers aloud on each trial; see Morey & Cowan, 2005) which may have increased the WM load of the task, leading to stronger consolidation effects for the squares stimuli than would be seen without the additional load. Thus, without the articulatory suppression task, the squares stimuli may look similar to the trained shapes in the current experiment (i.e., there would be no effect of delay for any set size), again suggesting a role for higher level, central executive processes in the consolidation process. In contrast, unfamiliar stimuli may be inherently more demanding on WM resources because they are more complex, showing effects of consolidation without an additional load. This resource limited view fits with the results seen in the current experiments for performance across the set sizes: as set size (i.e., load) increases, the impact of the masks also increases, suggesting more resources are required to consolidate information as the task increases in difficulty. A key follow up to the current set of experiments should examine the role of central executive resources in consolidation.

Second, training may help to reduce interference build-up across the experiment, as discussed in Experiment 1. That is, with greater similarity between to-be-remembered items (as is the case with the stimuli used in this experiment) comes greater levels of interference between them. Training may help to reduce this interference between items

as well as reduce the impact of the interfering visual masks. This reduction in interference may be facilitated by reducing the WM load and freeing up resources, making this alternative complementary with the alternative approach outlined above.

Taking together Experiment 1 and Experiment 2b, there is strong evidence that long-term representations facilitate short-term consolidation: both inherently familiar and explicitly trained stimuli not only show performance advantages over unfamiliar stimuli, but they are faster to consolidate into VWM. This provides strong empirical support for the workspace model of WM and is the first successful demonstration of training facilitating VWM processing. These results are also the first demonstration of a very early facilitative role of LTM in processing visual information.

## CHAPTER 4 - EXPERIMENT 2B

Experiment 2a demonstrated that training improved VWM performance and facilitated short-term consolidation into VWM, providing converging evidence for the advantage of squares stimuli found in Experiment 1. Taken together, these experiments support the idea that long-term representations can facilitate early VWM processing (a key assumption of the Workspace Model of WM). However, the positive effects of training seen in Experiment 2a are in contrast to the prior work examining the influence of explicit training on VWM performance (e.g., Chen et al., 2006). Experiment 2b contrasts the training effects found in Experiment 2a with the prior literature by more closely replicating the training task used by Chen et al. (2006) and examining how these two different training tasks influence short-term consolidation. That is, Experiment 2b asks two key questions: does the type of training matter in terms of how long-term visual representations are developed and how do different training methods influence short-term consolidation?

More specifically, the 4AFC training task used in Experiment 2a focused on the individual shapes, in particular discriminating a presented shape from three other similar shapes. In Experiment 2b (and in Chen et al., 2006), a change detection training task is used that emphasizes detecting changes in four item visual arrays, a task that is more similar to the backward-masking consolidation task and less focused on discriminating

specific shapes. Thus, Experiment 2b provides a contrast between training that emphasizes individual shapes with training on the change detection task.

This contrast is key because it provides insight as to what is being consolidated into VWM during short-term consolidation, something that has not been examined in the short-term consolidation literature. If visual information is being consolidated, then training should facilitate consolidation with a 4AFC task and less so for the change detection task because of the emphasis on visual distinctions between the items. But if other information is being consolidated (e.g., practice with task or specific associations between items) then the change detection training should also facilitate consolidation. Alternatively, training using a 4AFC task may influence consolidation (as demonstrated in Experiment 2a), but using a change detection task could lead to general improvements in performance by providing practice with the task and not specifically modulating consolidation (which Chen et al., 2006 demonstrated). Thus, Experiment 2b is an important extension of the effects found in Experiment 2a because the contrast between the experiments provides insights as to what information is being consolidated and what aspects of training influence VWM performance.

### **Method**

The method is exactly the same as Experiment 2a with the following exceptions.

#### **Participants**

Seventy introductory psychology students participated for optional, partial course credit.

## **Training**

A subset of 12 new polygons were selected at random for training and the same 12 shapes were used for each participant. To parallel the training used in Chen et al. (2006), participants completed 200 change detection trials that consisted of only the 12 polygons selected for training. In each trial, four of the 12 shapes were selected and presented on the screen in an invisible 2 x 2 grid. These shapes were presented for 500 ms, followed by a mask 484 ms after the offset of the shapes, followed by a 716 ms blank retention interval. Then the test array was presented and participants determined if one of the four shapes changed between study and test (making a same/different judgement). Participants received visual feedback during the training portion only.

## **Results**

Eight participants were dropped from the analyses due to low performance (overall accuracy that was greater than two standard deviations below the mean) leaving 62 participants for the following analyses. As in Experiments 1 and 2a, any sphericity assumptions were corrected using the Huynh-Feldt correction and pairwise comparisons were conducted using the least significance differences method. All analyses are significant at the  $p < .05$  level unless otherwise noted.

### **Change Detection Training & Recognition Performance**

First, performance on the training task and the subsequent recognition tests were analyzed to examine the effectiveness of training. Overall accuracy for the change detection training task was generally low ( $M = .66$ ,  $SD = .08$ ). When broken down into four blocks, performance did increase slightly (Block 1:  $M = .64$ ,  $SD = .11$ ; Block 2:  $M$

= .67,  $SD = .10$ ; Block 3:  $M = .67$ ,  $SD = .11$ ; Block 4:  $M = .66$ ,  $SD = .12$ ) but that increase was not statistically significant [ $F(2.62, 157.38) = 1.20$ ,  $p = .31$ ,  $MSE = .01$ ,  $\eta_p^2 = .02$ ].

This pattern suggests that the change detection task was difficult and remained so throughout training.

The recognition performance data support this conclusion. Recognition accuracy was low for both the intermediate recognition test ( $M = .73$ ,  $SD = .34$ ) and the final recognition test ( $M = .76$ ,  $SD = .31$ ), although performance on both tests was significantly different from chance [Intermediate:  $t(60) = 4.55$ ; Final:  $t(58) = 6.49$ ]. To examine how learning of the shapes varied across the experiment, a paired samples t-test was conducted contrasting intermediate recognition performance with final recognition performance. This analysis showed no significant differences between the two recognition tests [ $t(57) = -.64$ ,  $p = .52$ ] suggesting that whatever representations were created during training remained robust across the experiment time interval.

Note that performance with the change detection training task is lower than 4AFC training task [4AFC Intermediate Recognition:  $M = .85$ ; 4AFC Final Recognition:  $M = .83$ ], indicating that the change detection training task was more difficult and less effective at creating visual long-term representations. However, in order to fit Experiment 2b into the allotted 1.5 hour time slot, the number of training trials was reduced to 200 compared to 240 in Experiment 2a, so it is possible that with more training trials recognition accuracy would improve. Though the fact that there was no significant increase in performance across training blocks suggests that this is not the case, or at least that significantly more trials would be needed to improve recognition.

## Accuracy

To parallel the analyses in Experiments 1 and 2a, the current analysis first examines the overall pattern of results across all three variables (stimulus type, set size, and delay) to examine if differences in short-term consolidation exist between trained and novel shapes. The 2 (stimulus type: familiar vs unfamiliar) x 4 (set size: 1, 2, 3, & 4 items) x 5 (delay: 17, 134, 250, 367, & 484 ms) repeated measures ANOVA yielded a significant main effect for set size [ $F(3, 183) = 334.35$ ,  $MSE = .02$ ,  $\eta_p^2 = .85$ ], suggesting performance varied as the number of items increased, and the stimulus type x set size interaction approached significance [ $F(3, 183) = 2.27$ ,  $p = .08$ ,  $MSE = .01$ ,  $\eta_p^2 = .04$ ], suggesting the impact of set size varied as a function of stimulus type (see pairwise comparisons below). There was no significant main effect of stimulus type [ $F(1, 61) = 1.18$ ,  $p = .28$ ,  $MSE = .07$ ,  $\eta_p^2 = .02$ ] or delay [ $F < 1$ ] and none of the other interactions reached significance [stimulus type x delay:  $F(3.71, 226.05) = 1.16$ ,  $p = .33$ ,  $MSE = .01$ ,  $\eta_p^2 = .02$ ; set size x delay:  $F < 1$ ; stimulus type x set size x delay:  $F < 1$ ].

Thus, there was no advantage for trained shapes (i.e., no main effect for stimulus type) and no evidence that consolidation varied between the two stimulus types (i.e., there was no significant 3-way interaction; see Figure 9). Since masks did not interfere with either familiar or unfamiliar shapes, this suggests that there was a generalized benefit of training: unlike Experiment 2a in which the unfamiliar shapes showed consolidation effects but the familiar shapes did not, there were no consolidation effects in Experiment 2b with unfamiliar shapes being identical in performance to the familiar shapes.

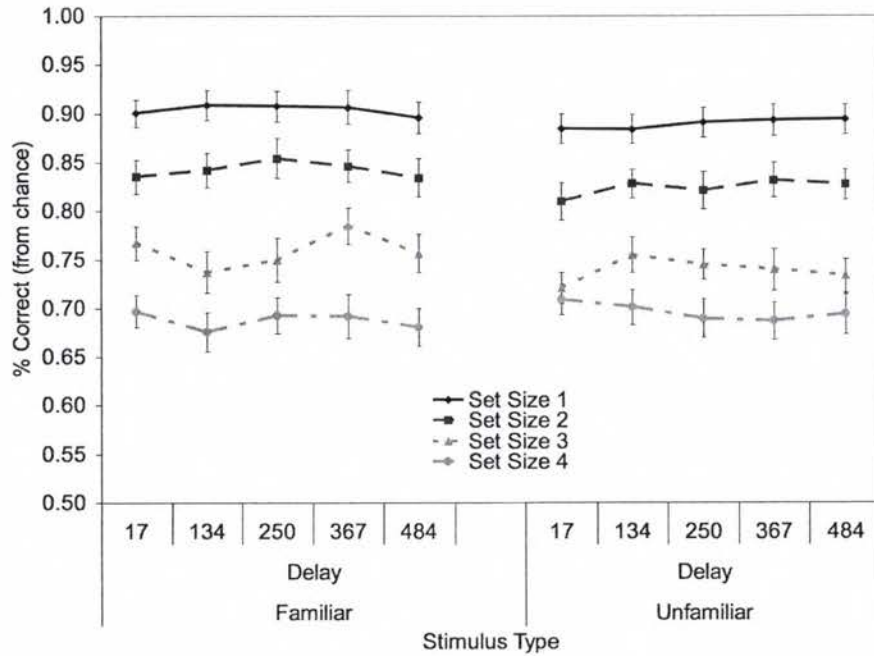


Figure 9. Unconditional Accuracy data for Experiment 2b. Error bars represent one standard error of the mean.

