Stimulus familiarity improves consolidation of visual working memory representations

Lisa Durrance Blalock

Published online: 27 February 2015
© The Psychonomic Society, Inc. 2015

Abstract Short-term consolidation is a process that stabilizes visual working memory (VWM) representations so that they are less susceptible to interference. The current study examined this process, specifically if training on specific shapes facilitated the consolidation of visual representations in working memory. Three experiments using two different training tasks compared performance between trained and novel stimuli using the backward masking paradigm. Experiment 1 used a four alternative forced choice task and found an overall advantage for trained shapes as well as evidence for faster consolidation for trained shapes and this cannot be explained by verbal labeling of the trained items (Experiment 3). Experiment 2 used a change-detection training task and showed no overall benefit of training but did show evidence of transfer of training to novel shapes. Taken together, these results show that long-term visual representations can facilitate VWM processing, but the type of training task used impacts the degree to which the long-term representations will affect VWM.

Keywords Visual working memory · Training · Short-term consolidation · Visual long-term memory · Encoding

When perceiving information in the visual world, sensory information leads to perceptual representations in the brain

Portions of this paper are from a dissertation submitted to the Academic Faculty of Colorado State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

L. D. Blalock (✉)
Department of Psychology, University of West Florida,
11000 University Parkway, Pensacola, FL 32514, USA
e-mail: lblalock@uwf.edu

that are fragile (Sperling, 1960). In order to retain visual information over longer periods of time some mechanism must select some of these visual representations and stabilize them so they can (1) persist after the visual information is no longer available in sensory memory, and (2) 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 visual working memory (VWM; Gegenfurtner & Sperling, 1993). This process of turning a transient perceptual representation into a more durable and stable VWM representation has been termed short-term consolidation (Jolicoeur & Dell'Acqua, 1998; Vogel, Woodman, & Luck, 2006). The primary focus of the current study was to examine how visual long-term memory (LTM) representations influence VWM processing, specifically the consolidation process.

Consolidation in VWM

Much of the work on short-term consolidation comes from the attentional blink literature. In a typical attentional blink task, a stream of stimuli is presented rapidly and participants respond to two target stimuli within that stream (also called a rapid serial visual presentation, or RSVP task). For example, a stream of numbers will be presented rapidly (~ten 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 (see Dux & Marois, 2010; Martens & Wyble, 2010 for recent reviews). 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, Shapiro, & Arnell, 1992). Thus, when attention is focused on processing the first target, the ability to not only report but also detect a second target is greatly

diminished, showing a temporary lapse, or blink, in attentional resources. A leading theory explaining this attentional blink argues that the second target cannot be processed while the first target is still being consolidated (Chun & Potter, 1995; Nieuwenstein, Chun, van der Lubbe, & Hooge, 2005; Shapiro, Arnell, & Raymond, 1997; Vogel et al., 2006).

Vogel and colleagues (2006) further examined this process by determining the time course of consolidation in VWM (also see Blalock, 2013). 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. These mask arrays should impair change detection performance only if presented while information is still being consolidated into VWM (i.e., at the 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 found 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 suggests the consolidation process is limited in capacity, requiring more time as the number of 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 relatively simple items, about 200–300 ms is required to create a stable representation in VWM. Gegenfurtner and Sperling (1993) also examined the time course of consolidation and found a considerably faster rate of 20–30 ms per item. However, 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 (Vogel et al., 2006). This idea of how familiarity with the stimuli affects consolidation was a primary focus of the current set of experiments.

Interactions between VWM and LTM

While there are many theories to describe the relationship between working memory (WM) and LTM (e.g., Baddeley, 2012; Baddeley & Logie, 1999; Cowan, 1999, 2001), most focus on the verbal domain (though see Logie, 2003; Logie & van der Meulen, 2009 for an exception). Luck (2008) notes that little work has focused on how visual representations in VWM and in visual LTM interact. The small amount of research specifically 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 and comparing already known stimuli with novel stimuli. The general hypothesis for

all of this research has been that if VWM consists of activated LTM representations, performance and/or capacity should increase. Thus far, the results are mixed.

Explicit training of stimuli

Currently, only one study has used an explicit training task in VWM (Chen, Eng, & Jiang, 2006). In Chen et al. (2006) participants trained on a subset of eight polygons by completing 320 change detection trials that used the same eight trained shapes. Participants then completed a change detection task again with both trained and novel polygons. 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 familiarity of stimuli may play a role only in 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, it remains an open question if training can facilitate early VWM processes such as consolidation.

Inherently familiar versus unfamiliar stimuli

Unlike the work of 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, work on face recognition has supported the notion that familiarity with stimuli leads to improvements in VWM. Buttle and Raymond (2003) compared change detection between famous and non-famous faces and found that 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 led to enhanced VWM performance.

In a similar experiment, Curby and Gauthier (2007) compared change detection performance for upright versus inverted faces. They used a backward masking paradigm similar to that of Vogel et al. (2006) to examine the time course of perceptual expertise. 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 supports the idea that experience can influence early visual processing. However, given that the stimuli used in this research were highly familiar and were not explicitly trained,

it is still unclear how explicitly trained stimuli that were novel before training can impact VWM processing.

Proactive interference in VWM

Long-term representations may also be formed over the course of an experiment, particularly when the same stimuli are used across trials. Often, these representations lead to interference, called proactive interference, in later trials. Several recent studies have examined the role proactive interference plays in commonly used VWM tasks with mixed results. For example, Lin and Luck (2012) examined proactive interference in a change detection task by manipulating the release from interference after a switch in stimuli (e.g., switching from color to orientation stimuli) and by manipulating the similarity of the stimuli. While both of those methods have demonstrated substantial proactive interference effects using verbal stimuli (e.g., Keppel & Underwood, 1962; Wickens, Born, & Allen, 1963), Lin and Luck (2012) showed no evidence of proactive interference in change detection tasks. Their results suggest that long-term representations of the stimuli are not formed over the course of a VWM task and have little impact on VWM capacity estimates.

