Journal of Experimental Psychology:
Human Perception and Performance
2012, Vol. 38, No. 2, 429–438

0096-1523/12/$12.00 DOI: 10.1037/a0023925

Parallel Consolidation of Simple Features Into Visual Short-Term Memory

Irida Mance
University of Oregon

Mark W. Becker and Taosheng Liu
Michigan State University

Although considerable research has examined the storage limits of visual short-term memory (VSTM), little is known about the initial formation (i.e., the consolidation) of VSTM representations. A few previous studies have estimated the capacity of consolidation to be one item at a time. Here we used a sequential-simultaneous manipulation to reexamine the limits of consolidating items into VSTM. Participants viewed briefly presented and masked color patches (targets), which were shown either sequentially or simultaneously. A probe color followed the targets and participants decided whether it matched one of the targets or was a novel color. In four experiments, we consistently found equal performance for sequential and simultaneous presentations for two targets. Worse performance in the simultaneous than the sequential condition was observed for larger set sizes (three and four). Contrary to previous results, suggesting a severe capacity limit of one item, our results indicate that consolidation into VSTM can occur in parallel and without capacity limits for at least two items.

Keywords: visual short-term memory, consolidation capacity, attention

Many tasks rely on our ability to utilize visual information after the sensory stimulus is no longer visually available. A long history of research has shown that sensory perceptions are quite volatile and must be stabilized in more durable forms if they are to survive interference from new information (Sperling, 1960). This storage of information has been shown to rely on a form of visual memory known as visual short-term memory (VSTM; Phillips, 1974; Hollingworth & Luck, 2008). While there is a vast literature on the topic of VSTM, most of the research has investigated storage capacity limits (Luck & Vogel, 1997; Phillips, 1974), how the resolution of items is influenced by memory load (Alvarez & Cavanagh, 2004; Awh, Barton, & Vogel, 2007; Vogel, Woodman, & Luck, 2001), and how individual items are selected for VSTM when multiple items compete (Schmidt, Vogel, Woodman, & Luck, 2002; Woodman & Vogel, 2008). Although much has been learned about these aspects of VSTM, surprisingly little is known about the actual processes involved in the initial formation of VSTM representations, that is, the consolidation of items into VSTM.

The few studies that have investigated the creation of working memory representations have found that the consolidation process also has a limited capacity or bandwidth (Jolicoeur & Dell'Acqua, 1998; Vogel et al., 2006). These studies have found that as the number of to-be-consolidated items increases, so does the time required to consolidate them into durable representations. For example, in a change detection experiment where the sample array was masked after variable intervals (Vogel et al., 2006), performance was very high even at very short array-mask intervals when the array contained a single item. However, as the number of items in the array increased, performance with short array-mask intervals began to suffer, indicating that the time required to consolidate the sample items into VSTM increased with the size of the array (Vogel et al., 2006).

Though the findings of Jolicoeur and Dell'Acqua (1998) and Vogel et al. (2006) demonstrate limits to the number of items that may be consolidated concurrently, they do not address whether this limit results because consolidation occurs serially or is a limited capacity parallel process; both serial and limited capacity parallel processes can lead to longer consolidation times for multiple items (Townsend, 1990). The goal of the present study was to clarify exactly how many items individuals are able to concurrently process when creating VSTM representations. That is, when selecting items for consolidation into VSTM, are individuals restricted to processing only one item at a time, or is it possible to select and process multiple items concurrently?

Recent research on the consolidation of simple items into working memory suggests that the consolidation process is a serial process that can only operate on a single object at a time. West, Pun, Pratt, and Ferber (2010) found extreme capacity limits for the consolidation of salient pop-out items. During their experiments, participants were presented with two successive arrays of pop-out and distracter objects and had to indicate whether a change occurred between the two successive displays. The results showed that as the number of pop-out targets increased, the participants' sensitivity to a change in one of the pop-out items declined. For the majority of participants, sensitivity significantly declined between one and two pop-out items, while for a few others, a significant decline occurred between two and three. This led the researchers to conclude that the automatic processing of salient pop-out items is

This article was published Online First May 30, 2011.

Irida Mance, Department of Psychology, University of Oregon, Eugene, Oregon; Mark W. Becker, Department of Psychology, Cognitive Science Program, Michigan State University, East Lansing, Michigan; Taosheng Liu, Department of Psychology, Cognitive Science Program, Neuroscience Program, Michigan State University.