To further examine the main effect for set size, pairwise comparisons were conducted. These comparisons showed the same set size 1 > set size 2 > set size 3 > set size 4 pattern as in the previous experiments, suggesting that performance drops as the number of to-be remembered items increases. However, the lack of any interactions between set size and delay suggests that while the task was more difficult with four items compared to one item, there is no evidence for any differences in consolidation between the set sizes. This is contrary to both the prior experiments and to Vogel et al. (2006) in which items were slower to consolidate as set size increased (suggesting that the consolidation process is resource limited). This has an interesting implication, specifically that a four item array was consolidated as an integrated global spatial configuration. Training on the change detection task could have improved the ability to create this configuration, leading to facilitation for both trained and novel shapes.

To further examine the stimulus type x set size interaction, two separate 4 (set size) x 5 (delay) repeated-measures ANOVAs were conducted for familiar and unfamiliar shapes. The familiar analysis showed a significant main effect for set size [ $F(3, 183) = 211.71, MSE = .01, \eta_p^2 = .78$ ]. The main effect for delay [ $F(4, 244) = 1.27, p = .28, MSE = .01, \eta_p^2 = .02$ ] and the set size x delay interaction [ $F < 1$ ] did not reach significance (see Figures 10 and 11). Pairwise comparisons to examine the set size main effect showed the same set size 1 > set size 2 > set size 3 > set size 4 pattern seen in the overall analysis and in the previous experiments, again suggesting that the task was more difficult as the number of to be remembered items increases.

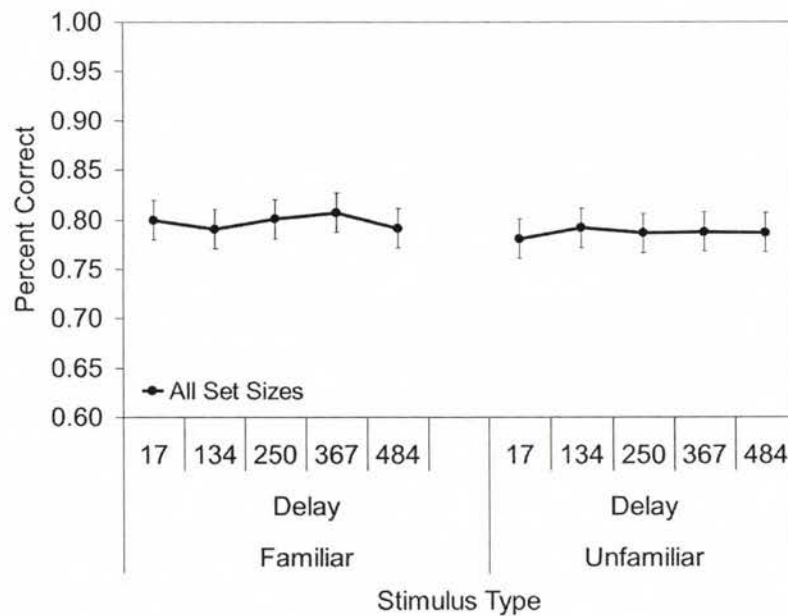


Figure 10. Experiment 2b accuracy across the five delay conditions collapsed across set size conditions for familiar and unfamiliar stimuli. Error bars represent one standard error of the mean.

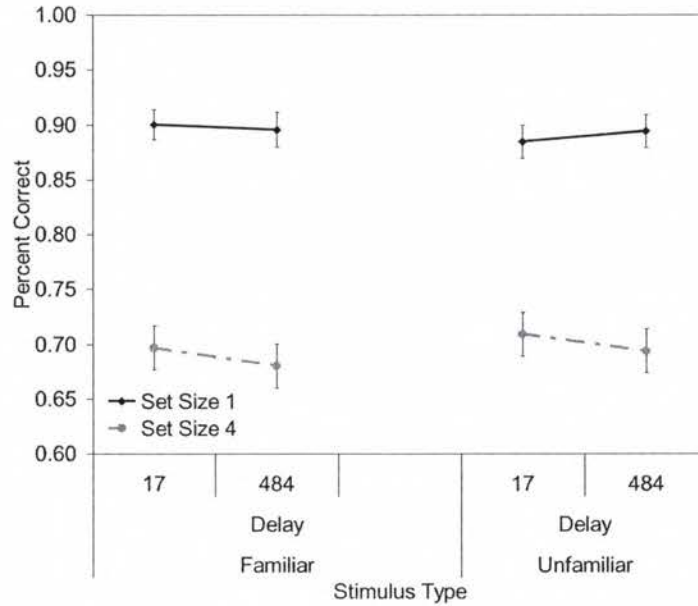


Figure 11. Experiment 2b accuracy for the shortest and longest delay conditions (17 ms and 484 ms) with Set Size 1 and Set Size 4 for familiar and unfamiliar items. Error bars represent one standard error of the mean.

The unfamiliar analysis also yielded a significant main effect for set size [ $F(2.79, 170.24) = 162.58, MSE = .02, \eta_p^2 = .73$ ]. The main effect for delay and the set size x delay interaction did not reach significance [ $F < 1$  for both]. Pairwise comparisons for the set size main effect also showed the set size 1 > set size 2 > set size 3 > set size 4 pattern as the familiar analysis. In contrasting the familiar and unfamiliar analyses, the data suggest that set size had a greater impact on performance for familiar stimuli ( $\eta_p^2 = .78$ ) than on unfamiliar stimuli ( $\eta_p^2 = .73$ ). Given the performance advantages found in Experiment 2a for familiar stimuli, this effect is surprising; one would expect familiar items to be less affected by an increase in task difficulty.

The disadvantage for trained shapes could be due to the fact that during training, the shapes were always presented in the same spatial configuration (a square) but in the consolidation task the positions of the items varied randomly. It is possible that during

training participants were learning not only the shapes (since recognition accuracy was above chance) but also the relationships between the shapes (i.e., where each shape was relative to the other shapes). When that relational information was disrupted following training, performance dropped. This is consistent with work from Jiang et al. (2000) who demonstrated that VWM performance declines if the spatial configuration of the shapes changes, even when that configuration is irrelevant to the task (as it was in the current experiment). The role that processing relational information between shapes will be revisited in the General Discussion.

### **Conditionalized Accuracy Analysis**

The above analysis suggests that there was a generalized benefit of training as both familiar and unfamiliar shapes were relatively impervious to visual masks even at short delay intervals. However, given the overall low recognition and training performance, it is difficult to draw conclusions regarding the influence of long-term representations in short-term consolidation. That is, there could be no evidence for masks influencing performance in the above analysis due to noise in the data from participants who did not learn the shapes. Thus, the same analysis as above was conducted on those participants that reached at least 90% accuracy (only missing one item on the recognition test) on the intermediate recognition test to examine if consolidation varies between trained and novel stimuli with high levels of recognition. Excluding those below 90% resulted in dropping 30 participants, leaving 32 participants for the following analysis.

To examine how consolidation differed between trained and novel shapes conditionalizing based on recognition performance, a 2 (stimulus type: familiar vs

unfamiliar) x 4 (set size: 1, 2, 3, & 4 items) x 5 (delay: 17, 134, 250, 367, & 484 ms) repeated measures ANOVA was conducted. This analysis yielded a significant main effect of set size [ $F(3, 93) = 173.31, MSE = .02, \eta_p^2 = .85$ ] and a significant stimulus type x delay interaction [ $F(4, 124) = 3.31, MSE = .01, \eta_p^2 = .10$ ]. There was no main effect for stimulus type [ $F < 1$ ] or delay [ $F < 1$ ] and none of the other interactions reached significance [stimulus type x set size:  $F(3, 93) = 1.17, p = .32, MSE = .01, \eta_p^2 = .04$ ; set size x delay:  $F < 1$ ; stimulus type x set size x delay:  $F(10.56, 327.44) = 1.24, p = .26, MSE = .01, \eta_p^2 = .04$ ], including the stimulus type x set size interaction that was significant in the unconditionalized analysis. This suggests that with a high level of knowledge of the shapes, participants were equally able to remember a set of one, two, three, or four trained shapes as novel shapes. Thus, overall, there are no large differences between trained and novel shapes, though unlike the unconditionalized analysis there is some evidence here that consolidation differed between trained and novel shapes (i.e., stimulus type x delay interaction).

Pairwise comparisons for the main effect for set size followed the same set size 1 > set size 2 > set size 3 > set size 4 pattern seen throughout this set of experiments. Again, this pattern replicates Vogel et al. (2006) and suggests that the consolidation task gets more difficult as the set size increases. Similar to the unconditionalized analysis there is no evidence that consolidation differed between the set sizes (i.e., there was no significant interaction between set size and delay), suggesting that at set size 4, participants did not require more time to consolidate information into VWM compared to set size 1, but overall accuracy was lower with more items. The lack of a main effect for

delay provides no evidence for any consolidation effects (i.e., no effects of visual masks) and no main effect for stimulus type suggests that there was no advantage for trained shapes.

To further examine the stimulus type x delay interaction, separate 4 (set size) x 5 (delay) repeated measures ANOVAs were conducted for each stimulus type. The familiar analysis yielded a significant main effect for set size [ $F(2.63, 81.51) = 115.54, MSE = .01, \eta_p^2 = .79$ ]. The main effect for delay [ $F(4, 124) = 1.18, p = .32, MSE = .01, \eta_p^2 = .04$ ] and the set size x delay interaction [ $F(12, 372) = 1.12, p = .34, MSE = .01, \eta_p^2 = .04$ ] did not reach significance. Pairwise comparisons for set size showed the set size 1 > set size 2 > set size 3 > set size 4 pattern, paralleling the pattern seen in the overall 3-way ANOVA. While there was no main effect for delay, pairwise comparisons were also calculated for delay since there was a significant stimulus type x delay interaction. First, to parallel the analyses for the previous experiments, the difference in accuracy between the 17 ms and 484 ms conditions was examined for set sizes 1 and 4. This analysis showed no significant differences in accuracy with set size 1 between the 17 ms delay condition ( $M = .91, SD = .11$ ) or the 484 ms delay condition ( $M = .92, SD = .10, t(31) = -1.29, p = .21$ ). This was also shown for a set size of 4 (17 ms:  $M = .71, SD = .11$ ; 484 ms:  $M = .71, SD = .14, t(31) = .29, p = .77$ ). Further pairwise comparisons showed only one comparison reached significance, with accuracy following a 250 ms delay being significantly higher than a 484 ms delay. Thus, there was a difference between two non-adjacent delay intervals, but without a significant main effect for delay this difference likely reflects noise in the data.

For the unfamiliar analysis, there was also a main effect for set size [ $F(3, 93) = 90.80, MSE = .01, \eta_p^2 = .75$ ]. The main effect for delay [ $F(3.62, 112.19) = 1.52, p = .21, MSE = .01, \eta_p^2 = .05$ ] and the set size x delay interaction [ $F < 1$ ] did not reach significance. Pairwise comparisons for set size showed the set size 1 > set size 2 > set size 3 > set size 4 pattern. To parallel the familiar analysis, first the difference in accuracy between the 17 ms and 484 ms conditions was examined for set sizes 1 and 4. This analysis showed no significant differences in accuracy with set size 1 between the 17 ms delay condition ( $M = .91, SD = .10$ ) or the 484 ms delay condition ( $M = .92, SD = .10; t(31) = -.59, p = .56$ ). This was also shown for a set size of 4 (17 ms:  $M = .73, SD = .12$ ; 484 ms:  $M = .73, SD = .13; t(31) = -.15, p = .88$ ). Further comparisons indicated that accuracy following a 17 ms delay was significantly lower than accuracy following a 367 ms or 484 ms delay, suggesting that performance following shorter delays was worse compared to performance following longer delays. Again, this pattern is difficult to interpret given the lack of a main effect for delay, but the differences between familiar and unfamiliar stimuli in terms of the pattern of accuracy across the delay conditions is likely driving the stimulus type x delay interaction.