In contrast to this, however, other research has shown both moderate (Hartshorne, 2008; Makovski & Jiang, 2008) and substantial proactive interference effects (Endress and Potter, 2014) in VWM tasks. Endress and Potter (2014) argued that the use of longer encoding intervals in previous research allowed visual items to consolidate into LTM representations, thereby underestimating the influence of proactive interference on VWM performance. In their study, participants viewed varied-length sequences of real-world objects presented at a fast speed (250 ms). At the end of each sequence participants saw a probe object and determined if that object was shown in the preceding sequence. When new items were used on each trial (low proactive interference), accuracy remained high even with long sequences (up to 100 items) reflecting a large capacity. When the same objects were used across trials (high proactive interference), however, Endress and Potter (2014) showed significantly lower performance indicating a much lower capacity (three to four items). They proposed a framework that includes a passive, temporary unconsolidated visual LTM store that is virtually without limit so long as proactive interference is not present and the stimuli are meaningful (allowing for categorization at encoding). This temporary store is separate from the WM store, which is used if active maintenance or additional processing is required; however, only three or four items can be manipulated at one time in WM. Overall, they argue for a close relationship between LTM and VWM but their study did not address how they might interact (e.g., whether or not the LTM store can assist the more limited VWM store).

Current study

The present study addresses three questions that are currently ambiguous in the literature. First, how can we train long-term visual representations? Second, how and when do long-term representations interact with VWM? Finally, can the limited capacity of VWM be improved with the help of long-term visual representations? Answering these questions is important for developing comprehensive theories and models of VWM and how it interacts with other cognitive constructs, specifically LTM. Given the mixed results in the current literature, this is an important next step towards understanding the interactions between VWM and LTM. Additionally, it provides insight as to how VWM operates in real-world contexts in which prior knowledge influences behavior (e.g., Hollingworth 2004).

Specifically, these experiments 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, 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 consolidated into 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 different mask SOAs for trained and novel shapes. Individuals were trained on a set of abstract polygons and performance was compared between those trained shapes and novel shapes in the backward masking task. Two different training methods are compared: Experiments 1 and 3 used a four alternative forced choice task (4AFC) and Experiment 2 used a change detection task. If visual long-term representations can influence early VWM processing, then trained shapes should consolidate faster; that is, they should be less susceptible to interference from visual masks. If the predicted results are found, it would support a direct relationship between LTM and VWM at very early processing stages.

Experiment 1

Method

Participants and design

One hundred two introductory psychology students participated for optional, partial course credit. Participants were required to have normal or corrected vision. A 2 (stimulus type: trained vs. novel) $\times$ 4 (set size: 1, 2, 3, & 4 items) $\times$ 5 (SOA:

517, 634, 750, 867, & 984 ms) within-subjects design was used.

Stimuli

Attneave shapes (Attneave, 1957; Attneave & Arnoult, 1956) were used as stimuli (see Fig. 1). The shapes had four to seven sides with an approximate visual angle ranging from 1.0° to 2.0° . The shapes were selected randomly from a set of 400 total shapes. The visual masks consisted of 4×4 matrices with half of the cells filled in black and half filled in white (see Fig. 1) with a visual angle of $1.3^\circ \times 1.3^\circ$. The items were assigned randomly to 1 of 16 positions on an invisible 4×4 grid that subtending $9.80^\circ \times 7.30^\circ$ of visual angle, with an approximate minimum separation of 2.0° depending on the size of the stimulus.

Procedure

Experiment 1 involved three main tasks that were completed in a single session lasting approximately 90 min: a training task, a recognition task, and the consolidation task. Participants read through on-screen instructions for each task that included example stimuli. No instructions were provided on any of the tasks on whether to prioritize speed or accuracy.

Training Twelve polygons were selected randomly for training and used for all participants. Participants performed a 4AFC task during the training session. In this task, 1 of the 12 polygons was presented for 1 s in the center of 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. The location of the shapes and numbers were equal to the second and third rows of the invisible 4×4 grid used in the consolidation task. Participants responded by pressing the number that corresponded to the shape 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 1 and 2), and participants had to select which of the two shapes they studied previously in the experiment. The locations of the shapes and labels were the center four locations from the 4×4 grid used in the consolidation task. The 12 distractor shapes used in this task were not used in any other portion of the experiment.

Consolidation task Trained and novel shapes were presented in separate blocks of 200 trials each (four blocks total in the experiment, for a total of 800 trials, 400 trained and 400 novel) and the order of blocks was randomized across participants. This resulted in 20 trials per stimulus type $\times$ SOA $\times$ set size combination. For trained trials, all of the shapes presented were from the set of 12 that were trained earlier in the session. In the novel condition, the shapes were completely novel and not presented in any other portion of the study. The novel shapes were selected randomly from the 400 total shapes, excluding the trained shapes and the distractor shapes used in the recognition test. Thus, on each novel trial all the stimuli were novel and were not repeated. To avoid any comparison errors in comparing the memory array with the test array for the relatively homogenous polygon 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.

Each trial began with a fixation cross for 500 ms. The memory array was then presented for 500 ms, which was followed by a blank screen until the mask array onset. The SOA varied randomly between 517 ms, 634 ms, 750 ms, 867 ms, and 984 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, which

Fig. 1 Trial procedure for the backward masking task with a set size of 4. Cue in test array is presented in magenta. Not to scale

remained on the screen until a response was made or 2,000 ms passed, was presented. 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 shapes). Participants responded same (mapped to the “z” key on a standard keyboard) if the cued item did not change between study and test or they responded different (mapped to the “/” key on a standard keyboard) if they did. There was a 500-ms blank interval after the participant response before the next fixation cross (1,000 ms total after response before the next memory array was presented including the fixation cross).

Final Recognition Task To determine whether or not the long-term representations of the trained shapes changed during the consolidation task, 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, 14 were dropped due to low performance (overall accuracy 2 standard deviations below the mean), and 17 were dropped because they did not finish all four phases in the allotted 1.5-h 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 and 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 [$F < 1$], indicating that the training task was relatively easy.

Overall recognition accuracy was moderately high for both the intermediate recognition test ($M = .85$, $SD = .31$) and the final recognition test ($M = .83$, $SD = .28$). 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 during training and these representations were robust since there was no significant decline across the experiment.

