

# The Time Course of Consolidation in Visual Working Memory

Edward K. Vogel  
University of Oregon

Geoffrey F. Woodman  
Vanderbilt University

Steven J. Luck  
University of Iowa

How long does it take to form a durable representation in visual working memory? Several theorists have proposed that this consolidation process is very slow. Here, we measured the time course of consolidation. Observers performed a change-detection task for colored squares, and shortly after the presentation of the first array, pattern masks were presented at the locations of each of the colored squares to disrupt representations that had not yet been consolidated. Performance on the memory task was impaired when the delay between the colored squares and the masks was short, and this effect became larger when the number of colored squares was increased. The rate of consolidation was approximately 50 ms per item, which is considerably faster than previous proposals.

*Keywords:* visual working memory, masking, attention, consolidation

The visual working memory (WM) system is responsible for the short-term retention and manipulation of visual information from the immediate environment (Baddeley & Hitch, 1974; Phillips, 1974). A fundamental property of this memory system is that it is subject to severe storage capacity limitations. Specifically, visual WM can maintain information about approximately three objects at any given time, and this information appears to be coded in the form of integrated object representations, rather than as a collection of disconnected visual features (Irwin & Andrews, 1996; Jiang, Olson, & Chun, 2000; Lee & Chun, 2001; Luck & Vogel, 1997; Sperling, 1960; Vogel, Woodman, & Luck, 2001; but see Wheeler & Treisman, 2002; Xu, 2002b).

Recent research has indicated that storage capacity is not the sole limitation of the visual WM system. Evidence from several paradigms has suggested that the process that consolidates information into a durable WM representation is also severely limited in its processing capacity (Chun & Potter, 1995; Duncan, Ward, & Shapiro, 1994; Jolicœur & Dell'Acqua, 1998; Potter, 1976; Raymond, Shapiro, & Arnell, 1992; Vogel, Luck, & Shapiro, 1998; Ward, Duncan, & Shapiro, 1996). By consolidation, we mean the process of transforming a fleeting perceptual representation into a durable WM representation that can survive the presentation of

new sensory inputs (Chun & Potter, 1995; Enns & Di Lollo, 1997; Jolicœur & Dell'Acqua, 1998; Miller, Li, & Desimone, 1993).<sup>1</sup>

The consolidation process has been assumed to be highly capacity limited and very slow, with some theorists proposing that it takes as much as 500 ms to store a single item in WM (Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998; Ward et al., 1996). However, these estimates have been inferred primarily from dual-task paradigms, and these paradigms may overestimate the duration of consolidation by including processes other than consolidation. Consequently, it is currently unclear exactly how much time is required for the visual system to construct a durable WM representation, and it is also unclear whether this process is slower for multiple-element stimulus arrays than for individual objects. Thus, the goals of the present study were (a) to measure the time course of visual WM consolidation and (b) to determine whether the rate of consolidation depends on the number of items being consolidated.

---

<sup>1</sup> The term *consolidation* has become conventional in both long-term memory and WM research, but *vulcanization* might be a better term. Vulcanization is a process in which rubber is combined with sulfur under heat and pressure, which makes the rubber both stronger and more resilient (as well as less odorous). Without vulcanization, rubber tends to become brittle over time and will crumble when stressed. By analogy, a vulcanized representation would be both strong and flexible and would not crumble over time as new perceptual inputs arrive. In contrast, the term *consolidation* implies the compaction of previously disparate pieces of information, which may not be a good characterization of the process that creates durable WM representations. However, we use the term *consolidation* in this article to provide continuity with previous studies (and to avoid adding more jargon to the field).

---

Edward K. Vogel, Department of Psychology, University of Oregon; Geoffrey F. Woodman, Department of Psychology, Vanderbilt University; Steven J. Luck, Department of Psychology, University of Iowa.

This study was made possible by National Institute of Mental Health Grants MH56877, MH63001, and MH65034, National Science Foundation Grant SBR 98-09126 to Steven J. Luck and Grant BCS 0617681 to Edward K. Vogel, and Human Frontier Science Program Grant RG0136.

Correspondence concerning this article should be addressed to Edward K. Vogel, Department of Psychology, 1227, University of Oregon, Eugene, OR 97403-1227. E-mail: vogel@darkwing.uoregon.edu

### Early Evidence for Slow Visual WM Consolidation

Potter (1976) conducted a series of experiments to measure the amount of time required for perceptual identification and memory encoding. Photographs of complex scenes were presented at fixation in a rapid serial visual presentation (RSVP) paradigm. Sixteen photographs were presented on each trial, and the presentation rate was one, three, four, six, or eight stimuli per second. In a recognition condition, the observers viewed an RSVP sequence and were then given a photograph and asked whether it had been presented in that sequence. Recognition accuracy was very high at the slowest rate of presentation but fell precipitously at the faster rates. This drop in performance could have been due to either an inability to identify the stimuli or an inability to form a durable WM representation.

To rule out the possibility that perception was impaired at the fastest rates, Potter's study also included a picture-preview condition, in which the observers were shown a picture prior to the RSVP sequence; at the end of the sequence, they indicated whether they had detected the previewed picture in the sequence. Performance was highly accurate in this condition at all stimulus presentation rates, suggesting that perception was not impaired at the highest rates. However, the observers could have performed this task by detecting low-level features of the previewed picture rather than by fully identifying each picture.

To rule out this possibility, a word-preview condition was tested in which the observers were shown a word at the beginning of each sequence; at the end of the sequence, they indicated whether they had detected a scene matching that word in the sequence. For example, the word might be *boat* and the RSVP sequence might contain a photograph of a sailboat on a small lake. As in the picture-preview condition, accuracy was high in this condition at all stimulus presentation rates. Because the word specified the concept represented by the photograph rather than the low-level features of the photograph, this result indicates that the observers could identify the scenes at a conceptual level at the highest rates, even though they were highly inaccurate in the recognition condition at the highest rates. From these results, Potter (1976) concluded that the visual system can rapidly extract the gist of a photograph and form a conceptual representation of it. However, the process of forming a durable representation in WM is slower, such that scenes can be identified faster than they can be consolidated.

This study is an important landmark in setting the stage for more recent work on visual WM consolidation, but it leaves three important questions unanswered. First, exactly how much time is required for the consolidation process? The RSVP paradigm requires the observers to both perceive and consolidate each photograph in a short period of time, and it is difficult to provide independent measures of perception and consolidation timing. Second, how much time is required to consolidate one photograph? When many photographs are presented in succession, the consolidation of one photograph may interfere with the consolidation of the preceding and subsequent photographs, making it difficult to measure the consolidation time for a single photograph. Third, how much time is required to consolidate one object? The photographs were very complex and contained multiple objects, and it is unknown whether the consolidation time would be shorter for

individual visual objects as opposed to complex arrangements of objects. The present study was designed to provide answers to these three questions.

### Recent Evidence From Dual-Task Experiments

More recent work on the time course of the consolidation process has stemmed from research on the attentional blink phenomenon, which is thought to occur primarily as a consequence of limitations in the consolidation process. In the attentional blink paradigm, an observer attempts to detect two targets embedded in a rapid stream of stimuli. When the second target (T2) is presented 200–600 ms after the first target (T1), the observer is impaired at reporting the identity or even the presence of T2 (Broadbent & Broadbent, 1987; Raymond et al., 1992; Reeves & Sperling, 1986; Shapiro, Raymond, & Arnell, 1994). Currently, most models of the attentional blink agree that the deficit is primarily the result of a failure to adequately consolidate T2 into a stable WM representation while T1 is undergoing consolidation (Chun & Potter, 1995; Shapiro, Arnell, & Raymond, 1997; Vogel & Luck, 2002; Vogel et al., 1998). By this logic, the long duration of the attentional blink is a result of the slow consolidation of T1, during which T2 cannot be consolidated and is consequently susceptible to being masked by or confused with other items in the stream (Giesbrecht & Di Lollo, 1998; Isaak, Shapiro, & Martin, 1999). From these results, one could infer the time course of consolidation by simply measuring the duration of the attentional blink, resulting in an estimate of approximately 500 ms.

However, such an estimate may be inflated in several ways. First, even if the attentional blink solely reflects interference between the consolidation of T1 and the consolidation of T2, the offset of T1 consolidation will reflect the sensory and perceptual processes that precede T1 consolidation as well as the T1 consolidation process itself. Consequently, the duration of the attentional blink does not just reflect the duration of the T1 consolidation process but instead reflects the duration of all processes between T1 onset and the offset of T1 consolidation. Second, a significant period of time may be required to switch attention to T2 after T1 consolidation is complete (Weichselgartner & Sperling, 1987). Third, if the T1 and T2 tasks are not identical (e.g., a letter identification task for T1 and a letter detection task for T2), then additional time may be required to perform task switching (Potter, Chun, Banks, & Muckenhoupt, 1998). Fourth, Jolicoeur (1998) has demonstrated that a speeded response to an auditory T1 produces an attentional blink for a visual target, which indicates that response processes associated with T1 processing may also interfere with the ability to consolidate T2. Thus, the assumption that the attentional blink solely reflects interference between T1 consolidation and T2 consolidation is not valid, which undermines any attempts to use the duration of the attentional blink as a measure of the duration of T1 consolidation. Thus, the attentional blink paradigm does not provide a good means of estimating the time course of WM consolidation.

To be fair, it should be noted that no one has explicitly claimed that the 500-ms duration of the attentional blink reflects a 500-ms consolidation time. Duncan et al. (1994) come close to such a claim, proposing that the duration of the attentional blink in a somewhat different paradigm reflects the dwell time of attention.