Overall, the results from Experiment 2b are slightly ambiguous. While recognition performance showed low overall recognition of the trained shapes, there was evidence that training improved performance for both familiar and unfamiliar shapes (i.e., there was no effect of masks even at the earliest delay). The conditionalized analysis provided some evidence that consolidation was better for trained items, suggesting that with high levels of stimulus knowledge long-term memory will facilitate consolidation

for those items. However, given the low power in the conditionalized analysis from excluding half of the sample, it is unclear if these effects are meaningful or simply noise in the data. Comparing these results with those in Experiment 2a, however, suggests that the change detection training task led to a generalized benefit of training

### **Capacity**

To parallel the analyses in Experiment 1 and 2a, capacity ( $K$ ) was calculated to compare capacity estimates between trained and novel shapes. This estimate was calculated for each stimulus type condition at set size 4 across all delay conditions. A paired-samples t-test between the familiar ( $M = 1.62$ ;  $SD = .90$ ) and unfamiliar ( $M = 1.60$ ;  $SD = .81$ ) stimuli showed no significant differences in VWM capacity [ $t(61) = .21$ ,  $p = .84$ ]. This is in contrast to Experiment 2a in which trained items had higher capacity than novel items. This indicates either that there was no impact on capacity or that the training used in this experiment led to an overall increase in capacity for both familiar and unfamiliar items. To examine this latter possibility, the overall  $K$  was calculated for Experiments 2a and 2b and compared in a paired-samples t-test. This test showed that capacity was significantly higher in Experiment 2a ( $M = 1.44$ ;  $SD = .35$ ) than in Experiment 2b [ $M = 1.30$ ,  $SD = .47$ ;  $t(61) = 2.00$ ,  $p = .050$ ]. Thus, capacity was not generally higher in 2b indicating that change detection training did not increase the overall capacity of VWM.

### **Discussion**

Experiments 1 and 2a provided evidence for a role of long-term memory in facilitating short-term consolidation. The goal for Experiment 2b was to extend these

results, specifically by generalizing them to a different training task. The results showed no significant differences between trained and novel shapes for either accuracy or capacity and no clear evidence that consolidation differed between the stimuli types. This is in contrast to Experiment 2a in which there was both an overall advantage for trained shapes (higher accuracy and capacity compared to novel shapes) and evidence that trained shapes were consolidated faster than novel shapes. Overall, the results from Experiment 2b suggest that change detection training led to a benefit for both familiar and unfamiliar items.

What do experiments 2a and 2b tell us about training visual stimuli and how training influences performance? These results suggest that the type of training task impacts the effect training will have on VWM performance: training tasks that focus on discriminating visual items facilitates short-term consolidation for the trained items but a training task that is more task specific leads to a generalized benefit of task practice. Specific to this set of experiments, a 4AFC training task in which participants recognized specific shapes led to stimulus specific improvements in VWM performance while the a change detection task was less effective at creating long-term representations of the stimuli (based on recognition performance) but the training transferred to novel shapes.

This difference between experiments suggests that short-term consolidation is multi-dimensional. During consolidation, visual features are being consolidated (as demonstrated in Experiment 2a) as well as task specific properties such as creating a global spatial configuration of the items (i.e., the overall spatial layout; Jiang et al., 2000). In Experiment 2b, the change detection training task required knowledge of the

shapes but it also required knowledge of the relations between the shapes. Thus, participants likely improved in their ability to create the global configuration, which in turn improved their consolidation performance. This is supported by the lack of a set size x delay interaction. Since four items did not require more time to consolidate than a single item, this indicates that each item was not consolidated independently but rather as an integrated spatial configuration.

This latter conclusion is in line with Chen et al. (2006). They also showed an overall benefit for training, but no specific benefit for trained versus novel shapes. They argued that LTM was not helpful in VWM processing because the resolution of the long-term visual representations was poor except for visual information that has been learned over long periods of time (e.g., a relative's face) or by extreme exposure (e.g., celebrity faces). While Experiment 2b replicates their results, Experiment 2a challenges the conclusion that long-term visual representations cannot improve VWM processing. The combined experiments suggest that the type of training used will greatly impact how long-term representations influence VWM processing.

More specifically, in order to create stable long-term memory representations the current results suggest that stimulus specific training is important, particularly distinguishing individual shapes from one another. Training in this manner allows for detailed representations that can assist VWM processing. For example, Brady et al. (2008) used a repeat detection task (in which participants see individual objects and have to hit a button if that object was shown earlier in the experiment) that emphasized individual objects and showed highly detailed long-term memory for those objects. Taken

together with Experiment 2a, this suggests that in forming long-term representations of visual stimuli, a key process is developing a unique representation for the visual properties of that shape that distinguish it from other shapes.

This work does not, however, provide insight as to how this visual information is represented or organized in long-term memory. Brady et al. (2008) suggest that to maintain detailed visual information, long-term representations may be stored at different levels of the visual system hierarchy (i.e., information stored in early visual areas in addition to higher level visual areas), with each level providing different types of information allowing for the most detailed representation possible. This view is in line with Kosslyn's model of mental imagery in the brain (Kosslyn, 2005; Kosslyn, Ganis, & Thompson, 2001; Kosslyn, Thompson, & Ganis, 2006), which argues for a role of both early visual processing areas (e.g., the primary visual cortex) and higher level visual association areas (e.g., extrastriate visual cortex) in creating and maintaining mental images based on long-term visual representations. Arguing for a role of early visual processing areas is consistent with the early facilitation effects found in Experiment 2a, and work using the backward-masking paradigm is a useful way to not only understand how information goes from perception to WM but also how long-term visual representations are represented in the mind.

However, the downside to stimulus specific training is that it will only improve VWM performance for familiar shapes. Experiment 2b suggests that in order to show transfer of learning to novel shapes, the training task must be similar to the criterion task. A critical result in the current experiment is that training did not just improve overall

performance, it led to faster consolidation of completely novel shapes. That demonstrates a high degree of transfer, which is seldom seen especially in visual or perceptual learning tasks (e.g., normal vs. inverted faces; see Buttle & Raymond, 2003; Curby & Gauthier, 2009). Additionally, since the training task and the consolidation task were not identical (as they were in Chen et al., 2006), this is a demonstration of far transfer in which learning of information transfers to new contexts (Barnett & Ceci, 2002).

Overall, the combined results from Experiments 2a and 2b provide support for the Workspace and Unitary models of WM by demonstrating that visual long-term representations can facilitate early VWM processing. However, how visual items were initially learned will have an impact on how effective those items will be in facilitating VWM processing (i.e., stimulus specific or generalized). Given that change detection training improved consolidation with novel shapes (compared to novel shapes in Experiment 2a), both the workspace and unitary models may need to be revised to explain how novel shapes operated like familiar shapes. It may be that practice on task leads to more efficient strategies that are implemented by the central executive component of WM. Currently, the Unitary Model makes more specific predictions about how the central executive can control the contents of WM (or the focus of attention) compared to the Workspace Model, though neither model specifically addresses how the central executive could stabilize VWM representations.

## CHAPTER 5 - EXPERIMENT 3

Experiments 2a and 2b demonstrated that explicit training facilitates short-term consolidation and improves VWM performance. Experiment 2a showed that stimulus specific training leads to faster consolidation and better performance compared to novel shapes. Experiment 2b, however, demonstrated that change detection training facilitated consolidation for both trained and novel shapes, suggesting that consolidation is processing more than just visual information. Rather, Experiment 2b suggests that in addition to visual information, relational information between items is also consolidated. In particular, the fact that there was no interaction between set size and delay in Experiment 2b suggests that individual items were not consolidated (which would require more time as the number of items increases as demonstrated in Experiments 1 and 2a as well as in Vogel et al., 2006) but rather that visual arrays were consolidated together as a global spatial configuration.

The goal for Experiment 3 is to replicate this effect and more directly examine the influence of configural information in short-term consolidation using a different approach to training. Specifically, Experiment 3 uses an incidental learning paradigm to examine how repeating the same set of shapes in the same configuration over the experiment impacts consolidation for those repeated items over novel items (Olson & Jiang, 2004; Olson, et al., 2005). If both learning the configuration of the shapes and the shapes

themselves are part of the consolidation process, then only repeated shapes should show facilitation (due to stimulus exposure) and there should be no interactions with set size for repeated items because a set of four items can be consolidated as an integrated configuration. Thus, in Experiment 3 the configuration of the items is trained in addition to the visual stimuli.

This approach to incidental learning dates back to Hebb (1961). Hebb was interested in how short-term representations become long-term representations. In particular, he was interested in whether or not previously remembered stimuli left a record in the nervous system that could later become a long-term representation (Stadler, 1993). If information is incidentally repeated across an experiment, would that information leave a trace that could in turn influence behavior at a later point? In his experiment, participants completed a digit span task in which a string of nine numbers were presented and they had to repeat the numbers back in serial order. Unbeknownst to the participants, however, on every third trial the digits were presented in the same order. Hebb (1961) showed that performance for those repeating numbers was better compared to the random trials.

Hebb (1961) argued that encoding and recalling a list of numbers creates a momentary representation of that list, called an activity trace (Stadler, 1993). Over the course of an experiment, this activity trace could either be rewritten by subsequent trials or a more stable structural trace could be left behind that could be used to improve performance for future trials. Given that repeated trials were recalled better than novel

trials, Hebb (1961) concluded that previously remembered information, even if only maintained briefly, can have a permanent effect on memory.

This framework speaks to the current set of experiments. First, the current set of experiments can address the issue of what constitutes an activity trace. This trace could consist of a representation of the individual shapes, a representation of the global spatial configuration of the set of items, or both. Currently the results suggest that both are playing a part and Experiment 3 further examines the role of configuration information and how it influences VWM processing. Second, Hebb's (1961) framework provides a useful way to address a current problem in the WM literature: how is novel information, such as unfamiliar polygons, learned (i.e., how novel information transitions between WM and LTM)? This is particularly the case for the Unitary Model, which clearly explains how activated LTM gets into WM but has more difficulty explaining how novel information goes from WM to LTM. If repeated shapes show an advantage over novel shapes and there is no set size x delay interaction for trained shapes, that would suggest a role for configural information in creating long-term representations of visual items.