Change detection accuracy

Accuracy was higher overall for the trained items ($M = .83$, $SD = .07$) compared to the novel items ($M = .81$, $SD = .07$), and this difference was especially apparent at larger set sizes (see Fig. 2).¹ Accuracy also dropped significantly as set size increased (set size 1: $M = .92$, $SD = .01$; set size 2: $M = .88$, $SD = .02$; set size 3: $M = .78$, $SD = .01$; set size 4: $M = .71$, $SD = .02$). Additionally, the visual masks led to more disruption at earlier SOAs with novel shapes than trained shapes, suggesting that trained shapes were consolidated faster than novel items. This pattern of results was confirmed by a 2 (stimulus type: trained vs. novel) $\times$ 4 (set size: 1, 2, 3, & 4 items) $\times$ 5 (SOA: 517, 634, 750, 867, & 984 ms) repeated measures ANOVA. 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 SOA [$F(4, 280) = 8.19$, $MSE = .01$, $\eta_p^2 = .11$]. Additionally, there was a significant stimulus type $\times$ SOA interaction [$F(4, 280) = 3.38$, $MSE = .01$, $\eta_p^2 = .05$], a marginally significant set size $\times$ SOA interaction [$F(11.02, 771.13) = 1.68$, $P = .07$, $MSE = .01$, $\eta_p^2 = .02$], and a significant stimulus type $\times$ set size $\times$ SOA interaction [$F(10.44, 730.72) = 1.95$, $MSE = .01$, $\eta_p^2 = .03$]. The stimulus type $\times$ 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; visual masks disrupted VWM performance at different time intervals across the two stimuli types and this disruption varied depending on the set size of the visual array. 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.

Pairwise comparisons for set size showed significant drops in performance with each additional item (set size 1 > set size 2 > set size 3 > set size 4). Pairwise comparisons for the five SOAs showed that performance following 517 ms and 634 ms SOAs was significantly lower than that following 867 ms and 984 ms SOAs. Performance following 750 ms SOA was significantly less than at 867 ms SOA. Performance after a 867-ms SOA was significantly higher than performance following all other SOAs. Interestingly, this

¹ The accuracy analysis was not conditionalized based on recognition performance. 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.

Fig. 2 Accuracy in the backward-masking consolidation task as a function of stimulus type, stimulus onset asynchrony (SOA), and set size for Experiment 1. Error bars 1 standard error of the mean (SEM)

suggests that performance was lower at 984 ms compared to 867 ms—an effect that will be addressed in the Discussion.

To further examine the 3-way interaction, separate 4 (set size) $\times$ 5 (SOA) repeated measures ANOVAs were conducted for each stimulus type. The trained analysis yielded a significant main effect for set size [$F(2.69, 188.14)=315.40$, $MSE=.01$, $\eta_p^2=.82$]. The main effect of SOA [$F<1$] and the set size $\times$ SOA interaction [$F(10.34, 723.56)=1.06$, $p=.39$, $MSE=.01$, $\eta_p^2=.02$] did not reach significance. Pairwise comparisons for set size showed the same pattern as the overall analysis ($1>2>3>4$). The lack of an effect for a SOA, and the fact that the effect of SOA did not interact with set size, suggests that trained items were not susceptible to interference from the visual masks, indicating LTM involvement in short-term consolidation.

The novel analysis yielded significant main effects for set size [$F(2.69, 188.55)=421.63$, $MSE=.01$, $\eta_p^2=.86$] and SOA [$F(4, 280)=10.61$, $MSE=.01$, $\eta_p^2=.13$] as well as a significant set size $\times$ SOA interaction [$F(10.82, 757.67)=2.63$, $MSE=.01$, $\eta_p^2=.04$]. Pairwise comparisons for set size showed the same pattern as the overall and trained analyses ($1>2>3>4$). Pairwise comparisons for SOA showed that masks interfered with performance up to 867 ms: performance following a 517 ms SOA was significantly lower than performance following 750 ms, 867 ms, and 984 ms SOAs; performance following a 634 ms SOA was significantly lower than performance following a 867 ms SOA or a 984 ms SOA; and performance following a 867 ms SOA

was significantly higher than all other SOA conditions. This pattern is clearly different from the trained items since there was no evidence for mask interference with trained items. These results replicate Vogel et al. (2006) using more complex stimuli, providing additional support for a consolidation process.

Discussion

Experiment 1 examined the impact the visual long-term representations have on short-term consolidation into VWM. The results provide strong evidence that visual long-term representations facilitate the consolidation process: trained items were unaffected by visual masks even at an SOA of 517 ms, whereas performance did not peak for novel items until an SOA of 867 ms. These results suggest that VWM and LTM have a direct relationship with LTM influencing processing of information in VWM.

Interestingly, there was an overall main effect for trained versus novel stimuli—a result that is in contrast with that of Chen et al. (2006). There are a couple 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 2 examines this possibility by replicating the current experiment using a training procedure similar to that of 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 facilitates VWM processing only in the face of interfering information. 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 (Engle & Kane, 2004; Kane & Engle, 2000); thus, if training improves consolidation only when interfering masks are used, it could suggest a role for general cognitive resources in short-term consolidation.

Overall, there was a trend towards performance following the longest SOA (984 ms) being lower than performance following the second longest SOA (867 ms), particularly for novel shapes. This pattern is surprising considering that the total time between the presentation of the memory array and the test array was equal, the only difference being when the masks were presented. It is possible that with more complex shapes the memory traces begin to decay, which makes them more susceptible to interference from visual masks (e.g., Altmann & Gray, 2002). Future work should more specifically examine the persistence of complex visual information in the face of interfering information.

In summary, Experiment 1 provides evidence that long-term representations facilitate short-term consolidation: explicitly trained stimuli not only show performance advantages over novel stimuli, but they consolidate faster into VWM. This provides the first successful demonstration of explicit training facilitating both VWM performance and VWM consolidation.

Experiment 2

Experiment 1 demonstrated that training improves VWM performance and facilitates short-term consolidation. However, the positive effects of training found in Experiment 1 were in contrast to prior work on the impact of explicit training on VWM performance (e.g., Chen et al., 2006). Experiment 2 examined this discrepancy by replicating the training task used by Chen et al. (2006). Thus, Experiment 2 asks two primary questions: (1) does the type of training matter in terms of how long-term visual representations are developed, and (2) how do different training methods influence short-term consolidation?