However, many authors at least seem to imply that the long duration of the attentional blink is related to a slow consolidation process (e.g., Chun & Potter, 1995). Indeed, previous work from our own laboratory has assumed that the long duration of the attentional blink implies that WM consolidation is much slower than the process of object recognition (Woodman, Vogel, & Luck, 2001).

Jolicœur and Dell'Acqua (1998) developed an approach for directly examining short-term consolidation using a dual-task procedure that combines a verbal WM task with an auditory speeded-response task. On each trial, the observers were first presented with a visual array of to-be-remembered alphanumeric characters (T1), and this was followed by a low- or high-pitched tone (T2). The observers made an immediate speeded response to the tone, and memory for the character array was tested at the end of the trial. When the delay between T1 and T2 was short (350–550 ms), responses to the tone were very slow, but the responses became progressively faster as the delay between T1 and T2 increased, eventually reaching asymptote after several hundred milliseconds. In addition, the response times to the tone were much slower when the observer had to remember arrays with several items rather than just remembering a single item, and this difference was largest at short delays. These results indicate that the consolidation of the memory array into WM interfered with the performance of the auditory task at a later, amodal stage of processing (i.e., response selection), suggesting that WM consolidation and the selection of a motor response require similar “central” mechanisms. These results provide two important details regarding the time course of consolidation. First, they demonstrate that consolidation exists as a limited-capacity amodal process that can be distinguished from other forms of stimulus encoding, such as modality-specific sensory and perceptual processes. Second, the results indicate that consolidation is limited in capacity, requiring more time when more items must be remembered.

However, these results are not well suited for quantifying the rate of consolidation in visual WM. First, the dual-task slowing observed at short delays presumably reflects interference between T1 consolidation and T2 response selection processes. It is not clear whether these interference effects are due to a postponement of T2 response selection until T1 consolidation is complete, or whether both consolidation and response selection for the two tasks operate simultaneously but with a severe slowing for both processes (Jolicœur & Dell'Acqua, 1998; Ward et al., 1996). In addition, because the visual stimuli were alphanumeric characters, the results probably reflect a combination of verbal and visual WM consolidation. Moreover, the process of transferring verbal information in WM appears to involve an articulatory process (Baddeley, 1986; Cowan, 1997), and the finding that verbal WM encoding interferes with making a manual response to a tone could reflect interference between two response selection processes (subvocal articulation for T1 and an overt button-press for T2). Thus, the results of Jolicœur and Dell'Acqua (1998) do not provide strong evidence regarding either the time course or the central resource requirements of visual WM encoding. However, a more recent study by Dell'Acqua and Jolicœur (2000) used a tone followed by a masked, nonverbal memory array. Speeded responses to the tone were found to interfere with the encoding of the memory array, which indicates that visual WM consolidation does

depend on an amodal, central resource. However, the time course of visual WM consolidation was not explored in that study.

### Recent Evidence From Masking

Gegenfurtner and Sperling (1993) explored the time course of WM consolidation using a combination of a masking procedure and the partial report procedure that Sperling (1960) previously developed for assessing the time course of iconic memory. In the Gegenfurtner and Sperling study, observers were presented with three rows of 3–4 letters, followed at a variable delay by a tone that cued which row of letters should be reported. In addition, the letters were followed after a variable delay by a pattern mask (the cue delay and mask delay were manipulated independently). By presenting a mask at a variable delay after the array of letters, the experimenters were able to determine how much time was required to transform the iconic representation of the letters into a durable representation that could survive masking. The use of a cue to direct attention to one of the three rows is a complication that makes it difficult to assess the timing of WM consolidation independently of the timing of attention. Fortunately, this study included many trials in which the cue was presented at or after the time of the mask and had essentially no effect. These are called nonselective transfer trials.

On nonselective transfer trials, Gegenfurtner and Sperling (1993) found that the observers could consolidate between 3.5 and 4.5 letters within 100 ms of stimulus onset (i.e., with a delay of 100 ms between letter-array onset and mask onset). Thus, this study estimated that it takes 20–30 ms to consolidate each letter. Shibuya and Bundesen (1988) obtained a similar estimate of the consolidation rate using a similar paradigm. Of note, these estimates are more than an order of magnitude faster than the 500 ms/item consolidation rate implied by attentional blink experiments.

However, several factors complicate this estimate of the time course of visual WM consolidation. First, alphanumeric stimuli were used, and the observers gave verbal reports, so the results may reflect the time course of verbal rather than visual WM consolidation. The verbal and visual WM storage systems have been demonstrated to be very distinct mechanisms (Baddeley, 1986; Smith & Jonides, 1997). Consequently, the process of consolidating verbal material may be quite different from the process of consolidating visual material. Second, the observers were highly trained, experiencing at least 1,000 trials before data collection began and then thousands of subsequent trials, and this may have led to a substantial increase in consolidation speed. Third, the rate of transfer appeared to be much faster for the central row than for the top and bottom rows, presumably owing to acuity and attentional differences between central and peripheral stimuli. Fourth, the masks may have influenced perception of the letters as well as disrupting consolidation. Thus, although the Gegenfurtner and Sperling (1993) study provides a rich set of data about WM consolidation, it does not provide a means of estimating the rate of consolidation in visual WM.

### Overview of the Present Study

The goals of the present study were (a) to provide an initial estimate of the time course of visual WM consolidation and (b) to

determine how much additional processing time is necessary for encoding multiple items into memory. To do this, we used a procedure similar to that used by Gegenfurtner and Sperling (1993), in which consolidation was interrupted by masks at various points throughout its time course, which makes it possible to determine how much information has been consolidated in visual WM at that moment.

As illustrated in Figure 1, we used a visual WM paradigm in which observers were presented with two arrays of simple colored squares that were separated by a 900-ms interstimulus interval. The observer's task was to report whether the two arrays were the same or whether they differed in the color of one of the squares. To interrupt the consolidation process for the memory array, we presented pattern masks at the locations of each of the items in the array after various time intervals following the offset of the memory items. The masks were intended to disrupt processing after perceptual analysis was largely completed but before the representations had been consolidated, and this was achieved by using simple stimuli that could be perceived rapidly and relatively long delays between the stimuli and the masks. In addition, follow-up experiments were conducted to demonstrate that the masks did not significantly interfere with the perceptual analysis of the memory array.

In the first experiment, we manipulated both the number of items in the memory array and the delay between the memory array and the masks. Manipulating these two factors allowed us to measure the rate of information accrual in WM and to determine whether that rate is contingent on the number of objects to be encoded in the memory array. In the second experiment, we held the number of items constant and tested a very dense set of mask delays to provide a precise measure of the rate of consolidation. In Experiments 3, 4, and 5, we provided evidence that the masks in this paradigm specifically impair the WM consolidation process rather than earlier perceptual processes or later decision processes.

### Experiment 1

In Experiment 1, we sought to provide an initial estimate of the time course of visual WM consolidation for memory arrays containing one, two, three, or four colored squares. In previous studies of storage capacity with nearly identical stimuli and procedures, we have shown that observers are near perfect with memory arrays

of up to three items and are about 90% accurate with four items (Luck & Vogel, 1997; Vogel et al., 2001; Woodman et al., 2001). To interrupt consolidation, we presented pattern masks at the locations of each of the memory items shortly after the offset of the memory array (see Figure 1). We sampled five delays between the offset of the memory array and the onset of the masks, ranging from 17 ms to 484 ms. The masks were large multicolored squares centered over the locations of each of the items in the memory array, and they were intended to disrupt the perceptual representations of the items in the memory array before they could be consolidated in visual WM.

We chose highly discriminable colored squares as the memory items for three reasons. First, it is currently unknown whether consolidation takes longer for complex, multifeatured objects, and testing single-featured items should provide a lower bound to the consolidation time estimate. Second, using simple suprathreshold colored squares rather than complex objects reduces the amount of time necessary to form an adequate perceptual representation of the items in the array. Thus, the masks are less likely to interrupt the perceptual processing of the memory array and should primarily disrupt the postperceptual processes that produce a durable representation in visual WM. Third, unlike the alphanumeric characters used by Gegenfurtner and Sperling (1993), these colors should not vary greatly in discriminability as a function of eccentricity over the relatively small portion of the visual field that was tested (Anderson & Thibos, 1999; Brown, Govan, & Block, 1983).

### Method

**Participants.** Ten college students participated for course credit. These participants were between the ages of 18 and 30, reported normal or corrected-to-normal visual acuity and normal color vision, and reported no history of neurological problems.

**Stimuli.** The stimuli were presented on a video monitor with a gray background ( $8.20 \text{ cd/m}^2$ ). The center of each item was located within a  $9.80^\circ \times 7.30^\circ$  region, viewed at a distance of 70 cm. Each item was randomly placed within this region with the constraint that the items in a given array were separated from each other by at least  $2^\circ$  (center to center). The color of each object was randomly selected without replacement from a set of seven colors: white, red, blue, green, black, yellow, and violet (see Vogel et al., 2001, for the color coordinates). In the memory and test arrays, each square subtended  $0.65^\circ \times 0.65^\circ$  of visual angle, and each of these arrays consisted of one, two, three, or four squares.

Each mask was composed of four colored squares that were abutted to form a  $2 \times 2$  checkerboard pattern ( $1.3^\circ \times 1.3^\circ$  total size). The colors of the squares in each mask were randomly selected with replacement from the same set of colors used in the memory arrays. A mask was presented at the location of every item that had been presented in the memory array.