This assertion is supported by work on implicit memory in the visual domain. This work has shown that the visual system is very sensitive to repeating information. For example, Chun and Jiang (1998; see also, Jiang, Song, & Rigas, 2005) showed that visual search was faster with search displays that repeated over the course of the experiment (without participants' knowledge) compared to novel displays, an effect called contextual cueing. Jiang et al. (2005) suggest that the visual system maintains a high capacity for learning repeating contexts to compensate for the severe capacity limitations of VWM. In

the context of VWM, Jiang et al. (2000) showed that memory for an item in a visual array was significantly reduced when the locations of the irrelevant distractor items in the array changed, suggesting that context also plays a key role in VWM. Thus, implicit learning mechanisms influence learning and remembering of visual information and as such it is reasonable to assume that these mechanisms could play a role in consolidating information into VWM.

This experiment assesses the role that incidental learning of visual items in a consistent spatial configuration plays in facilitating short-term consolidation. If repeated shapes are consolidated more quickly than novel shapes, that would suggest a role for LTM in short-term consolidation. Additionally, if set size does not vary as a function of mask delay, that would support the conclusion from Experiment 2b that spatial configuration information being consolidated into VWM along with visual information.

## **Method**

### **Participants and Design**

Twenty-four introductory psychology students participated in this experiment for optional, partial course credit. Participants were required to have normal or corrected to normal vision. A 2 (stimulus type: repeated vs. novel) x 4 (set size: 1, 2, 3, & 4 items) x 5 (mask delay: 12, 134, 250, 367, & 484 ms) within-subjects design was used.

### **Stimuli**

The entire set of polygons used in Experiments 2a and 2b will be used in Experiment 3. Four of these shapes will be randomly selected for repetition across the experiment.

## Procedure

The procedure for Experiment 3 is similar to the previous experiments, particularly the unfamiliar trials in Experiment 1. Participants completed a total of 300 consolidation trials with every third trial repeating one to four of the subset of four shapes depending on the set size for that trial. Therefore, 100 trials in Experiment 3 repeated the same four shapes, resulting in five trials per stimulus type x delay combination for repeated trials. The shapes presented in these repeating trials were be presented in the same location on each trial to maximize learning (Olson, et al., 2005). The remaining 200 trials consisted of novel shapes<sup>6</sup>, resulting in 10 trials per stimulus type x delay combination.

Following the consolidation task, participants were tested for awareness of the repeating shapes. As a part of the demographics questionnaire participants were asked if they were aware of any repeating information and what they thought that information was. Also, participants completed a brief recognition task in which they were shown a 4 x 4 matrix of 16 shapes and asked to circle the four shapes that were repeated during the experiment. The other eight shapes were completely novel shapes that were not presented in any part of the experiment. These measures combined will provide insights as to whether or not participants were aware of the repeating shapes to examine if the representations formed during the experiment were implicit or if there was some explicit knowledge about the shapes.

---

<sup>6</sup> Due to constraints in the ability to generate Attneave shapes, shapes for the novel trials were selected from a pool of 400 shapes that repeated once during the experiment. However, the shapes were separated by at least 100 trials making any learning of those shapes highly unlikely.

## Results

One participant was dropped from the analyses because their overall accuracy was two standard deviations below the mean, leaving 23 for the following analyses. As with the prior experiments, any sphericity violations were corrected using the Huynh-Feldt correction and pairwise comparisons were conducted using the least significant differences method. All analyses are significant at the  $p < .05$  level unless otherwise noted.

### Recognition Test Accuracy

To assess participants' explicit knowledge of the repeating shapes, accuracy in selecting the correct four shapes out of a set of 16 was calculated. Participants were very accurate at selecting the shapes with the average correct recognition at 86%. This is in contrast to the self-report on the demographics questionnaire: only 5 participants (22%) reported that they noticed repeating information and that the repeating information was repeating shapes (though no participant said specifically that the shapes repeated on every third trial). This data suggests that participants developed some explicit awareness of the repeating information, meaning any effects detailed below are likely a combination of explicit and implicit knowledge of the shapes.

### Accuracy

To examine if consolidation differed between the stimuli types, a 2 (stimulus type: repeating vs. novel) x 4 (set size: 1, 2, 3, & 4) x 5 (delay: 17, 134, 250, 367, & 484) repeated-measures ANOVA was conducted on the accuracy data. This analysis yielded significant main effects for stimulus type [ $F(1, 22) = 17.09$ ,  $MSE = .02$ ,  $\eta_p^2 = .48$ ] and set

size [ $F(3, 66) = 123.92, MSE = .02, \eta_p^2 = .85$ ] as well as a significant stimulus type x delay interaction [ $F(4, 88) = 2.83, MSE = .02, \eta_p^2 = .11$ ]. There was no main effect for delay [ $F(4, 88) = 1.67, p = .16, MSE = .02, \eta_p^2 = .07$ ] and none of the other interactions reached significance [stimulus type x set size:  $F(3, 66) = 1.67, p = .18, MSE = .02, \eta_p^2 = .07$ ; set size x delay:  $F(12, 264) = 1.05, p = .40, MSE = .02, \eta_p^2 = .05$ ; stimulus type x set size x delay:  $F < 1$ ]. These results suggest that consolidation did differ between repeating and novel shapes, specifically that there were differences between the two stimulus types in terms of the impact of the mask delay intervals but this pattern did not interact with set size as in previous experiments (see Figure 12). This suggests that in this case, consolidation was not resource limited as the impact of visual masks with a set of four items was equivalent to the impact visual masks had at set size 1 and supports the notion that spatial configuration information is important for the consolidation process.

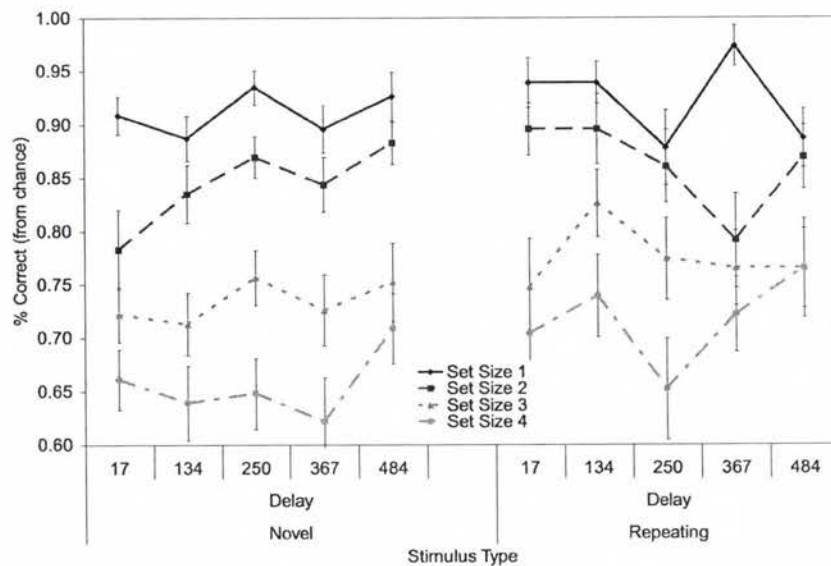


Figure 12. Accuracy data for Experiment 2b. Error bars represent one standard error of the mean.

Pairwise comparisons were conducted to examine the main effects for stimulus type and set size. Accuracy was significantly higher for repeating shapes than for novel

shapes, suggesting that implicit training improved VWM performance. Since spatial information was repeated in addition to visual information, however, it is not clear if this advantage is due to repeating the spatial context of the items or to learning the shapes or both. Given that recognition of the shapes was so high, it is likely that both knowledge of the shapes combined with the repeating locations led to higher performance for repeating shapes.

In terms of set size, the same set size 1 > set size 2 > set size 3 > set size 4 pattern was found that the previous experiments also demonstrated. Again, this pattern suggests that the consolidation task gets more difficult as the set size increases. Similar to Experiment 2b, there is no evidence in Experiment 3 that consolidation differed between the set sizes (i.e., there was no significant interaction between set size and delay), suggesting that at set size 4, participants did not require more time to consolidate information into VWM compared to set size 1, but overall accuracy was lower with more items. As noted above, this suggests that consolidation was not resource limited and that spatial configuration information aided the consolidation process.

To examine the stimulus type x delay interaction, separate 4 (set size) x 5 (delay) repeated-measures ANOVAs were conducted for novel and repeating items. The novel analysis showed no significant set size x delay interaction did not reach significance [ $F < 1$ ]<sup>7</sup>. That is in contrast with the predictions for this experiment, which expected that set size would not vary as a function of delay interval for repeating shapes only. Showing

---

<sup>7</sup> There were a significant main effects for set size [ $F(3, 66) = 142.64, MSE = .01, \eta_p^2 = .87$ ] and delay [ $F(4, 88) = 3.03, MSE = .02, \eta_p^2 = .12$ ]

this pattern for novel shapes indicates that repeating shapes in the same configuration may also transfer to novel shapes, mirroring the transfer effects shown in Experiment 2b.

To examine the main effect for delay, first accuracy at the shortest and longest delay intervals were compared to see if accuracy increased with more time between the memory array and the mask array. A paired-samples t-test between the 17 ms ( $M = .76$ ;  $SD = .09$ ) and 484 ms ( $M = .81$ ;  $SD = .09$ ) conditions showed that accuracy was significantly higher in the 484 ms condition than in the 17 ms condition [ $t(22) = -2.91$ ]. This suggests that visual masks led to interference when presented too soon following the memory array and additional time was needed to stabilize visual representations so that they were not susceptible to interference.

Further pairwise comparisons were conducted between all the delay conditions to get a sense of when consolidation occurred. These comparisons showed that accuracy following both a 250 ms delay and a 484 ms delay was significantly higher than accuracy following a 17 ms, 134 ms, or 367 ms delay (see Figure 13). This pattern is unusual in comparison with the previous experiments, and suggests that consolidation occurred as soon as 250 ms after the offset of the memory array (since accuracy was higher than the two shortest delays), but after that point the story becomes less clear since performance was significantly lower at 367 ms and then went back up at 484 ms. Overall, however, the data suggest that consolidation occurred approximately 250 ms following the offset of the memory array.

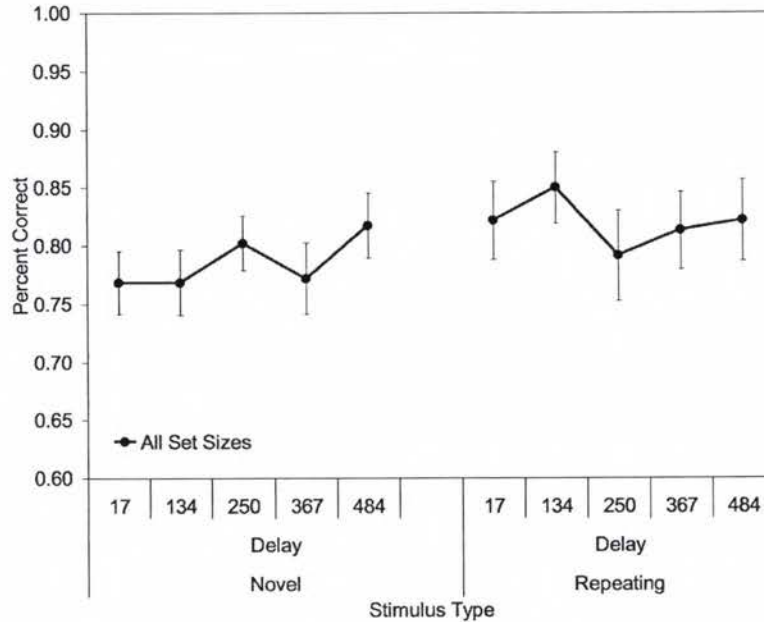


Figure 13. Experiment 3 accuracy across the five delay conditions collapsed across set size conditions for familiar and unfamiliar stimuli. Error bars represent one standard error of the mean.