We thank Joshua Kramer for assistance in data collection. This research is supported in part by a grant from the Provost Undergraduate Research Initiative (PURI) from Michigan State University (I.M.).

Correspondence concerning this article should be addressed to Taosheng Liu, Department of Psychology, Michigan State University, East Lansing, MI 48824. E-mail: tsliu@msu.edu

limited to one item at a time, with an upper bound of two for high-capacity individuals.

Additional work by Huang and colleagues (Huang, 2010; Huang & Pashler, 2007; Huang, Treisman, & Pashler, 2007) also suggests that even with very simple items serial processing may be necessary. As part of their Boolean map theory (Huang & Pashler, 2007), they propose that one could encode multiple perceptual features (e.g., color, size) about an object concurrently, but could only access a single feature per perceptual dimension (e.g., only one color) at any instant.[1] This one-feature-per-dimension principle was demonstrated through multiple experiments using very simple colored shapes (Huang & Pashler, 2007, Experiment 3–4; Huang et al., 2007, Experiment 1). During each trial, participants were briefly shown two colored stimuli (targets) presented either sequentially or simultaneously, followed by a test stimulus, for which participants indicated whether it matched one of the targets. Matching performance was worse for targets presented simultaneously than sequentially, suggesting that people could not consolidate two items at the same time.

However, both West et al. (2010) and Huang and colleagues (Huang & Pashler, 2007; Huang, Triesman, & Pashler, 2007) may have underestimated the number of items that individuals are able to simultaneously consolidate. West et al. (2010) used a change detection paradigm and complex displays to assess capacity limits. Evidence suggests that change detection tasks may underestimate the capacity of visual encoding (Hollingworth, 2003; Mitroff, Simons, & Levin, 2004; Simons & Rensink, 2005). Specifically, successful change detection requires one to both encode the to-be-changed item and to compare that representation to the subsequent object at the same location. A failure in the comparison process can produce change detection errors, despite the successful encoding of the prechange item (Mitroff et al., 2004).

The experiments by Huang et al. (Huang & Pashler, 2007, Experiment 3–4; Huang et al., 2007, Experiment 1) may also underestimate capacity. Neurophysiological studies have shown that when multiple objects appear within a neuron's receptive field, they compete to drive the neuron's firing (Desimone & Duncan, 1995; Duncan, 2006; Moran & Desimone, 1985). For a given cortical visual area, competition is also stronger when stimuli are closer together (Kastner et al., 2001). Consistent with this, behavioral studies have shown that stimuli in the same hemifield compete for attentional resource, while stimuli in different hemifields consume independent attentional resources (Alvarez & Cavanagh, 2005; Reardon, Kelly, & Mathews, 2009; Sereno & Kosslyn, 1991). Specifically, the findings of Alvarez and Cavanagh (2005) and Sereno and Kosslyn (1991) of a bilateral advantage in tracking and discrimination tasks suggest that the simultaneous presentation of items in the same hemifield could reduce performance due to visual competition. Given that the colored square stimuli presented by Huang et al. (2007, Experiment 1) were quite small and close to each other, there might be strong competition between them to access attentional resources during simultaneous presentations, but not during the sequential presentations, leading to an advantage in the latter condition. Thus, neural competition, along with its dependence on stimulus configuration, might modulate the capacity limit of VSTM consolidation. In sum, although previous studies suggest that the process of consolidating information into VSTM has severe capacity limits, evidence for a strictly serial process is less clear.

In the present study, we also used the sequential/simultaneous paradigm to measure the capacity of consolidating simple stimuli into VSTM. As suggested by Scharff and colleagues (Scharff & Palmer, 2008; Scharff, Palmer, & Moore, in press), the sequential/simultaneous paradigm provides a sensitive measure of whether multiple items may be processed in parallel, or if sequential processing is required. Here we used simple colored objects presented either sequentially or simultaneously. The sequential condition served as a baseline measure of performance as only one item had to be consolidated into VSTM at a time, while during simultaneous trials multiple stimuli competed for consolidation resources. In both conditions, the same amount of information had to be retained and decided upon, while the amount of information that had to be *concurrently* consolidated differed. Equivalent performance in the simultaneous and sequential trials would be strong evidence of a parallel consolidation process that had not exhausted its capacity. By contrast, better performance in sequential trials would indicate that the consolidation process was either a serial process or a parallel process that had exceeded its capacity.