**Procedure.** As illustrated in Figure 1, each trial consisted of a 100-ms memory array followed after a blank delay by a 200-ms array of masks, another blank delay, and then a 2,000-ms test array. The blank delay between the memory and mask arrays was 17, 134, 250, 367, or 484 ms, varying randomly from trial to trial. The stimulus onset asynchronies (SOAs) were thus 117, 234, 350, 467, and 584 ms. The SOA between the memory array and the test array was always 1,000 ms, and the delay between the mask offset and the test onset therefore depended on the delay between the memory and mask arrays. The memory and test arrays were identical with the exception that on half of the trials, the color of one of the squares was different between the two arrays. The new color was selected at random from the other possible colors. Participants indicated whether the two arrays were the same or different by pressing one or the other of two

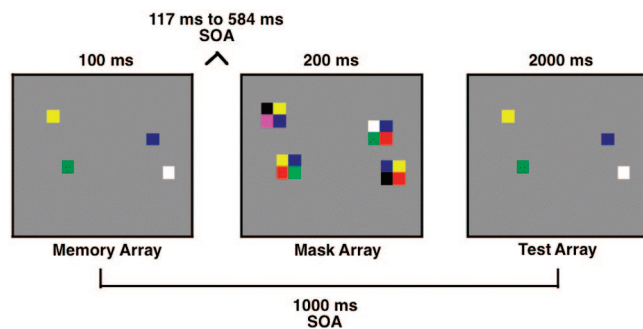


Figure 1. Example of the stimuli and timings used in Experiment 1. SOA = stimulus onset asynchrony.

buttons. Accuracy was stressed rather than speed. Participants performed 40 trials for each combination of mask delay and set size.

To rule out contributions from verbal WM in this task, we used an articulatory suppression procedure in which two randomly selected digits were presented at the center of the screen for 500 ms at the beginning of each trial. After a blank interval of 1 s, the sequence of memory, mask, and test arrays was presented. Participants were instructed to quickly say these two digits aloud repeatedly throughout the course of each trial (e.g., "three-seven, three-seven, three-seven . . ."). Previous research has shown that this is an effective method for discouraging verbal recoding of the visual stimuli (Baddeley, 1986; Besner, Davies, & Daniels, 1981).

### Results and Discussion

Figure 2 shows the observers' accuracy (percentage correct) on the visual WM task plotted as a function of the mask delay for each of the four memory-set sizes.<sup>2</sup> For arrays with only one item, performance was near ceiling across all mask SOAs. For arrays of two items, performance was slightly impaired at the shortest SOA (117 ms) but reached plateau at the 234-ms SOA. Performance for arrays of three or four items was very poor at the short SOAs and increased monotonically as the SOA increased. A two-factor analysis of variance (ANOVA) confirmed this pattern of results, yielding highly significant main effects of memory array size,  $F(3, 42) = 22.62, p < .001$ , and mask delay,  $F(4, 56) = 13.63, p < .001$ , and a highly significant interaction between these two factors,  $F(12, 168) = 49.20, p < .001$ . Separate ANOVAs were also conducted for each array size. There was no significant effect of mask delay for one-item arrays ( $F < 1$ ), but significant mask-delay effects were observed for arrays of two, three, or four items (all  $p_s < .001$ ).

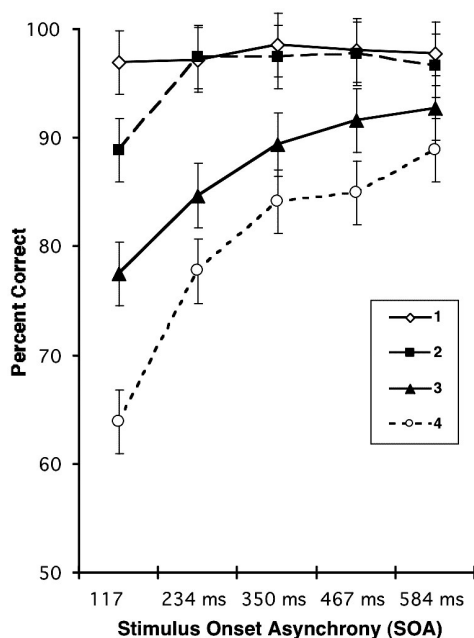


Figure 2. Results from Experiment 1. Mean accuracy for each set size plotted as a function of the stimulus onset asynchrony between the memory array and the mask array. Error bars represent 95% confidence intervals as described by Loftus and Masson (1994).

These results reveal the time course of creating durable visual WM representations. When the masks were presented shortly after the offset of the memory array, performance on the visual WM task was very poor for arrays containing several items. This effect presumably occurred because the perceptual representations of the items in the memory array were not yet consolidated and could therefore be disrupted by the masks. As the amount of uninterrupted processing time was increased, performance greatly improved, eventually reaching levels that are typical for no-mask performance (see, e.g., Vogel et al., 2001). Moreover, observers required more uninterrupted processing time to reach asymptote for larger arrays than for smaller arrays, suggesting that the consolidation process is limited in capacity. That is, the amount of time required to form a durable representation of four items is substantially longer than the amount of time required to form a durable representation of a single item. These results are consistent with the finding that the amount of dual-task slowing observed for a probe tone following a memory array is greater when the memory array contains more items (Jolicœur & Dell'Acqua, 1998).

The finding that more time is required to consolidate larger memory arrays may seem to imply that consolidation is performed in serial. However, this pattern is equally compatible with a limited-capacity parallel consolidation process, in which all of the items in the array are encoded simultaneously but each item is consolidated more slowly as the load increases. The current results cannot distinguish between these two candidate modes of processing.

Although these results appear to demonstrate that the visual WM consolidation process is extended over time and is subject to capacity limitations, several important issues remain. First, these results do not provide a direct measure of the rate of consolidation; this issue is addressed in the next experiment. Second, although the memory arrays consisted of highly discriminable simple features and were presented for 100 ms, it is conceivable (although unlikely) that some or all of the masking effects were due to a disruption of the formation of perceptual representation rather than the consolidation of these representations in WM. This possibility is addressed in Experiments 3 and 4. Finally, the differences in time course across array sizes may be influenced by decision noise. That is, the more items are present, the more decisions must be made, which will necessarily lead to lower performance. This issue is addressed in Experiment 5.

### Experiment 2

Experiment 2 was designed to quantify the rate at which consolidation occurs. The methods were similar to those of Experiment 1, except that we focused on arrays of four items and tested a fine-grained set of short mask delays (17, 33, 50, 67, and 83 ms), along with a medium delay (167 ms) and a long delay (217 ms). In

<sup>2</sup> Although we primarily quantified performance as percentage correct in this study (pooled over change and no-change trials), we also computed  $A'$  values for each experiment and found the same pattern of results. To compute the  $A'$  values, change trials were coded as signal-plus-noise trials and no-change trials were coded as noise-only trials. In this and all subsequent experiments, misses were approximately twice as common as false alarms across all conditions.

addition, we used a quantitative approach that was initially developed by Pashler (1988) and later refined by Cowan (2001). These investigators developed a simple equation for estimating the number of items,  $K$ , that is available in WM in a change-detection task. This 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. To correct for guessing, this procedure takes into account the false alarm rate. In Pashler's (1988) original formulation,  $K = [S \times (H - F) / (1 - F)]$ , where  $K$  is the memory capacity,  $S$  is the set size of the array,  $H$  is the observed hit rate, and  $F$  is the observed false alarm rate. Cowan (2001) noted that this formulation does not adequately account for guessing, and he developed a modified equation, which we use here.<sup>3</sup> His corrected formula is identical to the original equation except multiplied by the correct rejection rate ( $C$ ); that is,  $K = C \times [S \times (H - F) / (1 - F)]$  or  $K = S \times (H + C - 1)$ .

This approach was originally designed to estimate the total storage capacity of WM, but it is equally valid as a measure of how much information is actually stored in WM in a given experimental condition, even when this is substantially less than the potential capacity of WM. In the present study, we used  $K$  to estimate how many objects' worth of information have been consolidated in WM for a given mask delay. By measuring the slope of the function relating the mask delay to the number of objects' worth of information stored in WM, we can determine the rate at which information is consolidated. For example, if observers can store one object in WM with a mask delay of 50 ms and can store two objects with a mask delay of 150 ms, the consolidation rate would be one object per hundred milliseconds, or 10 objects per second.

The computation of  $K$  assumes that observers store  $K$  high-fidelity representations from an array of  $S$  items, with no information stored about the other items. This is certainly plausible, because observers can use an attention-directing cue to selectively remember a subset of the items in an array of objects, remembering almost nothing about the uncued items (Schmidt, Vogel, Woodman, & Luck, 2002; Vogel, 2000; Woodman, Vecera, & Luck, 2003). In the absence of an explicit cue, however, it is not yet known whether observers will selectively store  $K$  high-fidelity representations rather than a larger number of lower fidelity representations. Thus, when  $K$  is used to quantify performance, we refer to the number of objects' worth of information stored in WM rather than the number of objects.

## Method

The stimuli and procedures were identical to those used in Experiment 1 with the following exceptions. Ten new observers participated in the task. The set size was always four items, and the SOA between the onset of the memory array and the onset of the mask array was 117, 133, 150, 167, 183, 267, or 317 ms. Because the duration of the memory array was 100 ms, these SOAs correspond to delays of 17, 33, 50, 67, 83, 167, and 217 ms between memory array offset and mask array onset. The two longest SOAs were included so that the participants could succeed on a reasonable proportion of trials and would not give up on the task.