The repeating analysis also did not show a set size x delay interaction [ $F(12, 264) = 1.18, p = .30, MSE = .03, \eta_p^2 = .05$ ]<sup>8</sup>. Note that there was no significant effect of delay with the repeating shapes while there was an effect for delay with novel shapes. This difference is driving the overall stimulus type x delay interaction found in the 3-way ANOVA, and suggests that the visual masks did not interfere with consolidation at any delay. This mirrors the effects seen in Experiment 2a and 2b, with trained shapes showing no evidence for consolidation and suggests that trained items were not susceptible to interference from the visual masks even with larger set sizes indicating an early role for long-term memory in short-term consolidation. Critically, the lack of the set size x delay interaction suggests that participants were able to create a global configuration of the

<sup>8</sup> This analysis did yield a significant main effect for set size [ $F(3, 66) = 38.03, MSE = .03, \eta_p^2 = .63$ ]. The main effect for delay did not reach significance [ $F(4, 88) = 1.90, p = .12, MSE = .02, \eta_p^2 = .08$ ].

repeating items supporting a key role in processing relational information in short-term consolidation.

### **Capacity**

To parallel the analyses in the previous experiments, capacity ( $K$ ) was calculated to examine any capacity differences between repeating and novel shapes. This estimate was calculated for both stimulus type conditions at set size 4 and compared using a paired-samples t-test. This analysis showed that capacity for repeating items ( $M = 1.58$ ;  $SD = .95$ ) was marginally significantly higher than novel items ( $M = 1.27$ ,  $SD = .72$ ;  $t(22) = -1.97$ ,  $p = .06$ ). This suggests that training a set of four shapes incidentally leads to higher VWM capacity for those repeating shapes.

### **Discussion**

Experiment 3 expanded on Experiments 2a and 2b by examining the role that incidentally learned visual items play in short-term consolidation. The results showed a significant benefit for repeated shapes, and some evidence that incidentally learned shapes facilitate short-term consolidation. These results support the conclusions drawn in the previous experiments: long-term memory influences very early processing in VWM and short-term consolidation is a multidimensional process that processes both visual and relational information. Thus, Experiment 3 provides strong converging evidence that long-term representations not only improve VWM performance but can facilitate early VWM processing, supporting the Workspace Model of WM.

This effect is consistent with the literature on contextual cueing, suggesting that visual memory is sensitive to repeating context (i.e., the same shapes in the same four

locations every third trial) and can utilize that repeating information to facilitate processing even at very early stages (e.g., Chun & Jiang, 1998; Jiang et al., 2005). However, given the high level of recognition in this experiment, it is possible that the repeating information became explicit over the course of the experiment and it was these explicit representations that facilitated consolidation. Olson and Jiang (2004) also showed high levels of object recognition in their experiment, though they did not find any advantage for repeated over novel shapes. Thus, the training effects found in current set of experiments are likely due to a mixture of explicit and implicit memory processes.

The current experiment offers additional insight as to what is being consolidated into VWM. In particular, because repeating shapes were presented in the same spatial locations across the experiment, Experiment 3 suggests that another key aspect to consolidation has to do with the spatial configuration of the items in a visual array (i.e., how the items are arranged relative to each other). That is, the location of the items in the array relative to the other items seems to be consolidated in addition to the visual information. This is also supported by the lack of any interactions with set size with repeated shapes. When given a repeating spatial configuration, no additional time is required to consolidate additional items because they are integrated into a unitary configuration. This is consistent with Jiang et al. (2000; see also, Blalock & Clegg, 2010) who argued that VWM is organized based on this relational global spatial configuration, and this organization is a fundamental property of VWM. The current experiment suggests that not only is this true, but that maintaining relational information can aid VWM at very early processing stages.

However, this pattern was also shown for novel trials. This, along with the transfer demonstrated in Experiment 2b, suggests that training configural information can transfer to new shapes and new configurations. In Experiment 3, however, this training was not enough to lead to equivalent performance between repeating and novel shapes. Instead, the incidental training used in the current experiment allowed participants to reduce the overall load of the task at high set sizes by integrating items into a single global configuration.

Additionally, demonstrating that the spatial configuration of items is also consolidated into VWM helps to explain the effect of set size. In particular, while Vogel et al. (2006) argued that consolidation is a resource limited process because it took longer to consolidate four items compared to one item, that does not explain why there was no evidence for consolidation with a single item. That is, even with a single visual item, there should be some effect of visual masks and that is not the case for the current experiment or for Vogel et al. (2006). Experiment 3 suggests that one reason why set size 1 operates differently than set sizes 2, 3, or 4 is because there is no configural information that needs to be consolidated. This effect, combined with the results from Experiment 2b provide compelling evidence that not just visual information is consolidated, but also the relational information between items in the array. As the number of items increases, so does the amount of relational information, leading to the interactions with set size in the earlier experiments. When the configuration information is trained, this interaction goes away, suggesting that with training participants are faster to create the configuration and thus are able to consolidate the items as a single unit versus item by item.

The current experiment results are in contrast to Olson and Jiang (2004) who did not show any advantage for repeating shapes. One reason for this discrepancy is that in the current experiment each repeating item was presented in the same spatial location whereas the repeating shapes were shuffled to new locations in Olson and Jiang (2004). This consistent context for the repeating shapes may have led to the improvements in performance for repeating shapes in the current experiment. Another explanation, however, is that as with the explicit training done by Chen et al. (2006), Olson and Jiang (2004) did not present any interfering visual masks following the memory array. Thus, repeating information may only be useful in the face of interference, again suggesting a role for higher level central executive processes in the consolidation process.

Overall, the current experiment demonstrates converging evidence that long-term representations facilitate consolidation into VWM (whether they were developed explicitly or incidentally) and that both visual and relational information is consolidated into VWM. This provides strong support for the workspace and unitary models of WM, and provides a useful paradigm to examine how information goes from WM to LTM through repeated exposure. This link is particularly important for the unitary model as currently the model does not specify how novel information goes from being an activated WM representation to LTM information during learning. In the current case, providing a consistent configuration between the items facilitated this process indicating a crucial role for relational information in forming long-term representations for visual information.

## CHAPTER 6 - GENERAL DISCUSSION

The current set of experiments set out to answer one key question: how do long-term representations facilitate short-term consolidation of information into VWM? Experiment 1 examined this question by comparing the short-term consolidation process between two extreme stimulus types (simple squares and complex polygons) and demonstrated that complex shapes with no long-term memory representations require more time to consolidate into VWM than simple squares that do have visual long-term representations. Experiment 1, thus, provided preliminary evidence that long-term representations can facilitate short-term consolidation into VWM.

Experiment 2a eliminated any confounds due to stimulus differences in Experiment 1 by training participants on a subset of complex shapes then comparing consolidation between these trained shapes and novel shapes. Complementing the results from Experiment 1, Experiment 2a demonstrated that not only were trained shapes better remembered compared to novel shapes, they were also consolidated faster providing strong evidence for a role of long-term memory in facilitating short-term consolidation. Given that this positive effect of training was unique to the literature, Experiment 2b examined the impact of long-term memory on short-term consolidation using a different training task that more closely mirrored prior work. Using this different training task, both trained and novel shapes showed resistance to visual masks, demonstrating a

generalized effect of change detection training on VWM performance. These two experiments combined provided insights about the circumstances under which long-term representations will influence memory as well as what information is being consolidated into VWM (discussed below).

Finally, Experiment 3 took a different approach to examining how long-term memory representations that were trained incidentally can influence short-term consolidation into VWM. In this experiment, the same four shapes were presented on every third trial without participants' awareness and performance on these repeated shapes was compared to novel trials. Similar to Experiment 2a, repeated shapes were remembered better and there was some evidence that repeated shapes were consolidated faster than novel shapes. Additionally, Experiment 3 demonstrates a key role for processing configuration information in facilitating consolidation and in developing visual-long term representations.

Table 1.  
*Summary of Experimental Results and Conclusions for Experiments 1, 2a, 2b, and 3.*

	<b>Method</b>	<b>Key Result</b>	<b>Main Conclusion</b>	<b>Model?</b>
<b>Experiment 1</b>	Squares vs. Polygons	Squares > Polygons	Stimulus complexity & familiarity influence consolidation	Workspace & Unitary
<b>Experiment 2a</b>	4AFC training vs. Novel	Trained > Novel	LTM facilitates early VWM processing; direct link between perception & WM	Workspace & Unitary
<b>Experiment 2b</b>	Change Detection Training vs. Novel	Trained = Novel	LTM facilitates early VWM processing; task specific training leads to transfer	Workspace & Unitary
<b>Experiment 3</b>	Repeating vs. Novel	Repeating > Novel	Incidentally learned shapes facilitate consolidation; spatial configuration information key to VWM processing	Workspace & Unitary

Across the experiments, while some of the effects sizes were relatively low, the fact that there was any effect of visual masks at all provides insight into how this

information is processed. That is, the primary question going into these experiments was not how large of an impact there was but if there was an impact at all. Taken together, these results demonstrate that not only does long-term memory influence VWM processing, it is involved at a very early stage, facilitating the process of stabilizing fragile WM representations. These findings have important implications for current WM theories in terms of how they account for the relationship between WM and LTM, as well as for understanding the process of short-term consolidation which plays such a critical role in everyday visual processing by strengthening visual information to reduce interference.

#### **What does this tell us about the relationship between WM and LTM?**

Currently, one of the largest differences between WM models is how they account for the relationship between WM and LTM (Miyake & Shah, 1999). The three models highlighted in the Introduction are good examples: they vary in how they conceptualize not just the flow or order of information (e.g., WM first or LTM?) but also in the basic storage mechanisms (e.g., separate memory systems or unitary?). Understanding how long-term knowledge influences memory over the short-term is not only a critical concern for WM models, but it is also key in understanding how cognition operates in real-world environments (e.g., language processing, reasoning, driving, typing, etc.). Given that WM encompasses much of our current conscious awareness and facilitates all complex behavior to some extent (Baddeley & Logie, 1999), better understanding the WM system and how it uses prior knowledge addresses the broader question of what influences human performance in complex tasks.