More specifically, the 4AFC training task used in Experiment 1 focused on the individual shapes, in particular discriminating a presented shape from three other similar shapes. In Experiment 2 (and in Chen et al., 2006), a change

detection training task was 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 2 provides a contrast between item-based and task-based training.

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 details are being consolidated, the item-based 4AFC training should facilitate consolidation but not the task-based change detection training. 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 1), 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 2 is an important extension of the effects found in Experiment 1 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 methods used in Experiment 2 were the same as the methods in Experiment 1, except where noted below.

Participants

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

Training

A new subset of 12 polygons was 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 2×2 grid (the center four positions of the 4×4 grid used in the consolidation task, subtending $5.0^\circ \times 5.0^\circ$ of visual angle). 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 judgment). There was a 500 ms blank interval between trials. 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 Experiment 1, any sphericity assumptions were corrected using the Huynh-Feldt correction and pairwise comparisons were conducted using the least significance differences method. All analyses were significant at the $p < .05$ level unless otherwise noted.

Change detection training and 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 above 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 compared 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 recognition 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 2 into the allotted 1.5-h time slot, the number of training trials was reduced to 200 compared to 240 in Experiment 1, 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.

Change detection accuracy

To parallel the analyses in Experiment 1, the unconditionalized data² were analyzed across all three variables (stimulus type, set size, and SOA) to examine if differences in short-term consolidation existed between trained and novel shapes. Accuracy was slightly higher for trained items ($M = .80$, $SD = .17$) compared to novel items ($M = .79$, $SD = .16$), though unlike Experiment 1 this slight advantage for trained items was not greater at larger set sizes and there was considerably more variability in performance (see Fig. 3). Similar to Experiment 1, accuracy also dropped significantly as set size increased [set size 1: $M = .90$, $SD = .01$; set size 2: $M = .83$, $SD = .01$; set size 3: $M = .75$, $SD = .02$; set size 4: $M = .69$, $SD = .02$]. Additionally, visual masks did not appear to disrupt performance at earlier SOAs for either novel or trained shapes, another contrast to Experiment 1.

These results were confirmed by a 2 (stimulus type: trained vs. novel) $\times$ 4 (set size: 1, 2, 3, & 4 items) $\times$ 5 (SOA: 517, 634, 750, 867, & 984 ms) repeated measures ANOVA, which 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 $\times$ 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 were no significant main effects of stimulus type [$F(1, 61) = 1.18$, $p = .28$, $MSE = .07$, $\eta_p^2 = .02$] or SOA [$F < 1$] and none of the other interactions reached significance [stimulus type $\times$ SOA: $F(3.71, 226.05) = 1.16$, $p = .33$, $MSE = .01$, $\eta_p^2 = .02$; set size $\times$ SOA: $F < 1$; stimulus type $\times$ set size $\times$ SOA: $F < 1$].

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 Fig. 3). Masks did not interfere with either trained or novel shapes, suggesting encoding was less susceptible to masking even at early SOAs. This is in contrast to Experiment 1 in which the novel shapes showed consolidation effects but the trained shapes did not.

To further examine any possible transfer, paired samples t -tests were conducted between overall change detection training performance and overall set size 4 performance (which has an equivalent difficulty level) for both trained and novel shapes. Experiment-wise error was corrected using the Bonferroni correction ($\alpha = .025$). Performance was significantly better for novel shapes in the consolidation task compared to the training task [$t(61) = -2.88$, $p < .025$], suggesting that

² As in Experiment 1, a conditionalized analysis was conducted for Experiment 2 with recognition accuracy greater than 90 %. However, in Experiment 2 conditionalizing the results led to a significantly smaller sample size ($n = 32$) and low statistical power. Thus, those data are not reported here.

Fig. 3 Accuracy in the backward-masking consolidation task as a function of stimulus type, SOA, and set size for Experiment 2. Error bars 1 SEM

change detection performance improved after training and this extended to novel shapes. There was a trend for consolidation performance to be better than training for trained items, but it did not reach significance [$t(61)=-2.27, p=.027$]. These results suggest that the change detection training led to small improvements on the change detection task itself for both novel and trained stimuli.

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 Experiment 1, suggesting that performance drops as set size increases. However, the lack of any interaction between set size and SOA 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 Experiment 1 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).

To further examine the stimulus type x set size interaction, two separate 4 (set size) x 5 (SOA) repeated-measures ANOVAs were conducted for trained and novel shapes. The trained analysis showed a significant main effect for set size [$F(3, 183)=211.71, MSE=.01, \eta_p^2=.78$]. The main effect for SOA [$F(4, 244)=1.27, p=.28, MSE=.01, \eta_p^2=.02$] and the set size x SOA interaction [$F<1$] did not reach significance. 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. The novel 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 SOA and the set size x SOA interaction did not reach significance [$F_s<1$]. 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 trained analysis.

Cross-experiment analyses

To further examine the differences in training tasks between Experiments 1 and 2, additional cross-experiment analyses were conducted. The goal was to determine how the different training tasks impacted overall VWM performance for both trained and novel shapes, thus these analyses look at overall performance on trained versus novel stimuli collapsing across set size and SOA. Accuracy was higher in Experiment 1 ($M=.82, SD=.07$) compared to Experiment 2 ($M=.79, SD=.10$) and trained items showed higher overall accuracy ($M=.81, SD=.09$) compared to novel items ($M=.80, SD=.08$).

These patterns were confirmed in a 2 (experiment: 1 versus 2) x 2 (stimulus type: trained versus novel) mixed ANOVA with experiment as a between-subjects factor and stimulus type as a within-subjects factor. This revealed significant main effects of stimulus type [$F(1, 131)=4.12, MSE=.002, \eta_p^2=.03$] and experiment [$F(1, 131)=4.22, MSE=.25, \eta_p^2=.03$]. The experiment x stimulus type interaction did not reach significance [$F<1$]. While this pattern does lend some

support to the idea that 4AFC training created more robust long-term visual representations (i.e., a main effect for experiment), the lack of a significant experiment $\times$ stimulus type interaction prohibits making any strong conclusions about the item-specific training effects between experiments.