## Results and Discussion

Figures 3A and 3B show the observer's accuracy—quantified both as percentage correct and as  $K$ —plotted across each of the seven mask delays. Estimated visual WM storage was very low (just over one item's worth of information) at the shortest mask delay and increased in a largely linear manner over the first 200 ms of mask delay, reaching an asymptote of approximately 2.5 items' worth of information after a delay of between 200 and 300 ms. A one-way ANOVA confirmed that accuracy increased as the SOA increased,  $F(6, 48) = 15.66$ ,  $p < .001$ . Our current estimate of  $K$  asymptote being approximately 2.5 items is very close to the mean  $K$  estimate of 2.8 for no-mask performance that we have reported previously (Vogel et al., 2001).

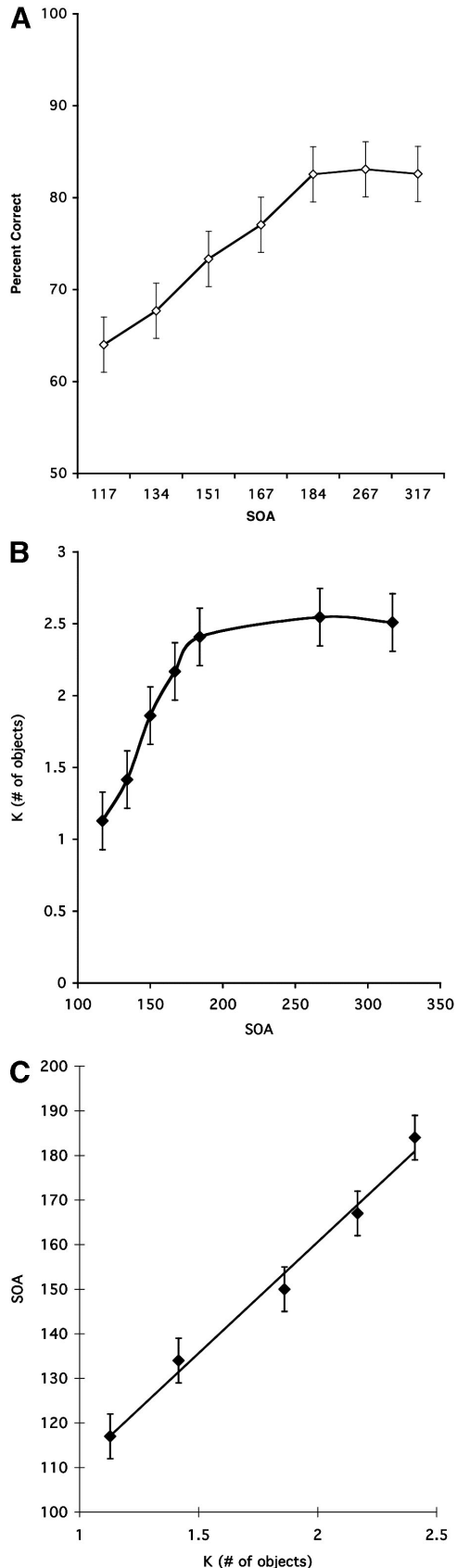
To estimate the rate of consolidation, we computed the best fit linear function for  $K$  across the range of SOAs prior to asymptote (SOAs of 117 ms to 184 ms). This function, which accounted for 99% of the variance, had a slope of 0.019 items per millisecond of SOA. That is, for every millisecond of SOA, 0.019 items were consolidated in visual WM. It is convenient to reverse the axes of this function, plotting SOA as a function of  $K$ , which then indicates the amount of time required to consolidate each item. As shown in Figure 3C, this function is nearly linear across SOAs of 117 to 183 ms, with a slope of 49.9 ms of SOA per item and a y-intercept of 60.4 ms. This estimate of approximately 50 ms per item is dramatically faster than the hundreds of ms per item that might be expected on the basis of attentional blink experiments (Chun & Potter, 1995; Raymond et al., 1992), but it is also considerably slower than the rate of 20–30 ms per item that was estimated by Gegenfurtner and Sperling (1993). We consider the reasons for these discrepancies in the General Discussion.

A straightforward interpretation of the data shown in Figure 3C is that the visual system requires an unmasked exposure of 60 ms to form a perceptual representation of the memory array (explaining the y-intercept) and then consolidates one representation every 50 ms (explaining the slope) until WM capacity is reached at around three items (explaining the asymptote). However, this interpretation must be regarded with caution. First, the y-intercept is based on extrapolation beyond the set of measured SOAs, and the function may not continue to be linear at shorter SOAs. Second, as discussed above, the consolidation process may operate in parallel rather than in serial. Nonetheless, the y-intercept of 60 ms is broadly consistent with previous studies of the timing of sensory integration between a target and a mask (Turvey, 1973), and the slope of 50 ms per item indicates that one item's worth of information is consolidated every 50 ms.

## Experiment 3

The results of the first two experiments demonstrate that presenting an array of pattern masks after a memory array can disrupt subsequent performance on a visual WM task. Although these masks were intended to selectively disrupt WM consolidation, it is

<sup>3</sup> We computed both versions of  $K$ , and the results differed only in the overall number of items that were estimated to be stored in WM. There was no difference in the effect of the mask delay on  $K$ , which is the key result of this experiment.



possible that the observed masking effects were caused by an interruption of processes other than consolidation, such as the perceptual analysis of the initial memory array. We intentionally chose stimuli (simple, suprathreshold, highly distinct colors) and timing parameters (memory array duration of 100 ms) that should have minimized the perceptual demands of the task. Nevertheless, it is conceivable that the observed masking effects were due, at least in part, to sensory integration of the memory array and the mask array or some other form of perceptual degradation. We addressed this possibility in Experiments 3 and 4.

In Experiment 3, we assessed whether there was significant sensory integration of the memory and mask arrays. Sensory integration is normally observed when the SOA between the pattern mask and target is less than 100 ms (Breitmeyer, 1984), owing to the temporal integration of information that occurs at several stages of the visual system. Although the shortest SOA in the previous experiments was 117 ms, it is still important to rule out the possibility that some integration occurred at the shortest SOAs. The sensory integration between a target and a mask is known to be equivalent for backward and forward masks (i.e., when the mask precedes the target) (Breitmeyer, 1984). Therefore, if sensory integration were contributing to these masking effects, we should see a significant effect of the masks even when they are presented shortly before (117 ms) the memory array. However, if sensory integration plays a negligible role, forward masks should be ineffective even at very short delays.

#### Method

The stimuli and procedures in this experiment were identical to those used in Experiment 2 with the following exceptions. The mask array either preceded (forward mask) or followed (backward mask) the memory array with equal probability. These two conditions were randomly intermixed within experimental blocks. The delay between the two arrays was either 17 ms or 484 ms. A new set of 10 observers participated in this experiment.

#### Results and Discussion

Figure 4 shows the observers' mean accuracy (percentage correct) on the visual WM task for the forward and backward mask conditions across the two SOAs. In the forward-mask condition, performance was high and was very similar for the two mask SOAs. This performance was also similar to no-mask performance for arrays of four items in previous studies (Vogel et al., 2001). In the backward mask condition, accuracy was very poor at the short mask delay but very high at the long delay, effectively replicating the results of Experiment 1. Accuracy in the long-SOA backward

*Figure 3.* Results from Experiment 2. A: Mean accuracy plotted as a function of the stimulus onset asynchrony (SOA) between the memory array and the mask array. B: Mean computed  $K$  values plotted as a function of the SOA between the memory array and the mask array. C: The first five SOAs between the memory and mask arrays plotted as a function of mean computed  $K$  values. A best fit linear function is drawn for  $K$  across the range of SOAs prior to asymptote. Error bars represent 95% confidence intervals as described by Loftus and Masson (1994).

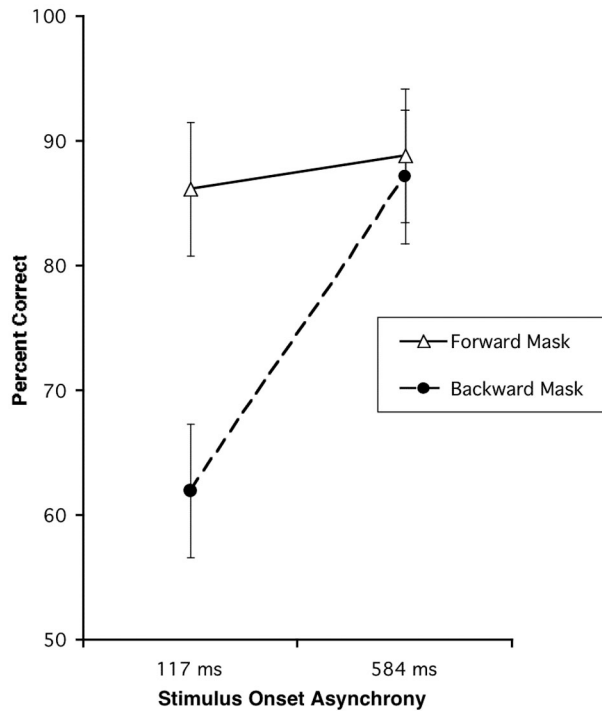


Figure 4. Results from Experiment 3. Mean accuracy for the forward and backward mask conditions plotted as a function of the stimulus onset asynchrony between the memory and mask arrays. Error bars represent 95% confidence intervals as described by Loftus and Masson (1994).

mask condition was comparable to accuracy at both SOAs in the forward mask condition.