## **Model Overview & Fit With Current Data**

**Gateway Model.** Overall, the current set of results are highly inconsistent with the gateway model because that model does not make any explicit predictions about how familiarity of items in VWM influence memory for those items which was a key prediction in the current set of experiments. In Baddeley's model, WM retrieves information from LTM, but it is not involved in stabilizing representations in WM. Additionally, the gateway model cannot easily account for why there were differences between the stimulus based training in Experiment 2a and the task specific training used in Experiment 2b. While none of the models necessarily predicted this effect there is no mechanism in the gateway model that could account for that result.

The gateway model does fit with the implication from Experiment 1 that executive processes are involved with stabilizing information in WM since the central executive component in the gateway models plays a key role in managing information in each of the domain-specific subsystems. However, both of the other models discussed (the unitary and workspace models) can also account for that effect and they can more easily account for the effects of LTM (see below). Thus, the gateway model provides a poor fit to the current set of data due to the limited relationship between WM and LTM.

**Unitary Model.** Because of the direct relationship between WM and LTM in the unitary model, this model fits the current set of results well. The unitary model fits these experiments in three ways. First, the unitary model would predict that familiar items are faster to consolidate than novel items. Specifically, Cowan has suggested that highly familiar information in LTM (particularly activated LTM) is less attentionally demanding

and will enter the focus of attention via more automatic processes instead of effortful ones (Cowan, 1999). Cowan has also suggested during encoding, items that are familiar are encoded with more detail, leading to better overall memory for those items (Cowan, 1999). Additionally, Cowan has argued for a search mechanism that scans information in the focus of attention to reactivate it during maintenance. Training could facilitate this process, particularly by speeding it up (Cowan, 1999). Thus, there are several ways in which training could facilitate processing of information in WM in the gateway model.

Second, the unitary model can account for the transfer effects seen in Experiment 2b and not in 2a. Cowan (1999) argues that forming novel LTM representations (as would occur during training) is accomplished by creating links between items in the focus of attention and in activated memory. In the current experiments, the change detection training used in Experiment 2b required participants to create these kinds of relationships between items which under the unitary model, could have lead to an improved ability to create these relationships between similar, novel items. In Experiment 2a, however, the focus was not on creating relational links between different items but the relations of features within an item. In both cases, training improved relational processing but in one case the improved processing was specific to stimuli and in another it was specific to relations between items. Interestingly, this mechanism within the unitary model fits nicely with Jiang et al.'s (2000) experiments demonstrating a fundamental role for relational processing in VWM. Cowan's model suggests that the role of relational processing may be broader in scope than just processing visual information but rather

may be a crucial process in learning novel information or novel combinations of familiar information, such as presenting a set of visual items in a new configuration.

Finally, as with the gateway model, the unitary model can easily account for any role of the central executive in stabilizing representations in WM. In Cowan's model, the central executive plays an important role in directing attention to representations within the focus of attention and activated memory as well as to external stimuli (e.g., during encoding). The central executive is resource limited and when completing a complex verbal task and a complex visual task simultaneously that both draw on these resources, performance will drop (Morey & Cowan, 2005). Because of the key role that the executive plays in attending to WM representations, it is highly likely that with fewer available resources, creating stable WM representations would be more difficult and/or time consuming.

Thus, the data fit with the unitary model well and that model can account for several of the key effects demonstrated here. However, there are a couple aspects of the current data set that the unitary model cannot easily account for. First, in the current set of experiments there was evidence for decay of the VWM representations at the longest delay interval (484 ms) with the more complex polygon shapes. Within the unitary model, the shapes being currently maintained would be in the focus of attention but Cowan (1999) argues that there are no decay based processes within the focus of attention (decay occurs in activated memory). Rather, the focus of attention is limited only in capacity. The current set of data suggest that decay is possible in the focus of

attention or at least that the representations became somewhat degraded over the delay interval making them more susceptible to interference from visual masks.

Second, Cowan (2001) proposed that the capacity limit of the focus of attention is  $4 \pm 1$  items, when controlling for any chunking using LTM. In the current set of experiments, the capacity estimates were much lower (2-3 items) for the complex polygon shapes despite the fact that for some of those stimuli there were long-term representations that would allow for chunking. Thus, if anything the current estimates for familiar shapes are inflated. While there was some evidence that capacity increased with familiar shapes (which is inline with what the unitary model would predict), the overall low estimates are inconsistent with the  $4 \pm 1$  view.

There are also some general considerations surrounding the unitary model. For example, one key advantage of the unitary model is that it is based on a wide range of data, making it generalizable to different modalities and tasks. However, it runs the risk of being overly general to the point that the model can account for everything, making it unfalsifiable. Additionally, proposing a unitary memory system cannot easily explain neuropsychological dissociations between LTM and WM (e.g., Shallice & Warrington, 1970). If either system is impaired, then that should lead to a deficit in the other system which is not the case.

Despite these drawbacks, the unitary model can explain the current set of data well. Additionally, the current set of data inform the unitary model, particularly concerning the transfer of information from WM to LTM. That is, the current evidence for a role of relational information during the consolidation process suggests that

processing that information aids in creating new LTM representations, in line with Cowan's (1999) view.

**Workspace model.** The current set of data are highly consistent with the view that there is a direct relationship between perception and LTM, a key feature of the workspace model. Because of this feature of this model, it can easily account for the current set of results: the workspace model would predict that familiar items would be easier to consolidate into WM than novel items and it would predict that the impact of LTM would occur at early information processing stages. More specifically, the workspace model can account for the early facilitation of WM by LTM in two ways: familiar items are more quickly activated in LTM by perception or the activated LTM information is more quickly transferred from activated LTM to WM.

Additional support for the workspace model is found by comparing performance between familiar and unfamiliar (or novel and repeating for Experiment 3) trials with a 17 ms delay, collapsing across all set sizes. Paired samples t-tests showed that in all experiments, except for Experiment 2b in which no differences between trained and novel shapes were found, accuracy for familiar shapes was significantly higher than unfamiliar shapes [Experiment 1:  $t(30) = 7.28$ ; Experiment 2a:  $t(70) = 3.60$ ; Experiment 3:  $t(22) = 2.33$ ]. Thus, in each experiment in which an advantage was demonstrated for familiar items, LTM was facilitating processing as early as 17 ms following the memory array (117-517 ms total, counting the presentation time for the stimuli). In Experiment 2b, both trained and novel shapes demonstrated resistance to masks as early as 17 ms even though

there was no difference between stimulus types. This extremely early effect of LTM is can only be explained by a direct relationship between perception and LTM.

In general, because the workspace model argues that activated LTM representations are what resides in WM, it can account for much of the same effects as the unitary model. However, a major drawback for the workspace model is that it is much less specified than the unitary model, so it cannot as easily account for the transfer effects demonstrated or the differences found between stimulus-based training and task-based training. This does not mean that the workspace model cannot account for these data at all, just that the higher level of specificity in the unitary model allows for it to explain a wider range of data. This ability to generalize further is due to the fact that the unitary model was developed based on data from a variety of tasks while the workspace model was developed to explain a particular set of data (i.e., VWM performance).

Thus, the workspace model can account for the current set of experiments but it does have some drawbacks compared to the unitary model, which can also account for the current data. However, the unitary model cannot as easily account for the early LTM effects since it does not make an explicit prediction about how perceptual input can activate information in LTM. The workspace model can also more easily account for the neurological dissociations between LTM and WM than the unitary model.

**Reconciling the models.** As the above overview highlights, both the unitary and workspace models can account for the current data and they both have pros and cons in terms of explanatory power. One important thing to note, however, is that the two models are not all that different; they both assume WM consists of activated portions of LTM but

the workspace model proposes a multi-component system while the unitary model does not. So in many ways, the preferred model will vary by researcher depending on how they view the WM system as a whole. Because of this fundamental difference between the unitary and multiple-component views, it is difficult to reconcile them based on the current data alone, since the current data do not address how many WM components there are. However, the current data do offer some suggestions as to how these models may be reconciled.

It is clear that while the workspace model has the potential to account for both the LTM effects demonstrated here and the neuropsychological evidence showing dissociations between WM and LTM the model needs more specification as to what controls the activation of information in LTM and how that information is subsequently processed in WM. Given the strengths of the unitary model, in explaining those shortcomings, it makes sense to apply the control processes in Cowan's model to Logie's workspace model. Essentially, this would mean using Cowan's model but assuming that that the system has separate WM and LTM stores (i.e., pull the focus of attention out of activated LTM). Additionally, some sort of decay based limit would be required to account for the drop in performance at 484 ms with more complex shapes.

The current set of experiments cannot dissociate between the Workspace Model and the Unitary Model. While the unitary model can account for most of the effects demonstrated in this experiment, the workspace model can as well and it can more easily account for neuropsychological data. Combining the two models may make sense, but given the difference between a unitary system and a multiple component system, the

current data are not well suited to do so. However, overall the current data provide strong evidence that there is a highly interactive relationship between WM and LTM and that LTM aids in WM processing very early.

### **An Alternative View: Cue-based Models of WM**

An alternative view to the more traditional, structural models of working memory detailed above are cue-based models of immediate memory (e.g., Nairne's Feature Model; Nairne, 1990; Nairne, 2002). Cue-based models are unitary models of memory that emphasize retrieval cues instead of activated information within immediate memory. These models assume similar processes exist for short-term and long-term memory (they are both cue-driven) and no distinctions are made based on modality of information (e.g., verbal versus visual memory systems; Nairne, 2002). Nairne's feature model provides a good example of a cue-based approach to immediate or primary memory. The feature model makes a distinction between primary and secondary memory in which primary memory consists of currently activated memory traces that are susceptible to interference and secondary memory is a more permanent store (Nairne, 1990), similar to Cowan's embedded processes model. This distinction is not structural, however, and merely makes a distinction between active memory traces and long-term representations.

Items within primary or immediate memory are not stored as activated information (as the standard models argue), but are represented as vectors containing features of that item that serve as cues for that item (Nairne, 1990). These vectors of cues within primary memory are then matched with groups of relevant traces in secondary memory (i.e., a secondary memory search set) based on similarity (as measured by

feature overlap between vectors) and then information from the search set is selected based on trace distinctiveness (Nairne, 1990). Thus, successful retrieval will depend on how distinctive the retrieval cues (or features) are in delimiting the search set to dissociate target memory items from other similar memory traces in secondary memory (Nairne, 1990, 2002).

Cue-based memory models offer advantages over structural memory models by being able to account for the impact of retrieval cues and long-term memory effects on memory. However, these models are very limited for two key reasons. First, cue-based models are less able to account for neurological data showing dissociations in brain damaged patients and in patterns of neural activity. Second, cue-based models are almost exclusively verbal in nature with no work currently addressing how such cues would operate with a visual task. Thus, the current data supporting cue-based model are based on verbal memory in neurologically intact participants, severely limiting the generalizability of these models.