The same analysis was conducted on conditionalized data (including participants who had at least 90 % accuracy on the intermediate recognition test). If the differences shown between the experiments above can be attributed to better LTM representations, then participants who were able to create those representations should perform equivalently regardless of the training type and there should be a significant effect of stimulus type. While the effect of experiment did go away after conditionalizing the data, so did the effect of stimulus type [main effect for stimulus type: $F(1, 86)=2.56$, $p=.13$, $MSE=.001$, $\eta_p^2=.03$; main effect for experiment and the interaction both had $F_s<1$], again making strong conclusions difficult to draw. While this does provide some evidence that the better overall performance in Experiment 1 was due to more robust LTM representations formed during training, it is also based on a significantly reduced sample size with unequal groups ($N_{\text{Experiment 1}}=56$; $N_{\text{Experiment 2}}=32$). This cross-experiment analysis does indicate, however, that future research examining the effects of LTM on VWM processing should use item-specific, discrimination training for the best chances at creating robust LTM representations using a relatively short training period.

Discussion

Experiment 1 provided evidence of a role for LTM in facilitating short-term consolidation. The goal for Experiment 2 was to extend these results by generalizing them to a different training task. The results showed no significant differences between trained and novel shapes and no clear evidence that consolidation differed between the stimuli types. This is in contrast to Experiment 1 in which there was both an overall advantage for trained shapes and evidence that trained shapes were consolidated faster than novel shapes. Overall, the results from Experiment 2 suggest that change detection training can lead to general task-based improvements on performance (i.e., performance increased after training for novel and trained items), replicating the pattern from Chen et al. (2006). However, in comparison to Experiment 1 the effects of these improvements were modest and not item specific.

The results from Experiments 1 and 2 suggest that the type of training used will determine how training impacts VWM performance: training tasks that focus on discriminating visual items will facilitate short-term consolidation for the trained items and task-based training can lead to general task improvements but may not adequately develop visual LTM representations. The lack of any effects of masking in Experiment 2 may indicate participants were

better able to form a spatial configuration of the array, essentially treating each array as a single object (Blalock & Clegg, 2010; Jiang, Olson, & Chun, 2000). This is supported by the lack of a set size $\times$ SOA interaction (i.e., four items did not require more time to consolidation than a single item). Thus, participants may be improved in their ability to create a global configuration and this may be what led to the general task improvements shown in Experiment 2. However, the lack of any SOA effects could be due to the longer SOA intervals used in this study to accommodate the more complex stimuli thereby reducing the effects of masking. This is addressed in Experiment 3.

Overall, the combined results from Experiments 1 and 2 demonstrate that visual long-term representations can facilitate early VWM processing when training leads to robust LTM representations. While change detection training (i.e., task-based training) could lead to robust LTM representations with additional amounts of training, the item-specific 4AFC training (i.e., item-specific training) used in Experiment 1 created those robust representations more efficiently. Again, the results support a close relationship between LTM and VWM at early processing stages.

Experiment 3

There are two possible alternative explanations for the effects shown in Experiments 1 and 2. First, given the prolonged exposure to a relatively small number of novel items, participants may have adopted a verbal recoding strategy that could explain the overall benefit of trained items versus novel items. While recent research has shown that verbal strategies in change detections tasks do not significantly impact performance (Blalock & VanWormer, 2013), it is still necessary to rule verbal recoding strategies given the amount of training used in this study.

Second, the presentation times used in Experiments 1 and 2 were lengthened compared to previous research given the complexity of the stimuli. However, those times may have been too long to see any meaningful effects of masking and longer encoding intervals may have allowed for more LTM involvement (Lin & Luck, 2012; Makovski & Jiang, 2008). To address these concerns, Experiment 3 replicated the methods used in Experiment 1 (as the 4AFC training was more successful compared to the change detection training in Experiment 2) but included an articulatory suppression task during the trials and the presentation time was shortened to 300 ms. If any effects of training persist after ruling out these alternative explanations, it will provide a stronger case for the role of visual LTM in facilitating early VWM processing.

Method

The methods used in Experiment 3 were the same as Experiment 1 with the following exceptions.

Participants and design

Thirty-one undergraduate students participated for a US\$20 gift card to the university bookstore. Due to time constraints and issues with participant fatigue in conducting the study (e.g., many subjects were dropped in Experiments 1 and 2 to low performance or not finishing in time), the overall number of trials was reduced by examining set sizes 1 and 4 only. This resulted in a 2 (stimulus type: trained and novel shapes) $\times$ 2 (set size: 1 versus 4 shapes) $\times$ 5 (SOA: 317, 434, 550, 667, 784) within-subjects design.

Procedure

As noted above, the number of trials was reduced to 400, which were presented in four 100 trial blocks (the order of which was randomized), including 200 trained trials and 200 novel trials. This resulted in 20 trials per stimulus type $\times$ SOA $\times$ set size condition. Some additional measures were taken to reduce the overall attrition rate of participants. The study session was lengthened to 2 h (though most finished around the 1.5 h mark) and the experiment was paused every 50 trials to allow participants to take breaks. This reduced the overall fatigue and dropout rates for this experiment.

On each trial, the numbers “123,” “456,” or “789” were selected randomly then displayed in the center of the screen with a written prompt to begin repeating the numbers aloud at a rate of about two numbers per second. Participants repeated the numbers during the entire trial, ending when they made a response. A researcher was present during each session to ensure participants complied with the articulatory suppression task. After the numbers were presented, a fixation cross appeared for 500 ms followed by the memory items, which were presented for 300 ms. The delay intervals between the offset of the stimuli and the onset of the masks were the same in Experiment 3 as in Experiment 1 (17, 134, 250, 367, and 484 ms) but, due to the shorter presentation time for the items, the SOAs are 200 ms shorter (SOA: 317, 434, 550, 667, 784).

Results

Three participants were dropped from the analyses; one participant for not following instructions (pressed an incorrect key throughout) and two participants had an overall accuracy that was two standard deviations below the mean resulting in 28 total participants. Pairwise comparisons were conducted using the least significance differences method. All analyses are significant at the $p < .05$ level unless otherwise noted.