This pattern of results led to significant main effects of both mask order,  $F(1, 9) = 31.58, p < .001$ , and mask delay,  $F(1, 9) = 63.72, p < .001$ , as well as a significant interaction,  $F(1, 9) = 24.8, p < .001$ . Of note, a planned comparison between the short- and long-SOA conditions with the forward mask did not approach significance,  $F(1, 9) = 1.83, p > .20$ . To provide further evidence of the ineffectiveness of the forward mask, we computed a one-way 95% confidence interval for the difference in accuracy between the short- and long-SOA forward mask conditions. This analysis indicated that we can be 95% confident that the effect of masking was no more than 4.36%.

The forward masks were clearly ineffective, which indicates that our stimulus parameters led to negligible sensory integration of the memory and mask arrays. Thus, the proposal of a poor sensory processing of the memory array cannot account for the large masking effects observed in the first two experiments.

#### Experiment 4

In this experiment, we sought to provide a more general test of the possibility that the masks interfered with the perceptual analysis of the memory array. To accurately perform the memory task, it is necessary for the visual system to (a) accurately determine the color of each item and then (b) consolidate these color identities into a form that can survive the mask presentation. Although the

stimulation parameters were designed to prevent the masks from interfering with the first of these two steps, it is important to rule out this possibility conclusively. In other words, we needed to demonstrate that observers can identify four colored squares in as little as 117 ms.

We tested this directly by creating a task that used the same stimuli as in the previous experiments and required that the observers identify, but not remember, the colors in the array. Specifically, observers performed a visual search task using the masked arrays of colored squares, searching for a particular target color. Accurate performance in this search task requires that the observers identify the colors of the items in the array so that they can determine whether the target is present, but it does not require that they remember the colors of each item. Accurate color identification is especially important on target-absent trials, because all of the colors in the array must be compared with the target color to determine that the target is absent. If the masking effects observed in the initial experiments were due to an inability to adequately identify the colors in the memory array, then masks presented after a short delay should significantly impair search performance. However, if the visual system can easily identify four suprathreshold colored squares within 117 ms, performance on the search task should be excellent irrespective of the mask delay (see Standing & Haber, 1968, and Potter, 1976, for previous usages of this same rationale).

#### Method

The stimuli and timing parameters used in this experiment were the same as those for the memory and mask arrays in Experiment 1, except as noted. At the beginning of each trial, a single colored square was presented at fixation for 500 ms, indicating the target color for that trial (see the upper portion of Figure 5). After a delay of 500 ms, a search array of four colored squares was presented for 100 ms, followed by an array of masks after an SOA of either 117 ms or 584 ms. Participants were required to make an unspeeded button-press response to indicate whether the target color was present or absent in the search array. The target color was present on half of the trials. A new set of 10 observers participated in this experiment.

#### Results and Discussion

The lower portion of Figure 5 shows the observers' mean accuracy (percentage correct) separately for target-present and target-absent trials and for the two mask SOAs. Mean accuracy was 95% correct or higher for both of the mask delays. Moreover, observers were highly accurate on both target-present and target-absent trials. The main effects of mask SOA and target presence did not approach significance,  $F(1, 9) = 2.09, p > .18$ , and  $F < 1$ , respectively, and neither did the interaction between these factors ( $F < 1$ ). To provide additional evidence, we computed one-way 95% confidence intervals for the difference in accuracy between the two SOAs (pooling over target-present and target-absent trials). This analysis indicated that we can be 95% confident that the effect of masking was no more than 3.21%.

These results indicate that observers could easily identify the colors of the stimuli in the stimulus arrays used in the previous experiments. Consequently, the large masking effects observed in Experiments 1 and 2 cannot be explained by an impairment in color identification. This same logic was used in the study of Potter



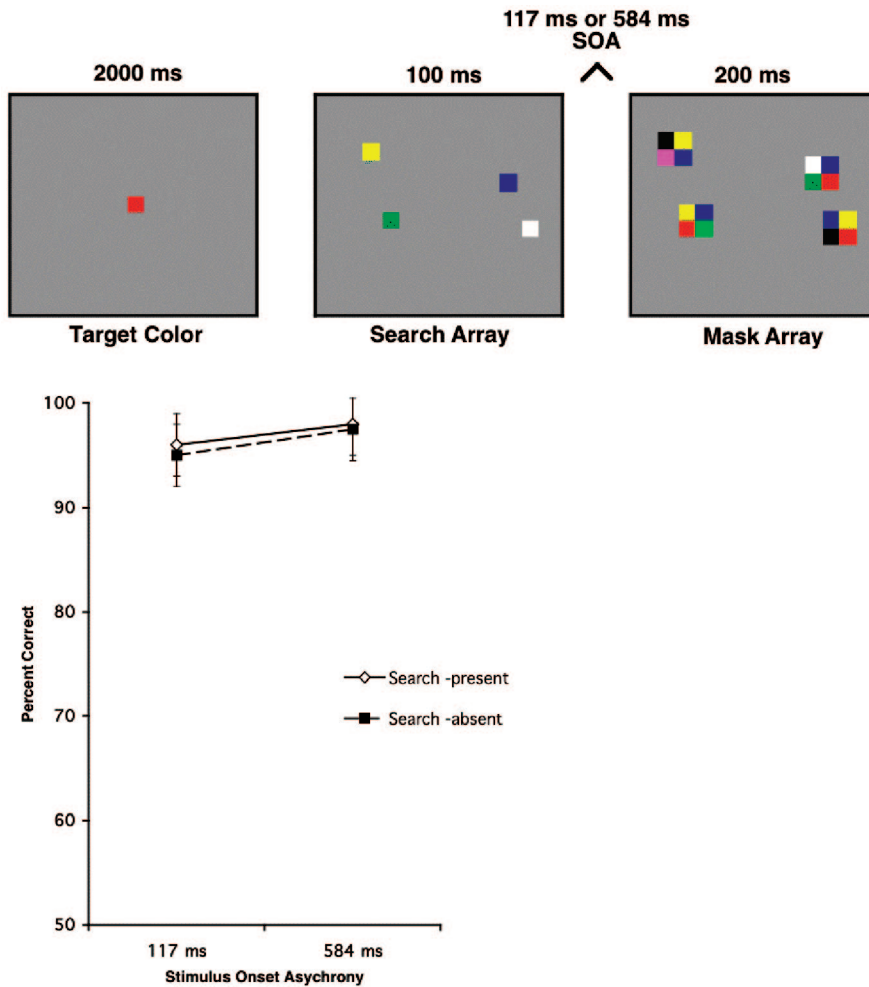


Figure 5. Example stimuli and results from Experiment 4. Mean accuracy for the target-present and target-absent conditions plotted as a function of the stimulus onset asynchrony (SOA) between the search array and the mask array. Error bars represent 95% confidence intervals as described by Loftus and Masson (1994).

(1976), which was described in detail in the introduction. Specifically, when memory for a rapid sequence of photographs was tested at the end of the sequence, performance was impaired for rapid presentation rates relative to slow presentation rates. In contrast, when the observers searched for a known target in the same stimulus sequences, performance was not substantially impaired at the fast presentation rates. From this pattern, Potter argued that the fast presentation rates led to an impairment in memory rather than an impairment in perception. In contrast, Standing and Haber (1968) compared search and memory tasks with arrays of letters to test the effects of masks presented at very short SOAs (< 100 ms) that would be expected to yield impaired perception. In this case, search and memory were equivalently impaired. Thus, previous studies have established the legitimacy of this general approach.

It is possible that the perceptual requirements of the search task used in this experiment were not as great as those of the memory task used in Experiments 1 and 2. Specifically, the search task required the observers to compare each item in the array with a

single, known color, whereas the memory task required the observers to be able to distinguish between each item in the memory array and each of the other possible colors in the color set. It is therefore possible that somewhat more precise perceptual representations were necessary for the memory task than for the search task. However, it is extremely unlikely that this subtle difference could explain the finding that observers performed the search task at approximately 95% correct with a 117-ms SOA but performed the memory task at approximately 65% correct with this same SOA. Thus, together with the results of Experiment 3, the present results indicate that the masking effects observed in Experiments 1 and 2 were primarily caused by an interruption of processing after the colors had been identified.

### Experiment 5

The results of Experiment 1 suggest that arrays with larger numbers of items take longer to consolidate than smaller arrays because observers require more uninterrupted processing time to

reach asymptotic performance levels with larger arrays. However, the poor performance observed for the larger array sizes at short mask delays may be at least partially the result of an accumulation of errors at the decision stage rather than solely being due to the time necessary for WM consolidation. Specifically, this memory task requires the observer to compare the remembered identity for each item in the memory array with each of the items in the test array. Each of these item-level decisions is subject to some probability of error, with the amount of error being dependent on the fidelity of the initial representation and the confusability of the colors. For example, if the masks interrupted the processing of the memory array before a high-fidelity representation of each of the colors was reached, the probability of making an error when determining whether a given item had changed would be nonnegligible. Because the task requires the observers to make a single response for the entire array, the likelihood of making an error necessarily increases at larger array sizes. From the accumulation of these potential decision errors, performance would be expected to be poorer with larger array sizes at short mask SOAs even without assuming any capacity limitations in the consolidation process (Palmer, Ames, & Lindsey, 1993; Shaw, 1982).

In Experiment 5, we evaluated the role of decision errors in Experiment 1 by using a partial-report procedure that allows the observer to make a single decision on each trial (Hawkins et al., 1990; Vogel et al., 2001). Specifically, one item was postcued in the test array at the end of each trial, and the participants were instructed that the cued item was the only item that could possibly have changed between the two arrays. The participant did not know which item would be cued until the onset of the test array, and it was therefore necessary to store the entire memory array in visual WM. However, when a decision was made at the end of the trial about whether an item had changed color, the decision process could be restricted to the cued location. This postcue condition was compared with a no-cue condition to determine whether limiting decisions to a single item would improve performance for a large memory array. If the differences between the array sizes observed in Experiment 1 were due to accumulated decision errors, then performance should be substantially better in the postcue condition than in the no-cue condition. However, if decision errors did not significantly affect performance, then there should be little or no difference between the postcue and no-cue conditions.