Despite these drawbacks, however, a cue based approach could account for some of the data reported here by arguing that the trained items developed more rich feature vectors (though given the lack of work in the visual domain for cue-based models it is unclear what those feature vectors would include), providing them with more retrieval cues. Additionally, in Experiment 3, the cue-based model may be able to argue that the repeating arrays were better remembered due to the repeating context of the items, thereby serving as a retrieval cue itself. However, the cue-based view is problematic for a couple of reasons. First, in Experiments 2a, 2b, and 3 the features of the shapes are very

similar: they are all black images with a certain number of points and some angles. Given the overlap in features, the cue-based model should predict an overall benefit of training for all shapes (i.e., not just trained shapes) since they share features, but this was only evident in Experiment 2b and that transfer can be explained by task specific strategies. Second, the cue-based model makes no specific predictions about the time course of consolidation and how that would vary depending on the level of familiarity with the stimuli, making the model ill-suited for understanding how visual long-term representations contribute to stabilizing perceptual traces.

### **Summary**

Overall, the results in these experiments demonstrate very close relationship between WM and LTM and they suggest that LTM influences processing in VWM at very early stages, as early as 17 ms following the presentation of a visual memory array. This finding is both unique and meaningful in that it informs current models of WM in how they account for the relationship between WM and LTM. Overall, the more recent Workspace Model most easily fits the current set of results, however the results are also consistent with the Unitary Model. Given the strengths and weaknesses for each of the models, combining the two models may best account for the current set of results.

### **What does this tell us about short-term consolidation?**

#### **What is being consolidated?**

The process of short-term consolidation is a crucial process that allows us to retain information past a few hundred milliseconds. While prior research demonstrated that consolidation was a key link between perception and WM (e.g., Jolicoeur &

Dell'Acqua, 1998; Vogel et al., 2006) as well as between perception and visual attention (e.g., Chun & Potter, 1995; Nieuwenstein et al., 2005; Shapiro et al., 1997), the current research demonstrates that LTM influences the process of consolidation between perception and WM, with familiar items being consolidated more quickly than unfamiliar items. Taken together, the research examining consolidation shows that this process lies at the intersection of several key cognitive processes: perception, attention, WM, and LTM.

In addition to demonstrating a clear role for long-term representations in facilitating consolidation, the current set of experiments can also address a question that has been largely ignored in the consolidation literature: what exactly is being consolidated? The set of results across the three experiments address this question. First, the advantage for the simpler square stimuli used in Experiment 1 over the complex polygons indicates that as visual complexity increases, more time is required for consolidation. Stated differently, consolidation is sensitive to visual features (such as color and shape) which suggests that the visual details of items are what is consolidated between perception and VWM.

The positive training result in Experiment 2a using a 4AFC training task supports this visual details idea. In particular, the 4AFC training task in Experiment 2a emphasized distinguishing visual details between a set of similar shapes which in turn led to faster consolidation. Demonstrating evidence that visual features are what is being consolidated into VWM is in-line with the prior work on short-term consolidation, particularly the attentional blink literature (e.g., Chun & Potter, 1995; Nieuwenstein et al., 2005; Shapiro

et al., 1997) and Vogel et al.'s (2006) conclusion that visual perceptual features are easily overwritten (i.e., by visual masks). Additionally, replicating the set size effects in the current set of experiments is largely consistent with the view that consolidation is resource limited, with more time needed as the number of items increases.

However, one interesting set size effect that has yet to be addressed in the literature has to do with performance at set size 1. Both in the current set of experiments and in the previous research (specifically Vogel et al., 2006), there was no evidence for consolidation with one item (i.e., performance was not worse for early delays compared to longer delays) but there was for all other set sizes. If consolidation is only strengthening visual information, then even with only a single visual item visual masks should interrupt consolidation at least at the earliest delay.

There are two possible explanations for this effect. One is that it is possible that with only a single item, consolidation could occur while the item was still visible (i.e., during the 100-500 ms memory array presentation). Thus, the masks would not interfere at any delay because the information was consolidated prior to mask presentation. However, a more theoretically interesting possibility comes from the data in Experiment 3, specifically how the relationships between items (called a global spatial configuration) in a visual array are processed.

In Experiment 3, the repeating shapes were presented in the same spatial locations, thus maintaining the global configuration of items on each repeating trial which, as noted above, likely led to the advantage for repeating shapes over novel shapes. This result suggests that the notion of visual characteristics being what is consolidated

into VWM may not be the whole story: the consolidation process may also be strengthening the spatial relations between the items, thereby improving retention not just for the visual features of items but also their spatial properties. Thus, with only a single item, there is no relational information to consolidate, making it relatively impervious to visual masks. For each set size following one, however, there is at least one spatial relationship, leading to longer consolidation times.

Prior research has demonstrated that maintaining a global spatial configuration is a fundamental component of VWM (Blalock & Clegg, 2010; Jiang et al., 2000), but little work has examined how this configuration is created. The current set of experiments suggests that the formation of this configuration could occur very early as information is consolidated into VWM. This suggests that short-term consolidation is multi-dimensional, strengthening all aspects of perceptual traces from specific visual details of items to the spatial properties of those items. This conclusion is also supported by the transfer effects demonstrated in Experiment 2b. Using a training task that required spatial relational processing in addition to visual knowledge of the shapes, novel items did not show any impact of visual masks compared to the other experiments.

### **How is information being consolidated?**

Beyond what is being consolidated during short-term consolidation is the equally important question of how information is consolidated into VWM. Vogel et al. (2006) outlined a general framework for how the consolidation process could work that is based on their visual attention model of VWM (Vogel, et al., 2001). In their framework, they argue that when visual items are presented, those items are represented at multiple levels

and in multiple visual areas. For example, the basic shape information would be represented in the primary visual cortex whereas higher level object identification would be represented in the extrastriate areas. These representations will quickly decay without a consolidation process that selects the appropriate information and then stabilizes those representations so that they can be maintained for a longer period of time and survive new incoming stimuli.

Vogel et al. (2006) proposed that once this process occurs, the visual information is stored in WM, though they are careful to note that the information is not literally moving to a new place. Instead, they suggest that, “the consolidation process may change the nature of existing representations” (Vogel et al., 2006, 1448). In their previous work (Vogel et al., 2001), they have argued that items in VWM are maintained via synchronous and asynchronous neural firing mechanisms; neurons tuned to the shared visual features (including location) of an item fire together thereby binding that information together (synchronous firing) and that neural firing rate is slightly different from features associated with another item (asynchronous firing). The two (or more) items are maintained separately by the asynchronous firing between items, but as more items are maintained, there is a greater likelihood that firing between items will accidentally overlap, leading to memory errors. Thus, the process of consolidation could essentially be stabilizing these neural firing mechanisms.

The current data are largely in line with this framework, though they do raise some issues about how that framework generalizes to more complex shapes and familiar shapes. In Experiment 1, the more complex polygons took significantly longer to

consolidate than the simple colored squares. This difference could be due to long-term representations of the colored squares, but it could also be due to increased complexity of the polygon shapes. Under the framework outlined above, this suggests either that complex shapes take longer to bind together or that due to the increased similarity of the shapes, there is greater overlap in the neural firings for each shape making them harder to distinguish. While Vogel et al.'s (2001; 2006) framework can account for this effect, future work should address which of these possibilities leads to both slower consolidation and lower overall performance.

This synchronous firing framework has a more difficult time accounting for the familiarity and transfer effects demonstrated in the current set of experiments. In the current framework, there is no mechanism that would allow for familiar items (or novel items in the case of Experiment 2b) to be consolidated more quickly than novel items at such an early stage. To account for this, some type of top-down mechanism would need to be added that could direct attention from the long-term representations in LTM to the early visual processing stages. Alternatively, there could be some type of bottom-up perceptual filter that is tuned to recently presented visual items that reside in LTM that facilitates processing for those items.

Recent work by Moore, Cohen, and Ranganath (2006) suggests that both top-down and bottom-up mechanisms are responsible for facilitative effects of visual expertise. Using functional magnetic resonance imaging (fMRI), they examined how visual expertise with a set of complex figures correlated with brain activity in the occipitotemporal cortex (which represents objects information) as well as in the

prefrontal and parietal networks that have been shown to support VWM encoding and maintenance (specifically the dorsolateral prefrontal cortex and the intraparietal sulcus; see Olesen, Westerberg, & Klingberg, 2004; Pessoa, Gutierrez, Bandettini, & Ungerleider, 2002; Todd & Marois, 2004; Xu & Chun, 2006).

Moore et al. (2006) showed that training of visual stimuli led to domain-specific improvements in encoding and maintenance. Specifically, they showed expertise effects in the dorsolateral prefrontal cortex, the intraparietal sulcus, and the occipitotemporal cortex. They argue for a “cortical division of labor” (Moore et al., 2006, p. 11194) in which visual expertise leads to stimulus-specific tuning of neurons in the occipitotemporal cortex that project onto dorsolateral prefrontal neurons (i.e., a bottom-up effect) and that information in the prefrontal area leads to the formation of a frontoparietal network that “represents category information in a manner that can guide visual attention, WM encoding, and maintenance” (i.e., a top-down effect; p. 11195). This conclusion closely mirrors Kosslyn’s model of mental imagery (i.e., the process of seeing with the mind’s eye without perceptual input), which also argues for top-down and bottom-up processes in retrieving and constructing mental images (Kosslyn, 2005; Kosslyn, et al., 2006).

Moore et al. (2006) also showed evidence that expertise led to increased intraparietal sulcus activation and argue that this supports previous work indicating that area is responsible for encoding more visual detail of items (e.g., Xu & Chun, 2006). While this does not directly address the issue of consolidation, this work suggests that trained items are facilitated in VWM through both top-down attentional mechanisms and

bottom-up tuned neural networks and that those items will be encoded with more visual detail compared to novel items.

The evidence for top-down mechanisms facilitating visual expertise support the speculation following Experiment 1 that there is role for central executive processes in short-term consolidation. Additionally, this view fits with the implications of Experiment 2b that top-down strategies from change detection training facilitate VWM performance. This has implications for the WM models outlined above: how does the central executive influence early visual processing and allow for trained items to be consolidated into WM faster? There are a few possible explanations for this. One is that the central executive can limit the information that becomes activated during encoding (similar to Cowan's Embedded Processes Model). Another possibility is that with increased expertise, the central executive is better able to resist the build-up of proactive interference (when previously learned information interferes with learning new information) across trials, a view in line with work by Engle and colleagues (e.g., Engle & Kane, 2004; Kane, et al., 2001; Kane & Engle, 2000).

Finally, as expertise is acquired the central executive may direct visual attention to identifying features of items, allowing for faster object identification and consolidation. This final possibility is consistent with the increased intraparietal sulcus activity shown in Moore et al. (2006). One direction for future research could examine the role of executive processes in consolidation by examining how consolidation varies across individuals; if participants show higher working memory capacity (i.e., capacity of the central executive), then consolidation should be faster for those high span individuals.