4AFC training and recognition performance

As in Experiment 1, accuracy on the 4AFC training task was high ($M = .97$, $SD = .03$). Performance was consistently high across four 60 trial blocks (Block 1: $M = .97$, $SD = .04$; Block 2: $M = .98$, $SD = .03$; Block 3: $M = .97$, $SD = .03$; Block 4: $M = .97$, $SD = .04$). A one-way repeated measures ANOVA conducted on the 4AFC performance data across blocks showed no significant difference over time [$F < 1$], mirroring the results from Experiment 1.

Recognition accuracy was high at the immediate recognition test ($M = .93$, $SD = .14$) but dropped at the final recognition test ($M = .87$, $SD = .17$). A paired-samples t -test comparing immediate and final recognition accuracy showed no significant difference, however [$t(27) = 1.61$, $p = .12$; note recognition data from one participant was missing]. This result suggests participants were able to form long-term representations and they were relatively stable, though there was a trend toward lower recognition at the final recognition test. Again this pattern is similar to Experiment 1.

Change detection accuracy

Overall, unconditionalized³ accuracy on the change detection task was higher for trained stimuli at both set size 1 ($M = .94$, $SD = .07$) and set size 4 ($M = .62$, $SD = .11$) compared to novel stimuli (set size 1: $M = .93$, $SD = .07$; set size 4: $M = .60$, $SD = .12$; see Fig. 4). Performance did increase with longer SOAs though there was a larger increase for novel items compared to trained items, suggesting novel items were more susceptible to mask interference. A 2 (stimulus type: trained and novel) $\times$ 2 (set size: 1 and 4) $\times$ 5 (SOA: 317, 434, 550, 667, & 784) repeated-measures ANOVA confirmed that this difference was reliable with a significant main effect of stimulus type [$F(1, 27) = 4.29$, $MSE = .01$, $\eta_p^2 = .14$] as well as significant main effects for set size [$F(1, 27) = 1,025.73$, $MSE = .01$, $\eta_p^2 = .97$] and SOA [$F(4, 108) = 3.08$, $MSE = .03$, $\eta_p^2 = .10$]. There was also a significant three-way interaction [$F(4, 108) = 3.12$, $MSE = .01$, $\eta_p^2 = .10$], suggesting the effects of set size and SOA varied depending on if the items were trained or novel. These results suggest that even with a shorter presentation time and a concurrent verbal task to rule out verbal recoding strategies, trained stimuli led to better VWM performance and that consolidation differed between stimulus types.

To further examine the consolidation differences between trained and novel stimuli, separate 2 (set size) $\times$ 5 (SOA) repeated-measures ANOVAs were conducted for each stimulus type. For trained stimuli, accuracy was significantly higher with a set size of one compared to four [$F(1, 27) = 737.47$,

³ A conditionalized analysis was conducted with recognition accuracy at 90 % or above; however, it yielded similar results. Thus, all data are included in the change detection analysis.

Fig. 4 Accuracy in the backward-masking consolidation task as a function of stimulus type, SOA, and set size for Experiment 3. Error bars 1 SEM

$MSE=.01$, $\eta_p^2=.97$] but there was no significant effect of SOA [$F(4, 108)=1.07$, $P=.38$, $MSE=.01$, $\eta_p^2=.04$] or significant set size $\times$ SOA interaction [$F(4, 108)=2.22$, $p=.07$, $MSE=.01$, $\eta_p^2=.08$]. These results are consistent with the results from Experiment 1. Trained items showed little effect of masking and that was true for both set sizes.

In the novel analysis, accuracy was also significantly higher with a set size of 1 compared to a set size of 4 [$F(1, 27)=742.87$, $MSE=.01$, $\eta_p^2=.97$] and there was a significant effect of SOA [$F(4, 108)=2.45$, $p=.05$, $MSE=.01$, $\eta_p^2=.08$]. The set size $\times$ SOA interaction was also significant [$F(4, 108)=2.50$, $MSE=.01$, $\eta_p^2=.09$]. Again, these results are consistent with Experiment 1; early masks impaired consolidation of items into VWM particularly at large set sizes. This also replicates the original work by Vogel et al. (2006).

Discussion

Experiment 3 replicated the effects shown in Experiment 1 while addressing two possible alternative explanations: using an articulatory suppression task to eliminate verbal strategies and shortening the overall presentation time to 300 ms to increase the effectiveness of the masking. The pattern of data was largely the same, indicating that the benefits of training were not due to verbal strategies and that the effects persist with shorter SOAs. This provides stronger evidence for a role of visual LTM in facilitating VWM processing; trained shapes were better remembered overall and consolidated more quickly compared to novel shapes.

While shortening the encoding time to 300 ms allowed for more effective masking, it did dramatically reduce overall accuracy in the set size 4 condition. The overall accuracy in Experiment 1 at set size 4 was .71 ($SD=.13$) while accuracy in Experiment 3 at set size 4 was .62 ($SD=.12$). Participants showed particularly low performance at the shortest SOA interval in the novel condition, though performance was still significantly above chance [one-sample t -test: $t(27)=4.05$, $p<.05$]. A possible solution for this in future research would be to lower the set size when using more complex stimuli and faster encoding times.

Experiment 3 also addressed another shortcoming of Experiments 1 and 2 by reducing the attrition rate of participants. Including breaks, offering payment instead of extra credit, and reducing the number of trials reduced the number of dropped participants and lowered overall participant fatigue. This suggests that the effects shown in Experiments 1 and 2 were not due to any type of bias between the participants who were included and those who were dropped.

General discussion

The current study examined how long-term representations facilitate short-term consolidation of information into VWM. In all experiments, trained and novel shapes were compared in a backward-masking task. Experiment 1 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 LTM in facilitating VWM processes. Given this positive effect of training was unique to the literature, Experiment 2 examined the impact of LTM on short-term consolidation using a different training task that more closely mirrored prior work. Using this different training task, only general task improvements were shown for trained and novel shapes with no significant differences between stimulus type. Experiment 3 replicated Experiment 1 (showing better VWM performance and faster consolidation for trained shapes) but eliminated verbal strategies and shortened the encoding interval to avoid other LTM influences on performance. Taken together these experiments show that item-specific training is best when examining LTM influences on VWM, and that robust long-term representations can influence VWM processing by improving performance and making VWM representations less susceptible to interfering masks.