In a third cuing condition, we sought to demonstrate that participants could indeed use a cue of this nature if it was presented at the time of the memory array. As in the postcue condition, the cue in the precue condition indicated the only item that could possibly differ between the memory and test arrays, but in this situation the cue made it possible to limit both storage and decision to a single item. That is, if the participants are capable of using the precue, they need only consolidate a single item in visual WM, which should result in more accurate performance even at a short mask SOA. However, if the participants are unable to use the cue, performance should be equivalent to the no-cue condition.

### Method

The stimuli and procedures used in this experiment were identical to those used in Experiment 1, with the following exceptions. The memory

arrays contained four items, and the SOA between the memory and mask arrays was either 117 ms or 584 ms. The no-cue condition was otherwise identical to Experiment 1. In the postcue condition, an outline square was placed around one randomly selected item in the test array to indicate which item might have changed. In the precue condition, the outline square was placed around one item in the memory array. Participants were instructed that an item would change color between the memory and test arrays on 50% of trials but that if an item did change, it was always the cued item. These three conditions were mixed within a single block of 240 trials.

### Results and Discussion

Figure 6 shows accuracy as a function of mask SOA for the three cuing conditions. At the 117-ms SOA, performance was very poor for the postcue and no-cue conditions but was considerably better for the precue. At the 584-ms SOA, performance was substantially improved for all three cuing conditions, reaching near ceiling levels in the precue condition. This pattern of results yielded significant main effects of cue condition,  $F(1, 9) = 23.05, p < .001$ , and mask SOA,  $F(1, 9) = 106.92, p < .001$ , but the interaction was not significant,  $F(2, 18) = 1.52, p > .24$ .

Planned follow-up ANOVAs were also performed, comparing the no-cue condition with the precue condition and with the postcue condition. Although accuracy was slightly greater in the postcue condition than in the no-cue condition, the ANOVA comparing these two conditions did not yield a significant effect of cuing condition ( $F < 1$ ). In contrast, the effect of cuing condition was significant when the precue and no-cue conditions were compared,  $F(1, 9) = 28.93, p < .001$ , confirming that participants can indeed use this type of cue.

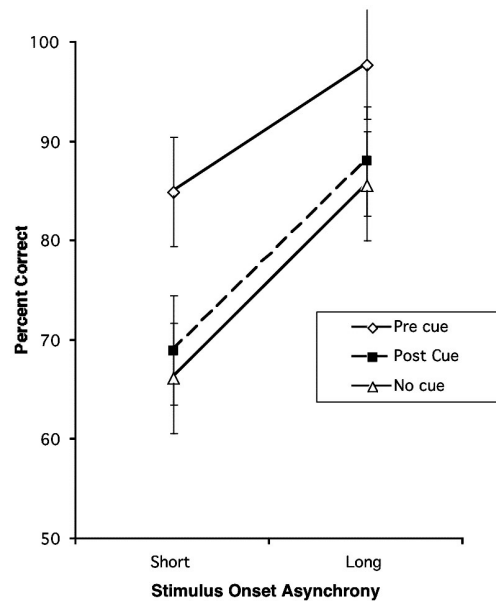


Figure 6. Results from Experiment 5. Mean accuracy for the three cuing conditions plotted as a function of the stimulus onset asynchrony between the memory and mask arrays. Error bars represent 95% confidence intervals as described by Loftus and Masson (1994).

These results indicate that the poor performance for large array sizes as compared with small array sizes in Experiment 1 appears to be due primarily to capacity limitations in the consolidation process rather than an accumulation of errors at the decision stage. Accuracy was slightly (but not significantly) better when the participants could limit decision processes to a single item in the test array, and we cannot rule out the possibility that accuracy is influenced somewhat by decision noise at larger set sizes in this paradigm. Indeed, decision noise must influence accuracy when multiple noisy decisions are made. However, the effects observed in the present experiment were far too small to fully account for the very large effect of set size that was observed at short mask delays in Experiment 1, and most of the set size effect must therefore reflect capacity limitations.

## General Discussion

### *Capacity Limitations in WM Consolidation*

The primary goal of this study was to determine how long it takes the visual system to create durable WM representations of simple stimuli and whether this process is limited in capacity. To accomplish this, we used masks to interrupt processing and measured how much information had reached WM after various amounts of uninterrupted processing time. In the first experiment, we found that larger memory arrays required more uninterrupted processing time than smaller arrays to reach asymptotic performance. This finding supports the general proposal that consolidation is limited in capacity, because each additional item added to the memory array caused a substantial increase in the amount of time required to reach asymptotic performance. Indeed, this pattern of increasing processing time for larger memory arrays is exactly what Joliceur and Dell'Acqua (1998) found with a very different procedure. The reasons for the capacity limitation are not currently known, but a possibility is discussed at the end of this article.

It may be natural to assume that the present results imply that one item is consolidated at each moment in time, as proposed by Standing and Haber (1968). However, a limited-capacity process that consolidates the memory items in parallel could also easily account for the current results, if the consolidation process simply takes longer as more items are being simultaneously encoded into visual WM. Additional research will be necessary to distinguish between these possibilities. For example, event-related potentials have recently been used to demonstrate that attention shifts from object to object in serial during certain difficult visual search tasks (Woodman & Luck, 1999, 2003b), and the same basic method might be able to determine whether the consolidation process is serial or parallel.

Similarly, it may be natural to assume that a given object is either consolidated fully or not at all. That is, observers may either create a high-fidelity representation of a given object in WM or fail to create any representation of that object. However, when presented with four items to be consolidated and a mask 117 ms later, leading to an accuracy level of 77% correct, it is possible that all four items were partially consolidated but with

low fidelity. As discussed by Vogel et al. (2001), we do not yet know whether visual WM stores a set of approximately three discrete, fixed-resolution object representations or a potentially large set of variable-resolution representations, and  $K$  should be treated as an estimate of the number of objects' worth of information stored in visual WM (see also Alvarez & Cavanagh, 2004). Thus, the present estimate of consolidation rate is an estimate of the number of objects' worth of information consolidated per unit of time.

### *The Rapid Time Course of Consolidation*

In Experiment 2, we measured the consolidation time for a memory array of four items by calculating the slope of  $K$  until it began to reach an asymptote of approximately three items at an SOA of about 200 ms. The slope of this consolidation function was found to be approximately 50 ms per item, which is dramatically faster than we expected on the basis of attentional blink experiments. We expected that the consolidation rate would be closer to 500 ms per item, an order of magnitude slower than the rate we observed. Our results suggest that consolidation is actually a relatively rapid process that can form about three durable visual WM representations within the first 200 ms of visual processing. This is comparable to typical estimates of the rate of attentional scanning in visual search (Treisman, 1988; Wolfe, 1994; Woodman & Luck, 1999), which raises the possibility that a common process underlies both search and consolidation. However, visual search rates vary widely as a function of the search stimuli, and the consolidation rate may also vary as a function of the nature of the targets and masks. Thus, it would be premature to propose a link between search and consolidation on the basis of the present findings.

There are several potential reasons for the large discrepancy between the observed consolidation rate and the rate we expected on the basis of attentional blink experiments. As discussed in the introduction, the hypothesis of slow consolidation is based on dual-task paradigms in which processes other than consolidation may strongly influence performance. Our 50 ms/item estimate, in contrast, was derived from a single-task procedure and may more directly measure consolidation time without the potentially deleterious influence of a second competing task.

A second potential reason for the large discrepancy in the consolidation estimates is that in most dual-task paradigms (e.g., the attentional blink) the targets are letters and numbers, as opposed to our use of simple colored objects. Alphanumeric characters are more visually complex objects than our simple colored squares, and it is possible that complex objects may take longer to consolidate in visual WM. However, this is unlikely because Ross and Joliceur (1999) found a normal-duration attentional blink (several hundred milliseconds) for color targets. Moreover, Gegenfurtner and Sperling (1993) used alphanumeric characters in a masking experiment and estimated that they were transferred into WM at a rate of 20–30 ms per item. This estimate is considerably faster than our current

estimate.<sup>4</sup> Thus, reduced stimulus complexity cannot easily explain the very fast consolidation rate observed in the present experiment.

An important third difference between the attentional blink paradigm and the present paradigm is that all of the targets (i.e., memory items) in the present paradigm are presented simultaneously, whereas the two targets are presented sequentially in the attentional blink paradigm. This difference may affect the consolidation rate in a number of ways. For example, the total consolidation time may depend on the number of events to be consolidated, with simultaneously presented targets being treated as a single event and sequentially presented targets treated as multiple events.

Another potential difference between simultaneous and sequential target presentations is that in the sequential presentation procedure it is necessary to maintain the representation of the first target in visual memory while the second target is being consolidated. It is possible that visual WM maintenance and consolidation require the same resources. Indeed, Logie's (1995) model of visual WM proposes that a single process performs both of these functions. If the maintenance process becomes less demanding over time, then the attentional blink could be explained by interference between the maintenance of the first target and the consolidation of the second target.