Experiment 1

Our first experiment was designed to reexamine Huang and Pashler's finding of extreme capacity limits for the consolidation of simple colored stimuli. As we outlined above, it is possible that the extreme capacity limit they reported might not reflect a general limit in the consolidation process, but may represent a special case when stimuli are so close to one another that they compete for the same neural and attentional resources.

Experiment 1a

Participants were shown two briefly displayed and masked colored squares (targets), either sequentially or simultaneously. The two targets and their masks appeared at two of the corners of an imaginary $1° \times 1°$ square that was centered on the fixation point. In half the trials, the two squares were aligned to one another vertically so that they both appeared in the same hemifield. In the other trials, the targets were aligned horizontally, so one square appeared in each hemifield. The presentation of targets was followed by a single color probe at fixation. Participants had to make a two-alternative forced choice indicating whether the probe color matched either of the targets or was a novel color. This produced a 2×2 within-subjects design with two levels of target presentation timing (sequential vs. simultaneous) and two levels of target presentation location (same hemifield vs. different hemifields). Both the amount of information that had to be retained in memory,

[1] Huang and colleagues have used the word "access" to describe "the limit on the content (or, in some sense, quantity) of visual information that is able to reach the stage of consciousness at any one moment" (Huang, 2010). In this article, we chose to use the word "consolidate" to describe the process of forming stable, reportable representations in visual short-term memory. This usage is perhaps broader than what is implied by "consolidation" in typical working memory studies. We believe these words (access vs. consolidation) mean essentially the same thing and it is somewhat difficult to dissociate perceptual versus memory-based accounts of capacity limit (for a thorough discussion of this issue, see Huang, 2010, pp. 176–177).

and the amount of information that had to be decided upon, was the same across all conditions.

If individuals are only able to consolidate one item at a time, then we should find better performance in the sequential than the simultaneous condition. Furthermore, if stimulus competition plays a role in the simultaneous condition, we would expect performance in this condition to depend on the stimulus configuration. Specifically, given that stimuli competition depends on distance and hemifield factors (Alvarez & Cavanagh, 2005; Awh & Pashler, 2000; Reardon et al., 2009; Sereno & Kosslyn, 1991), we would expect better performance when two stimuli are presented in different hemifields than when they are presented in the same hemifield in the simultaneous condition. Such a result would imply that the capacity to consolidate stimuli into VSTM is not a constant but depends on the spatial layout of the memory array.

Methods.

Participants. Eighteen participants from Michigan State University took part in the experiment. Two volunteered (including one author, I.M.), and 16 were compensated with course credit for their participation. All participants gave written informed consent. The study protocol was approved by the institutional review board at Michigan State University.

Stimuli. The stimuli, illustrated in Figure 1, were presented on a cathode ray tube (CRT) monitor with a 120-Hz refresh rate, using MGL (http://gru.brain.riken.jp/doku.php?id = mgl:overview), a set of custom OpenGL libraries for MATLAB (Mathworks, Natick, MA). Participants viewed the screen from 60 cm away, with their heads stabilized by a chin-rest.

Stimuli consisted of highly saturated red, green, yellow, or blue squares, measuring $1° \times 1°$. They were presented on a dark background and were drawn on the corners of an imaginary square, centered at fixation with 1° of eccentricity for the corners. The mask ($1° \times 1°$) consisted of an 8×8 multicolored checkerboard pattern composed of the same four colors as the stimuli; the color of each square in the mask was randomly assigned for each mask presentation. The probe square was a $1° \times 1°$ square presented at fixation. The exposure duration of all stimuli was verified with a photodiode and oscilloscope.