What do these results tell us about short-term consolidation?

This research shows the process of short-term consolidation lies at the intersection of several key cognitive processes: perception, attention, working, and LTM. 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 study demonstrates that LTM influences the process of consolidation between perception and WM, with trained items being consolidated more quickly than novel items.

In addition to demonstrating a role for long-term representations in facilitating consolidation, the current set of experiments also addresses a question that has not yet been addressed in the consolidation literature: what exactly is being consolidated? The positive result of training found in Experiment 1 using the 4AFC training task suggests that visual details of objects are consolidated into VWM. Specifically, the 4AFC training task emphasized distinguishing visual details between a set of similar shapes which in turn led to faster consolidation. Demonstrating that visual features are consolidated into VWM is in line with 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 overwritten easily by subsequent visual information (though see Blalock, 2013 for another view). This conclusion also fits with research showing improved resistance to visual interference when attention is directed toward specific objects in VWM (e.g., Makovski, Sussman, & Jiang, 2008; Schmidt, Vogel, Woodman, & Luck, 2002). The current experiments also replicated the set size effects (i.e., more time was needed

to consolidate more items), which is consistent with the argument that visual details are consolidated into VWM.

The item-specific performance increases following 4AFC training is in contrast to general task improvements following change detection training. Change detection training led to improved overall performance and less interference from masking in both trained and novel shapes, but these gains were modest in comparison to the 4AFC training. The lack of a set size $\times$ SOA interaction may indicate that participants were able to create a global configuration of the arrays (Blalock & Clegg, 2010; Jiang et al., 2000) and treated them as a single item leading to improved performance. However, the longer SOAs used in Experiment 2 may have reduced the effects of masking here so future work will have to examine what aspect of change detection training led to the task-based improvements shown in Experiment 2 and Chen et al. (2006).

Another possible explanation for the differences in performance between the 4AFC and change detection training tasks comes from work on implicit learning of shapes. Logie, Brockmole, and Vandenbroucke (2009) used a modified Hebb (1961) learning task in which shapes were repeated across trials. They showed learning occurred only when bound objects (i.e., the same shape-color-location pairing) were repeated every trial. If any novel trials were shown between repeating trials, repetition did not improve performance. Logie et al. (2009) argue that, in a change detection task, each new array of shapes overwrites the previous array leaving no traces for long-term learning. This could explain why less item-specific learning occurred following the change detection training task; the object representations did not persist across trials, making learning more difficult.

Training long-term visual representations

Previous research by Chen et al. (2006) indicated that explicit training led to an overall improvement on a change detection task but they found no advantage for trained 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 2 replicates their results, Experiments 1 and 3 challenge the conclusion that long-term visual representations cannot improve VWM processing. Instead, the experiments suggest that the type of training used will determine what impact long-term representations will have on VWM processing.

More specifically, in order to create stable LTM 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, using a repeat detection task, Brady, Konkle, Alvarez, and

Olivia (2008) showed highly detailed LTM for real-world objects. Taken together with Experiments 1 and 3, 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.

While there were significant effects of training in this study, the overall effect size of training was small. There are several possible reasons for this. First, there was a small amount of training. The fact that any positive effects of training were demonstrated after only 240 trials (in Experiments 1 and 3) is fairly remarkable, and future work should examine if more training leads to larger gains in performance. Second, the abstract and highly similar polygons used in this study may have limited any performance gains due to proactive interference effects (this is discussed in more detail below). It is possible that with more distinguishable stimuli (e.g., real-world objects) training would lead to greater performance gains. Finally, these small gains may still have important implications for real-world vision by detecting changes more quickly and accurately in familiar environments. Work by Hollingworth and colleagues (Hollingworth & Luck, 2009; Hollingworth, Richard, & Luck, 2008) has shown the critical role that VWM plays in navigating our complex visual environment, and any help from visual LTM in processing information in VWM could improve real-world performance. Additional research would be needed, however, to determine what that improvement may look like and what the limitations are.

Thus, these results suggest that while change detection training is not ideal in developing robust long-term representations (though it is still a possibility with enough training), it can lead to overall task-based improvements in VWM performance. Item-specific training tasks that focus on distinguishing shapes can improve item-specific VWM processing so long as the training leads to robust long-term representations. The current study suggests part of this performance boost comes from speeding up the consolidation process for familiar shapes, though it is also possible that masking interfered with maintenance as well (though that is less likely in Experiment 3 with the shorter SOAs). LTM likely influences each stage of VWM processing and fully understanding those influences will be important for developing VWM theories.

Relationship between VWM and LTM

In terms of the greater WM literature, the current results are consistent with theories that argue for a close relationship between LTM and WM. Cowan's Embedded Processes model for example (Cowan, 1999, 2001) argues that activated long-term representations make up the contents of WM. The current data are consistent with that idea, with objects already within LTM being consolidated more quickly. The current results are also consistent with Logie's Workspace Model, which argues that perceptual input passes through a long-

term knowledge base before entering WM (Logie & van der Meulen, 2009; van der Meulen, Logie, & Sala, 2009). Both models are able to explain the very early effects of LTM on VWM performance.

These results are also consistent with Endress and Potter's (2014) conclusion that VWM has a very close relationship with LTM, and they provide some additional insights to their framework. In the current study, the shapes were meaningless, highly similar, and did not allow for categorization. Based on Endress and Potter's framework, these kinds of stimuli would be stored in a capacity-limited WM as they are effortful to encode, require active maintenance, and are susceptible to substantial build-up of proactive interference even on novel trials given the high similarity of the shapes.