### *Generalizing to Other Stimuli*

The present estimate of 50 ms/item may depend on the nature of both the stimuli being consolidated and the masks that are used to interrupt consolidation. For example, although the storage capacity of visual WM is just as great for multifeature targets as for single-feature targets (e.g., Vogel et al., 2001; Wheeler & Treisman, 2002), the process of forming WM representations may be greater when more features must be linked together. Alternatively, the limitation on consolidation may be the same as the limit on storage, leading to identical consolidation rates for single-feature and multiple-feature objects. It is also possible that the consolidation rate for simple colored squares depends on the similarity among the colors (owing to greater interference). The consolidation rate may also depend on the degree of precision the observers expect to need at the time of test. That is, if the observers must detect very subtle changes in color between the memory and test arrays, they may need to form a higher resolution representation of each item, which may require more time. These possibilities would be relatively easy to address with the methods developed in the present study.

Given that the consolidation rate may vary widely as a function of these factors, what is the value of the present estimate of 50 ms/item? There are several answers to this question. The first is that the 50 ms/item value applies to a type of stimulus that has been widely examined in recent studies of visual WM (Jiang et al., 2000; Vogel et al., 2001; Wheeler & Treisman, 2002; Woodman et al., 2001, 2003; Xu, 2002b). The second is that even if the present estimate applies only to a narrow set of stimulus parameters, it is still faster than might have been expected on the basis of attentional blink studies (including those using color targets). The third is that it provides a starting point for assessing the consolidation rate across a broad range of stimulus parameters, which in turn

may provide important information about the nature of the consolidation process.

It is also worth considering how the consolidation rate may vary as a function of the number of parts in the object. Xu (2002a) has shown that visual WM storage capacity is just as great for multipart objects as for single-part objects, but only if observers are asked to store different feature dimensions from the different parts (e.g., the color of one part and the orientation of a different part). As in the case of multifeature single-part objects, it is possible that the consolidation rate would be higher for multipart objects than for single-part objects, even if different feature dimensions are stored for the different parts. When the same dimension must be stored for each part, each part may be stored in a separate "slot" in WM (see Alvarez & Cavanagh, 2004), and the consolidation of a two-part object may take the same amount of time as the consolidation of two single-part objects. This is an easily testable hypothesis.

It is also possible that the consolidation rate varies across individuals as a function of factors such as age, experience, and pathology. Indeed, a recent study using the methods developed here found that consolidation was dramatically slowed in patients with schizophrenia (Fuller, Luck, McMahan, & Gold, in press).

### *Comparison With Other Varieties of Masking*

The pattern masks that we used in this procedure appear to have selectively interrupted consolidation rather than sensory or decision processes. In Experiment 3, we tested for target-mask integration at an early sensory level by presenting the masks before the memory array, and we observed negligible sensory integration. In Experiment 4, we showed that visual search is unimpaired even at very short mask delays for these stimuli, which indicates that the masks did not limit the observer's ability to identify the colors. In Experiment 5, we demonstrated that accumulated decision errors do not play a major role in this paradigm by demonstrating that performance is not improved when the observer makes only a single decision for the entire array. These three results indicate that our paradigm effectively isolates intermediate-level cognitive processes that follow perceptual processes but precede decision processes.

The masking in the present paradigm is clearly different from the commonly studied varieties of perceptual masking, such as metacontrast, paracontrast, and integration masking. These varieties of masking would not be expected to occur with the stimulus parameters used in the present paradigm, and if they had occurred, they would have produced impaired perceptual performance in Experiments 3 and 4.

---

<sup>4</sup> Gegenfurtner and Sperling's estimate of 20–30 ms per item is notably faster than our estimate of approximately 50 ms per item. There may be a number of factors that contribute to this discrepancy. For example, they used alphanumeric characters, which are highly overlearned and may take less time to consolidate. Their observers also received thousands of trials, which may have increased the consolidation rate. In addition, the items in their arrays were presented in fixed positions, which would allow observers to focus spatial attention in advance of the display onset, whereas our items were presented at random locations for each trial and would therefore not facilitate spatial attention.

Enns and Di Lollo (Di Lollo, Enns, & Rensink, 2000; Enns & Di Lollo, 1997, 2000) have recently described a new variety of masking, called object-substitution masking, which appears to operate after objects have been perceived. In their experimental paradigm, the target item is surrounded by four small dots that remain visible after the target has terminated. These four dots dramatically impair the observers' ability to report the identity of the target, but only when the target is embedded in a large array of distractors. They propose that iterative reentrant processing between high and low cortical visual areas is necessary for the target to be represented in a reportable form. In their view, masking occurs when the initial representation of the target mismatches the low-level signal provided by the trailing mask. Because the mask signal is stronger than the fading target representation, the mask may overwrite or "substitute" for the target representation. We have recently provided electrophysiological support for this hypothesis, demonstrating that the masked target item is identified and even triggers a shift of attention (Woodman & Luck, 2003a). However, attention arrives after the target has been extinguished, so that the attended mask representation can overwrite the original target representation, leading to inaccurate behavioral performance.

Object-substitution masking shares two important properties with the variety of masking observed in the present experiment. First, both types of masking appear to operate after the initial identification of the masked information. Second, both types of masking have larger effects at larger set sizes. However, we suspect that the mechanisms of masking are not identical. First, object-substitution masking occurs only when the target discrimination task is attention demanding, and it would not be expected to occur for the simple colors used in the present study. Second, object-substitution masking leads to impaired target detection performance in search tasks, but we found no significant masking effects in the search task of Experiment 4. Third, our phenomenological experience differs between the two tasks: In object-substitution masking, the masked location appears to be completely empty, whereas in the present paradigm, observers report seeing the masked items clearly, even though they cannot remember their colors. Thus, although both varieties of masking may operate after the initial wave of perceptual processing, we propose that object-substitution masking operates at an earlier point than the masks used in the present study.

The variety of masking examined in the present study may be identical to the conceptual masking that has been described in picture-memory studies (e.g., Potter, 1976). In the domain of picture processing, conceptual masking is said to occur when a perceptual representation of a scene has been formed but is masked before it is transformed into a more durable representation. This is exactly the sort of postperceptual masking that we suggest occurs for simple colored squares in the present study. Conceptual masking has a similar time course to the masking observed in the present study, and it has also been shown to occur under conditions that do not produce perceptual masking (Loftus & Ginn, 1984; Potter, 1976). Thus, the consolidation process that has been quantified in the present study is probably also used for complex, natural scenes as well as for simple, artificial stimuli.

### *A Scheme for Consolidation and Masking*

In this final section, we outline a model of the masking and consolidation processes that we believe operate in the present study. When the memory array is presented, the visual system performs a perceptual analysis at several levels, leading to perceptual and conceptual representations in a variety of perceptual and postperceptual systems (e.g., within extrastriate areas of the occipital lobe, within the inferotemporal cortex, within the medial temporal lobe, within dorsolateral prefrontal cortex). These representations will passively decay over a period of a few hundred milliseconds unless a limited-capacity consolidation process is applied to the representations. The consolidation process stabilizes a subset of these representations, allowing the consolidated representations to be maintained over long periods of time (e.g., tens of seconds) and to survive the presentation of new stimuli.

Once the consolidation process has occurred, the representations are said to be stored in WM, but we do not mean to imply that the representations have somehow been moved from one place to another. Instead, the consolidation process may change the nature of existing representations. Specifically, we have previously proposed that visual WM maintenance relies on establishing correlated firing among the neurons that are already active as a part of the high-level perceptual representation of a given object (Luck & Vogel, 1998). A neural network simulation demonstrated that this general scheme can account for various aspects of visual WM performance, such as the ability to maintain three multifeature object representations over periods of seconds (Raffone & Wolters, 2001). This simulation did not directly address the time course of the creation of correlated firing, nor did it examine whether correlated firing would allow the representations to survive the presentation of new stimuli. However, it is plausible that the creation of independent correlations for multiple objects would be time consuming, and the stabilizing influence of correlated firing should make the representations less susceptible to interference from new information.

Neurophysiological studies indicate that representations at early and intermediate stages of the visual system (e.g., from the retina through occipital cortex) cannot be maintained beyond a few hundred milliseconds, and these studies have demonstrated that maintenance of object identity information occurs primarily in inferotemporal and prefrontal regions (Fuster & Jervey, 1982; Miller & Desimone, 1994; Miller, Erickson, & Desimone, 1996). Moreover, both single-unit and neuroimaging studies indicate that the representations in inferotemporal cortex do not survive the presentation of new stimuli, whereas the representations in prefrontal cortex do survive (Miller, Gochin, & Gross, 1993; Miller, Li, & Desimone, 1993). This suggests that the consolidation process may be limited to representations in prefrontal cortex. However, these studies used task-relevant stimuli to interrupt the WM representations, and it is possible that inferotemporal WM representations can survive the presentation of task-irrelevant information, such as the masks used in the present study. Indeed, psychophysical studies have found larger conceptual masking effects for task-relevant masks than for task-irrelevant masks (Intraub, 1981, 1984). Thus, it is not yet clear whether consolidation operates within inferotemporal cortex.

In the present paradigm, we propose that the masks receive extensive perceptual processing, and the perceptual representations of the masks simply overwrite the perceptual representations of the items in the memory array. If a given memory item has not been consolidated before it is overwritten, it will not be available for comparison with the test array, leading to errors in change-detection performance. At this point, we are agnostic about whether consolidation is an all-or-none or a continuous process. That is, if the mask occurs while an item is undergoing consolidation, we do not know whether this will lead to no information or to partial information about the item in WM. However, there is some evidence that sudden-onset objects such as masks are automatically stored in visual WM (Schmidt et al., 2002), and this suggests that an incomplete representation of a memory array item will be replaced in WM by a representation of the mask.