Procedure: Main task. Figure 1 shows a schematic of one trial in the two conditions. Sequential trials began with a 400-ms fixation, followed by the first target square (the duration was determined by a threshold procedure described below). A mask was then presented for 200 ms, which was followed by a 500-ms fixation period. The second target square was then shown for the same duration as the first, followed by a 200-ms mask, and another 500-ms fixation period before the presentation of the probe square. The probe remained on screen for 2000ms, or until a response was made, whichever occurred earlier. Participants responded by pressing keys on a standard computer keyboard. They pressed the “z” key if the probe did not match either of the two targets, or “1” key (on the numeric keypad) if it did. Feedback was provided, with a high tone indicating a correct response, and a low tone indicating an incorrect response. A 600ms period preceded the start of the next trial. Simultaneous trials were similar, with the exception that two colored squares were shown simultaneously, followed by two masks in the same locations as the target squares (Figure 1a). Participants were instructed to respond as accurately as possible, without the need to respond fast.

Participants completed 5 blocks of 48 trials of the sequential condition, and 5 blocks of 48 trials of the simultaneous condition. The order of the last 8 blocks was counterbalanced in an ABBABAAB sequence. The two stimulus locations were randomly determined on each trial, except they could not be on a diagonal configuration (i.e., upper left/lower right, or upper right/lower left). This produced a total of four possible location pairs: two vertical configurations (same hemifield) and two horizontal

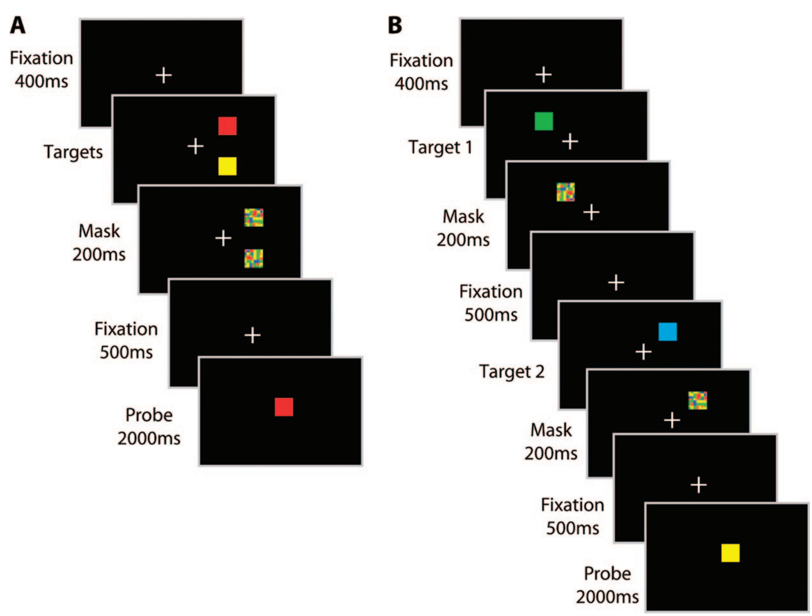

Figure 1. Schematic of stimuli and trial sequence. (A) An example trial in the simultaneous condition. Depicted here is a probe match trial. (B) An example trial in the sequential condition. Depicted here is a probe mismatch trial.

configurations (different hemifields). The first two blocks were used as practice to familiarize participants with the task, and were excluded from analysis (including them did not change the results).

Procedure: Thresholding exposure duration. We manipulated the stimulus exposure duration to calibrate the task difficulty. Prior to the main task, each participant ran two blocks (48 trials each) of the sequential condition, and two blocks (48 trials each) of the simultaneous condition. A method of constant stimuli was used with eight durations: 8 ms, 16 ms, 33 ms, 66 ms, 133 ms, 266 ms, 533 ms, and 800 ms. In these blocks, we only presented stimuli in the two vertical configurations, to obtain a sufficient amount of data to estimate threshold. Data from the sequential condition were used to estimate threshold; we included the simultaneous condition to equate participants' familiarity with both conditions. Proportion correct was calculated for each exposure duration and fitted with a Weibull function using psignifit (Wichmann & Hill, 2001). Stimulus duration that yielded an overall accuracy of ~80% correct was used for the main task.