## **Summary**

The current set of results demonstrate that the process of short-term consolidation is multi-dimensional and lies at a critical intersection of perception, WM, and LTM. These data are also consistent with a combined top-down and bottom-up view that suggests that consolidation is facilitated by tuned neural networks both in late visual areas and in executive function areas.

Two key questions that need to be addressed in the short-term consolidation literature that were not discussed here is clearly distinguishing consolidation from encoding and identifying consolidation processes in other domains. Vogel et al. (2006) do note that it is possible that consolidation only occurs in the face of interfering information (meaning encoding could operate with or without consolidation). Though again this does not explain the lack of any mask effects with a single item and given the high levels of visual interference in real-world visual tasks it does not make much practical sense to dissociate encoding and consolidation in this way. As far as consolidation operating in other sensory domains, it is certainly likely that it does. Currently, the visual system is arguably the most well understood of all the sensory processing systems, making the visual domain a natural place to begin understanding the relationship between perception and WM, but this work should be extended to other domains.

## **Concluding Remarks**

The processes of perception, WM, and LTM have historically been examined independently of one another with little discussion of how these constructs interact to facilitate complex cognition. Given the large amount of research conducted in each of

their respective domains, it is now possible to start to examine how these processes work together. The current set of experiments is a start in examining interactions between perception, WM, and LTM. The results lead to four key conclusions (also see Table 1): (1) there is a direct relationship between perception and LTM; (2) training improves the speed in which information gets into WM; (3) training a set of visual items can lead to improvements in VWM performance even for novel items; and (4) that short-term consolidation is a multidimensional process that strengthens both visual and spatial relational information to create more stable representations of perceptual input. Taken together, these conclusions provide strong evidence that WM, LTM, and perception are highly interactive and provide unique insights as to how these processes work together to facilitate everyday, complex cognition.

## CHAPTER 7 - REFERENCES

- Alvarez, G., & Cavanagh, P. (2004). The capacity of visual short-term memory is set by visual information load and by number of objects. *Psychological Science, 15*(2), 106-111.
- Atkinson, R. C., & Shiffrin, R. M. (1971). The Control of Short-term Memory. *Scientific American, 225*(2), 82-90.
- Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology, 53*, 221-227.
- Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin, 53*, 452-471.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*, 622-628.
- Baddeley, A. (1986). *Working Memory*. Oxford, England: Clarendon.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829-839. doi: 10.1038/nrn1201
- Baddeley, A. (2007). *Working Memory, Thought, and Action*. Oxford, England: Oxford University Press.

- Baddeley, A., & Hitch, G. (1974). Working Memory. *The Psychology of Learning and Motivation*, 47-89.
- Baddeley, A., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28-61). New York, NY: Cambridge University Press.
- Barnett, S., & Ceci, S. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, 128, 612-637.
- Beschin, N., Cocchini, G., Della Sala, S., & Logie, R. H. (1997). What the eyes perceive, the brain ignores: A case of pure unilateral representational neglect. *Cortex*, 33, 3-26.
- Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology*, 33A, 415-437.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.
- Blalock, L. D., & Clegg, B. A. (2010). Encoding and representation of simultaneous and sequential arrays in visuospatial working memory. *The Quarterly Journal of Experimental Psychology*, 63, 856-862.
- Brady, T. F., Konkle, T., Alvarez, G., & Olivia, A. (2008). Visual long-term memory has a massive storage capacity for objects. *Proceedings of the National Academy of Sciences*, 105, 14325.

- Buschkuhl, M., Jaeggi, S. M., Hutchinson, S., Perrig-Chiello, P., Däpp, C., Müller, M., et al. (2008). Impact of working memory training on memory performance in old-old adults. *Psychology and Aging, 23*, 743-753.
- Buttle, H., & Raymond, J. E. (2003). High familiarity enhances visual change detection for face stimuli. *Perception & Psychophysics, 65*(8), 1296-1306.
- Chen, D., Eng, H., & Jiang, Y. (2006). Visual working memory for trained and novel polygons. *Visual Cognition, 14*(1), 37-54.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28-71.
- Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10*, 360-365.
- Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 224-234.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 109-127.
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*(2), 163-191.
- Cowan, N. (1993). Activation, attention, and short-term memory. *Memory & Cognition, 21*, 162-167.

- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). New York, NY: Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*.
- Curby, K. M., & Gauthier, I. (2009). The temporal advantage for individuating objects of expertise: Perceptual expertise is an early riser. *Journal of Vision*, 9, 1-13.
- Dell'Acqua, R., & Jolicoeur, P. (2000). Visual encoding of patterns is subject of dual-task interference. *Memory & Cognition*, 28, 184-191.
- Della Sala, S., & Logie, R. H. (2002). Neuropsychological impairments of visual and spatial working memory. In A. Baddeley (Ed.), *The Handbook of Memory Disorders* (2nd ed., pp. 271-292). Chichester, UK: John Wiley & Sons.
- Di Lollo, V., Kawahara, J., Ghorashi, S. M. S., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research*, 69, 191-200.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory, capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 44, pp. 145-199). NY: Elsevier.
- Engle, R. W., Tuholski, S., Laughlin, J., & Conway, A. (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology: General*, 128(3), 309-331.

- Gao, Z., Li, J., Liang, J., Chen, H., Yin, J., & Shen, M. (2008). Storing fine detailed information in visual working memory - Evidence from event-related potentials. *Journal of Vision, 9*, 1-12.
- Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 845-866.
- Hebb, D. O. (1949). *Organization of Behavior: A Neuropsychological Theory*. New York, NY: Wiley, Inc.
- Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain Mechanisms and Learning* (pp. 37-46). Oxford: Blackwell.
- Hollingworth, A. (2004). Constructing visual representations of natural scenes: the roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 519-537. doi: 10.1037/0096-1523.30.3.519
- Hollingworth, A. (2008). Visual memory for scenes. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 123-161). New York, NY: Oxford University Press.
- Hollingworth, A., & Luck, S. J. (2009). The role of visual working memory (VWM) in the control of gaze during visual search. *Attention, Perception & Psychophysics, 71*(4), 936-949. doi: 10.3758/APP.71.4.936
- Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and

- gaze correction. *Journal of Experimental Psychology: General*, 137(1), 163-181.  
doi: 10.1037/0096-3445.137.1.163
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30, 685-701.
- Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 556-568.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, 105, 6829-6833.
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of experimental psychology Learning, memory, and cognition*, 26(3), 683-702.
- Jiang, Y., Song, J.-H., & Rigas, A. (2005). High-capacity spatial context memory. *Psychonomic Bulletin & Review*, 12, 524-529.
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36(2), 138-202. doi: 10.1006/cogp.1998.0684
- Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of WM capacity. *Journal of Experimental Psychology: General*, 130, 169-183.

- Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 336-358.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189-217. doi: 10.1037/0096-3445.133.2.189
- Kosslyn, S. (2005). Mental images and the Brain. *Cognitive Neuropsychology*.
- Kosslyn, S., Ganis, G., & Thompson, W. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*.
- Kosslyn, S., Thompson, W., & Ganis, G. (2006). *The Case for Mental Imagery*. New York, NY: Oxford University Press.
- Logie, R. H. (2003). Spatial and visual working memory: A mental workspace. In D. E. Irwin (Ed.), *Cognitive Vision: The Psychology of Learning and Motivation* (Vol. 42, pp. 37-78). San Diego, CA: Academic Press.
- Logie, R. H., Della Sala, S., Beschin, N., & Denis, M. (2005). Dissociating mental transformations and visuo-spatial storage in working memory: evidence from representational neglect. *Memory*, *13*(3-4), 430-434.
- Logie, R. H., & van der Meulen, M. (2009). Fragmenting and integrating visuospatial working memory. In J. R. Brockmole (Ed.), *The Visual World in Memory* (pp. 1-32). New York, NY: Psychology Press.

- Luck, S. J. (2008). Visual short-term memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 43-85). New York, NY: Oxford University Press.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281. doi: 10.1038/36846
- Luria, R., Sessa, P., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. (2009). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience*, *22*, 496-512.
- Makovski, T., & Jiang, Y. (2008). Indirect assessment of visual working memory for simple and complex objects. *Memory & Cognition*, *36*(6), 1132-1143. doi: 10.3758/MC.36.6.1132
- Miyake, A., & Shah, P. (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. New York, NY: Cambridge University Press.
- Moore, C. D., Cohen, M. X., & C, R. (2006). Neural mechanisms of expert skills in visual working memory. *The Journal of Neuroscience*, *2006*, 11187-11196.
- Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 703-713.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251-269.
- Nairne, J. S. (2002). Remembering over the short term: The case against the standard model. *Annual Review of Psychology*, *53*, 53-81.

- Nieuwenstein, M. R., Chun, M. M., van der Lubbe, R. H. J., & Hooge, I. T. C. (2005). Delayed attentional engagement in the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1463-1475.
- Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training working memory. *Nature Neuroscience*, *7*, 75-79.
- Olson, I. R., & Jiang, Y. (2004). Visual short-term memory is not improved by training. *Memory & Cognition*, *32*(8), 1326-1332.
- Olson, I. R., Jiang, Y., & Moore, K. S. (2005). Associative learning improves visual working memory performance. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 889-900.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369-378.
- Pessig, J., & Tarr, M. (2007). Visual object recognition: Do we know more now than we did 20 years ago? *Annual Review of Psychology*, *58*, 75-96.
- Pessoa, L., Gutierrez, E., Bandettini, P., & Ungerleider, L. (2002). Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron*, *35*, 975-987.
- Raymond, J. E., Shapiro, K., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849-860.
- Rensink, R. A. (2002). Change Detection. *Annual Review of Psychology*, *53*, 245-277.

- Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly Journal of Experimental Psychology*, 22, 261-273.
- Shapiro, K., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink: A view on attention and a glimpse on consciousness. *Trends in Cognitive Sciences*, 1, 291-296.
- Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, 7, 1-15.
- Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, Present, and Future. *Trends in Cognitive Sciences*, 9, 16-20.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, 1-29.
- Stadler, M. A. (1993). Implicit serial learning: Questions inspired by Hebb (1961). *Memory & Cognition*, 21, 819-827.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428, 751-754.
- Tulving, E. (2007). Are there 256 different kinds of memory? In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger III* (pp. 39-52). New York: Psychology Press.
- van der Meulen, M., Logie, R. H., & Della Sala, S. (2009). Selective interference with image retention and generation: evidence for the workspace model. *Quarterly Journal of Experimental Psychology*, 62, 1568-1580.

- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27(1), 92-114.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 32(6), 1436-1451. doi: 10.1037/0096-1523.32.6.1436
- Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 25, 1256-1271.
- Ward, R., Duncan, J. L., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, 30, 79-109.
- Wickens, C., Vincow, M., & Yeh, M. (2005). Design Applications of Visual Spatial Thinking: The Importance of Frame of Reference. In A. Miyake & P. Shah (Eds.), *The Cambridge Handbook of Visuospatial Thinking* (pp. 383-425).
- Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review*, 15, 763-771.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440, 91-95.