This also raises the possibility that masks used in the present study do not simply erase the perceptual representation of the to-be-remembered objects before they have been consolidated but instead compete with the to-be-remembered objects for WM storage capacity. By itself, this hypothesis cannot explain the present results, because equivalent competition would be expected at all mask delays. However, it is possible that the similarity between the memory items and the masks makes it difficult for attentional processes to select only the memory items for storage in WM at short delays. Longer delays may make the masks more distinctive from the memory items, yielding better selectivity for storage in WM. We have collected some preliminary evidence showing that the similarity between the memory items and the masks does influence the degree of masking in some cases, but we also found that many types of dissimilar masks are still highly effective (E. K. Vogel, A. T. Niese, & S. J. Luck, unpublished observations). This would argue against the possibility that the results of the present study are due to competition for memory storage combined with a failure of attentional selectivity at short delays.

A related possibility is that no consolidation is necessary unless masks are present. That is, it may be possible to encode and retain information in visual WM without consolidation or vulcanization of the WM representations, and these latter processes might be important only when subsequent stimuli might interfere with the WM representations. However, when humans view natural scenes, they typically make saccadic eye movements every 300–400 ms, and these eye movements presumably serve as highly effective masks. Moreover, the present estimate of the consolidation rate fits well with the typical timing of eye movements during scene perception: Given a visual WM capacity of four items and a consolidation rate of 50 ms/item, 300 ms should be just enough time to conduct a perceptual analysis of the currently fixated region and fill WM to capacity before initiating the next saccade. Thus, visual WM consolidation is likely to be an important and frequently used process in real-world vision.

### Concluding Comments

The present study has shown that the consolidation process is limited in capacity, which is consistent with previous findings, and it has also shown that consolidation is faster than might have been expected. More important, however, this study has provided a set of methods for measuring the consolidation rate and ruling out

contributions from factors such as perceptual masking and accumulated decision noise. We hope that other investigators will use these methods to study the consolidation process in greater detail.

### References

- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by total information load and by number of objects. *Psychological Science, 15*, 106–111.
- Anderson, R. S., & Thibos, L. N. (1999). Sampling limits and critical bandwidth for letter discrimination in peripheral vision. *Journal of the Optical Society of America, 16*, 2334–2342.
- Baddeley, A. D. (1986). *Working memory*. Oxford, England: Clarendon.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.
- Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology, 33A*, 415–437.
- Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.
- Broadbent, D. E., & Broadbent, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics, 42*, 105–113.
- Brown, L., Govan, E., & Block, M. T. (1983). The effect of reduced visual acuity upon Farnsworth 100-hue test performance. *Ophthalmic and Physiological Optics, 3*(1), 7–11.
- 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. (1997). *Attention and memory*. New York: Oxford University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.
- Dell'Acqua, R., & Jolicœur, P. (2000). Visual encoding of patterns is subject to dual-task interference. *Memory & Cognition, 28*, 184–191.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*, 481–507.
- Duncan, J., Ward, R., & Shapiro, K. (1994, May 26). Direct measurement of attentional dwell time in human vision. *Nature, 369*, 313–315.
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8*, 135–139.
- Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking. *Trends in Cognitive Sciences, 4*, 345–351.
- Fuller, R. L., Luck, S. J., McMahon, R. P., & Gold, J. M. (in press). Working memory consolidation is abnormally slow in schizophrenia. *Journal of Abnormal Psychology*.
- Fuster, J. M., & Jervey, J. P. (1982). Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *Journal of Neuroscience, 2*, 361–375.
- Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 845–866.
- Giesbrecht, B. L., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1454–1466.
- Hawkins, H. L., Hillyard, S. A., Luck, S. J., Mouloua, M., Downing, C. J., & Woodward, D. P. (1990). Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 802–811.

- Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 604–610.
- Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 115–125.
- Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125–155). Cambridge, MA: MIT Press.
- Isaac, M. I., Shapiro, K. L., & Martin, M. J. (1999). The attentional blink reflects retrieval competition among multiple RSVP items: Tests of the interference model. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1774–1792.
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2, 683–702.
- Jolicoeur, P. (1998). Modulation of the attentional blink by on-line response selection: Evidence from speeded and unspeeded Task<sub>1</sub> decisions. *Memory & Cognition*, 26, 1014–1032.
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36, 138–202.
- Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception & Psychophysics*, 63, 253–257.
- Loftus, G. R., & Ginn, M. (1984). Perceptual and conceptual masking of pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 435–441.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476–490.
- Logie, R. H. (1995). *Visuo-spatial working memory*. Hove, England: Erlbaum.
- Luck, S. J., & Vogel, E. K. (1997, November 20). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- Luck, S. J., & Vogel, E. K. (1998). Response from Luck and Vogel (response to commentary by Nelson Cowan). *Trends in Cognitive Sciences*, 2, 78–80.
- Miller, E. K., & Desimone, R. (1994, January 28). Parallel neuronal mechanisms for short-term memory. *Science*, 263, 520–522.
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, 16, 5154–5167.
- Miller, E. K., Gochin, P. M., & Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque monkey by addition of a second stimulus. *Brain Research*, 616, 25–29.
- Miller, E. K., Li, L., & Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience*, 13, 1460–1478.
- Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 108–130.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, 44, 369–378.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, 16, 283–290.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 509–522.
- Potter, M. C., Chun, M. M., Banks, B. S., & Muckenhoupt, M. (1998). Two attentional deficits in serial target search: The visual attentional blink and an amodal task-switch deficit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 979–992.
- Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, 13, 766–785.
- 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, 849–860.
- Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review*, 93, 180–206.
- Ross, N. E., & Jolicoeur, P. (1999). Attentional blink for color. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1483–1494.
- Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and involuntary attentional control of visual working memory. *Perception & Psychophysics*, 64, 754–763.
- Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink: A view on attention and a glimpse on consciousness. *Trends in Cognitive Science*, 1, 291–296.
- Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 357–371.
- Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology*, 14, 353–409.
- Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 591–600.
- Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33, 5–42.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74(Whole No. 498).
- Standing, L., & Haber, R. N. (1968). Visual search and memory under degraded and masked presentation. *Psychonomic Science*, 13, 81–82.
- Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, 40, 201–237.
- Turvey, M. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, 80, 1–52.
- Vogel, E. K. (2000). *Selective storage in visual working memory: Distinguishing between perceptual-level and working memory-level mechanisms of attention*. Unpublished doctoral dissertation, University of Iowa, Iowa City.
- Vogel, E. K., & Luck, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychonomic Bulletin & Review*, 9, 739–743.
- Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1656–1674.
- 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, 92–114.
- Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, 30, 79–109.
- Weichselgartner, E., & Sperling, G. (1987, November 6). Dynamics of automatic and controlled visual attention. *Science*, 238, 778–780.
- Wheeler, M., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131, 48–64.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238.

- Woodman, G. F., & Luck, S. J. (1999, August 26). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, *400*, 867–869.
- Woodman, G. F., & Luck, S. J. (2003a). Dissociations among attention, perception, and awareness during object-substitution masking. *Psychological Science*, *14*, 605–611.
- Woodman, G. F., & Luck, S. J. (2003b). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 121–138.
- Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, *10*, 80–87.
- Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science*, *12*, 219–224.
- Xu, Y. (2002a). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics*, *64*, 1260–1280.
- Xu, Y. (2002b). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 458–468.

Received December 15, 2002

Revision received October 10, 2003

Accepted November 3, 2003 ■

### New Editors Appointed, 2008–2013

The Publications and Communications Board of the American Psychological Association announces the appointment of six new editors for 6-year terms beginning in 2008. As of January 1, 2007, manuscripts should be directed as follows:

- *Behavioral Neuroscience* ([www.apa.org/journals/bne](http://www.apa.org/journals/bne)), **Ann E. Kelley, PhD**, Department of Psychiatry, University of Wisconsin–Madison Medical School, 6001 Research Park Boulevard, Madison, WI 53719.
- *Journal of Experimental Psychology: Applied* ([www.apa.org/journals/xap](http://www.apa.org/journals/xap)), **Wendy A. Rogers, PhD**, School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332-0170.
- *Journal of Experimental Psychology: General* ([www.apa.org/journals/xge](http://www.apa.org/journals/xge)), **Fernanda Ferreira, PhD**, The School of Philosophy Psychology and Language Sciences, The University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, United Kingdom.
- *Neuropsychology* ([www.apa.org/journals/neu](http://www.apa.org/journals/neu)), **Stephen M. Rao, PhD**, Division of Neuropsychology, Medical School of Wisconsin, 8701 West Watertown Plank Road, Medical Education Building, Room M4530, Milwaukee, WI 53226.
- *Psychological Methods* ([www.apa.org/journals/met](http://www.apa.org/journals/met)), **Scott E. Maxwell, PhD**, Department of Psychology, University of Notre Dame, Notre Dame, IN 46556.
- *Psychology and Aging* ([www.apa.org/journals/pag](http://www.apa.org/journals/pag)), **Fredda Blanchard-Fields, PhD**, School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332-0170.

**Electronic manuscript submission.** As of January 1, 2007, manuscripts should be submitted electronically via the journal's Manuscript Submission Portal (see the Web site listed above with each journal title).

Manuscript submission patterns make the precise date of completion of the 2007 volumes uncertain. Current editors, John F. Disterhoft, PhD, Phillip L. Ackerman, PhD, D. Stephen Lindsay, PhD, James T. Becker, PhD, Stephen G. West, PhD, and Rose T. Zacks, PhD, respectively, will receive and consider manuscripts through December 31, 2006. Should 2007 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2008 volumes.