Results and Discussion. The average presentation duration across participants was 60ms (range = 8–158 ms). For all experiments we analyzed the matching performance using both signal detection d' and accuracy (hits—false alarms). These two methods produced the same pattern of results, and here we report the accuracy results. A 2 × 2 repeated-measures analysis of variance (ANOVA) with two levels of presentation condition (simultaneous/sequential) and two levels of spatial arrangement (single hemifield/both hemifields) revealed no main effect of presentation condition, $F(1, 17) = .02$, no main effect of spatial arrangement, $F(1, 17) = .14$, and no presentation type by spatial arrangement interaction, $F(1, 17) = .18$. Our failure to find a sequential advantage suggests that people were able to consolidate the stimuli in the simultaneous condition in an unlimited parallel manner, and is inconsistent with claims that people can only consolidate one stimulus at a time (Huang et al., 2007; West et al., 2010). In particular, accuracy for horizontal and vertical stimulus configurations in the simultaneous condition was not significantly different (horizontal: $M = 0.44$, $SE = 0.06$; vertical: $M = 0.42$, $SE = 0.06$); nor were they different in the sequential condition (horizontal: $M = .43$, $SE = 0.06$; vertical: $M = .43$, $SE = 0.07$), suggesting that stimuli were not competing for representations when presented in the same visual hemifield.

We did not find support for our original intuition that stimulus competition contributed to performance on simultaneous trials. This might be due to the fact that our stimuli were rather simple compared to previous work showing a bilateral advantage (Awh & Pashler, 2000; Reardon et al., 2009; Sereno & Kosslyn, 1991), because competition presumably increases with the number and complexity of stimuli. Furthermore, bilateral advantage is generally smaller in working memory tasks than in tasks involving active spatial selection such as tracking (Delvenne, 2005; Umemoto, Drew, Ester, & Awh, 2010). Finally, our results were consistent with results from a change detection task that showed a lack of bilateral advantage in working memory for colors (Delvenne, 2005).

A bigger surprise, however, was that we also did not observe any difference between sequential and simultaneous conditions. This latter result contradicted findings from previous studies (Huang, 2010; Huang & Pashler, 2007; Huang et al., 2007). In our second experiment we sought to replicate Huang et al.'s (2007) finding of superior accuracy with sequential than simultaneous presentations, and to examine whether we could attribute our dissimilar results to differences in methodology.

Experiment 1b

In contrast to experiments reported by Huang and colleagues, we did not observe any difference between sequential and simultaneous presentation in Experiment 1a. Although the experiment was largely similar to Huang and colleagues' experiments, there were a few methodological differences. A potentially important difference is that there were multiple contingencies built in Huang et al.'s experiments.[2] First, the four possible colors were assigned to two groups: red/green and blue/yellow, with one color from each group randomly selected as targets on each trial. Thus, the possible target pairings were red-blue, red-yellow, green-blue, green-yellow, with no possibility for red-green or blue-yellow. This contrasts with our Experiment 1a, in which all six pairings were possible. Second, there were spatial contingencies in Huang et al.'s experiment such that each group (red/green vs. blue/yellow) consistently appeared along a particular diagonal, for example, for a given participant the stimulus that appeared on the 45° diagonal was always either red or green and the stimulus along the 135° diagonal was always blue or yellow. Finally, there was a temporal contingency in the sequential condition such that two groups (red/green vs. blue/yellow) appeared in specific order, for example, the red or green stimulus always appeared first followed by the blue or yellow stimulus. None of these contingencies existed in our Experiment 1a; both the location and order (in the sequential condition) of stimuli were randomized.

The existence of these contingencies may favor the sequential condition more than the simultaneous condition. For example, in the sequential condition, participants could anticipate either red or green in the first interval on the 45° diagonal, and anticipate blue or yellow in the second interval on the 135° diagonal. Thus during each temporal interval of the sequential condition, participants needed to monitor only two possible locations for the appearance of one color (e.g., red). If that color did not appear, the participant would know that the other color in the group must have appeared (e.g., green). In the simultaneous condition, the lack of temporal contingency would make such anticipation more difficult; at a minimum, participants would need to monitor four possible locations for the appearance of at least two colors (one color from each group).

Another methodological difference between Huang et al.'s experiment and Experiment 1a concerns the number of masks. In the experiments by Huang et al., simultaneous targets were followed by masks in all four locations, whereas each sequential target was followed by two masks, one in the stimulus location and one in that diagonal's empty location (four masks total in both conditions). By contrast, in our Experiment 1a, we masked only the locations occupied by the stimuli. Thus, we presented two masks in the simultaneous condition and one mask in each interval in the

[2] These contingencies were not reported in the original article (Huang et al., 2007), but were present in the computer program supplied by Dr. Liqiang Huang, who also acknowledged their existence (Liqiang Huang, personal communication, June 2010).

sequential condition (two masks total). Given this difference between masking procedures, it is possible that our results differed because we used fewer masks. Alternatively, it is possible that Huang et al.'s presentation of two masks on the diagonal in the sequential condition may have further accentuated the contingent structure of the experiment, thereby encouraging participants to pay attention to one diagonal in the first interval and the other diagonal in the second interval.