The trained items, however, could have been stored in the temporary, unconsolidated LTM store proposed by Endress and Potter, making those shapes faster to encode (and possibly maintain) compared to novel shapes. This would also likely lower the amount of proactive interference, as the exposure to the trained items would allow participants to distinguish those shapes even with the high level of similarity. Reliance on this temporary store can also explain relatively low recognition accuracy on the trained shapes, even in the 4AFC training; when correcting for guessing on the intermediate recognition test (averaging across Experiments 1 and 3), participants could distinguish a trained item on only 78 % of the trials. The nature of the training may have led to more temporary LTM representations instead of more stable ones. Future work will have to address this by increasing the amount of training to create more stable LTM representations. Work by Brady et al. (2008) indicates it is possible to create robust long-term representations, but more work is needed to know how much training is necessary. Even with these less stable LTM representations, however, the current study showed positive effects of training following a 4AFC training task providing evidence for LTM involvement in VWM processing.

Summary

The current set of experiments is a start to examining the interaction between perception, VWM, and LTM. The results lead to several key conclusions: (1) there is a close relationship between perception, VWM, and LTM; (2) training can improve the speed at which information gets into WM so long as the training leads to robust LTM representations; (3) item-specific training on a set of visual items can lead to improvements in VWM performance compared to novel items; and (4) that stimulus-specific training emphasizing distinguishing shapes from each other is the preferred training method to form robust long-term visual representations compared to change detection (or task-based) training. 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. Future work in this area will provide a better understanding of how VWM operates within the larger cognitive system.

References

- Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: the functional relationship of decay and interference. *Psychological Science*, 13(1), 27–33. doi:10.1111/1467-9280.00405
- Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology*, 53, 221–227.
- Attneave, F., & Amoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, 53(6), 452–471.
- Baddeley, A. D. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, 63(1), 1–29. doi:10.1146/annurev-psych-120710-100422
- Baddeley, A. D., & 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. 1–35). Cambridge, NY: Cambridge University Press.
- Blalock, L. (2013). Mask similarity impacts short-term consolidation in visual working memory. *Psychonomic Bulletin & Review*, 20, 1290–1295. doi:10.3758/s13423-013-0461-9
- Blalock, L., & Clegg, B. A. (2010). Encoding and representation of simultaneous and sequential arrays in visuospatial working memory. *The Quarterly Journal of Experimental Psychology*, 63(5), 856–862. doi:10.1080/17470211003690680
- Blalock, L., & VanWormer, L. A. (2013, November). Verbal strategies, aptitude, and change detection performance in visual working memory. *Poster Presented at the 2013 Annual Meeting of the Psychonomic Society*. Toronto, CA.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, 105(38), 14325–14329.
- Buttle, H., & Raymond, J. E. (2003). High familiarity enhances visual change detection for face stimuli. *Attention, Perception, & Psychophysics*, 65(8), 1296–1306.
- Chen, D., Eng, H. Y., & Jiang, Y. (2006). Visual working memory for trained and novel polygons. *Visual Cognition*, 14(1), 37–54. doi:10.1080/13506280544000282
- 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(1), 109.
- Cowan, N. (1999). An embedded-processes model of working memory (pp. 62–101). In A. Miyake & P. Shah (Eds.), *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.
- Cowan, N. (2001). *The magical number 4 in short-term memory: a reconsideration of mental storage capacity*, 24, 1–99.
- Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin & Review*, 14(4), 620–628.
- Dux, P. E., & Marois, R. (2010). The attentional blink: a review of data and theory. *Attention, Perception, & Psychophysics*, 71(8), 1683–1700. doi:10.3758/APP.71.8
- Endress, A. D., & Potter, M. C. (2014). Large capacity temporary visual memory. *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General*, 143(2), 548–565. doi:10.1037/a0033934
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of Learning and Motivation*, 44, 145–199.
- Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance*, 19(4), 845.
- Hartshorne, J. K. (2008). Visual working memory capacity and proactive interference. *PLoS ONE*, 3(7), e2716. doi:10.1371/journal.pone.0002716.s001
- Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain Mechanisms and Learning* (pp. 37–51). Oxford, UK: 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., & 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
- Jiang, Y., Olson, I., & Chun, M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 683–702.
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36(2), 138–202.
- 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, and Cognition*, 26(2), 336.
- Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, 1, 153–161.
- Lin, P. H., & Luck, S. J. (2012). Proactive interference does not meaningfully distort visual working memory capacity estimates in the canonical change detection task. *Frontiers in Psychology*, 3, 42. doi:10.3389/fpsyg.2012.00042/abstract
- 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* (pp. 1–32). New York: Academic.
- Logie, R. H., & van der Meulen, M. (2009). Fragmenting and integrating visuo-spatial working memory. In J. R. Brockmole (Ed.), *The Visual World in Memory* (pp. 1–32). Hove, UK: Psychology.
- Logie, R. H., Brockmole, J. R., & Vandenbroucke, A. R. E. (2009). Bound feature combinations in visual short-term memory are fragile but influence long-term learning. *Visual Cognition*, 17(1–2), 160–179. doi:10.1080/13506280802228411
- Luck, S. J. (2008). Visual short-term memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual Memory* (pp. 43–85). New York: Oxford University Press.
- Makovski, T., & Jiang, Y. (2008). Proactive interference from items previously stored in visual working memory. *Memory & Cognition*, 36(1), 43–52. doi:10.3758/MC.36.1.43
- Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2), 369–380. doi:10.1037/0278-7393.34.2.369

- Martens, S., & Wyble, B. (2010). The attentional blink: past, present, and future of a blind spot in perceptual awareness. *Neuroscience & Biobehavioral Reviews*, 34(6), 947–957. doi:[10.1016/j.neubiorev.2009.12.005](https://doi.org/10.1016/j.neubiorev.2009.12.005)
- 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(6), 1463–1475. doi:[10.1037/0096-1523.31.6.1463](https://doi.org/10.1037/0096-1523.31.6.1463)
- Raymond, J. E., Shapiro, K. L., & 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(3), 849.
- Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Attention, Perception, & Psychophysics*, 64(5), 754–763.
- Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences*, 1(8), 291–296.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74, 1–29.
- van der Meulen, M., Logie, R. H., & Sala, S. D. (2009). Selective interference with image retention and generation: evidence for the workspace model. *The Quarterly Journal of Experimental Psychology*, 62(8), 1568–1580. doi:[10.1080/17470210802483800](https://doi.org/10.1080/17470210802483800)
- 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](https://doi.org/10.1037/0096-1523.32.6.1436)
- Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 2, 440–445.