In sum, we reasoned that the contingencies and the mask configuration used by Huang et al., might have bolstered performance in the sequential condition relative to the simultaneous condition. That is, the difference they observed between the two conditions might have been caused by artificially high performance in the sequential condition due to participants' effective use of contingencies, rather than a decrement in performance in the simultaneous condition. Experiment 1b was designed to test this hypothesis by comparing performance in blocks of trials which replicated these contingencies to blocks in which the contingencies were removed.

Method.

Participants. Thirteen participants from Michigan State University took part in this experiment. Three participants volunteered, and 10 were compensated with course credit for their participation.

Stimuli. Stimuli were identical to those of experiment 1a.

Procedure. We first determined the exposure duration using the same procedure as in Experiment 1a. We then ran eight experimental blocks of simultaneous and sequential conditions. In half of the blocks (two simultaneous and two sequential), we repeated the random design as in Experiment 1a, such that there was no contingency in target color, location or temporal interval. We did, however, add two masks in the simultaneous condition and one mask (in the diagonal location) in each interval in the sequential condition. This increased the total number of masks per condition to four (same as Huang et al., 2007) and allowed a more direct comparison to the contingent conditions (see below).

In the other half of the blocks (two simultaneous and two sequential), we introduced the contingencies that appeared in Huang et al.'s experiment. In these blocks we fixed the location of the red/green and blue/yellow stimuli to a particular diagonal. This meant that in any trial the 135° diagonal could be either red or green, and the 45° diagonal could be either blue or yellow (this spatial contingency was flipped for half the participants). The presentation sequence was also fixed during sequential trials, such that one of the two possible colors of the 135° diagonal would always appear first, followed by one of the two possible colors of the 45° diagonal (this temporal contingency was also flipped for half the subjects). In addition, on simultaneous trials, all four locations were masked; on sequential trials, two masks followed each target stimulus, one over the stimulus location, and another in the diagonal location. This masking procedure matches that of Huang et al.'s experiment. For a particular subject, the spatial and temporal contingency was fixed, and participants were informed of such contingencies at the beginning of each block. There were four "superblocks", each containing one sequential and one simultaneous block with the same contingency level (either random or contingent). The order of simultaneous and sequential blocks within a superblock was randomized and the order of contingent and random superblocks was counterbalanced. Each block consisted of 48 trials.

Results and Discussion

Results confirmed our prediction (Figure 2). In the random blocks (no contingency), there was no difference between the two conditions, $t(12) = 0.44$, planned comparison, replicating results from Experiment 1a. In the contingent blocks, sequential presentation produced significantly higher performance than simultaneous presentation, $t(12) = 4.20$, $p < .05$, planned comparison. This pattern was further verified by a significant interaction effect, $F(1, 12) = 7.83$, $p < .05$ in a two (sequential vs. simultaneous) by two (contingent vs. random) repeated-measures ANOVA. Post hoc paired *t* tests revealed this interaction was driven by higher accuracy in the sequential contingent condition relative to all other conditions, all $t(12) > 2.6$, $p < .05$, which did not differ from each other (all $p > .15$).

In this experiment we were able to replicate the findings of Huang el al. (2007) of a sequential over simultaneous advantage with very simple stimuli. However, this was obtained only after we introduced contingencies in the design. We suggest that this was due to participants' ability to anticipate which of two possible diagonal locations would appear first, and which of two colors could appear in each diagonal in the sequential-contingent condition. Supporting this interpretation, our post hoc comparisons showed that the difference in sequential-contingent and simultaneous-contingent conditions was due to a better performance in the former, rather than an impaired performance in the latter, when compared to the random conditions (Figure 2). When the task parameters were fully randomized, the sequential and simultaneous conditions are indistinguishable; this replicates Experiment 1a, even though we doubled the number of masks. Thus, these results suggest that number of masks is not a critical factor in obtaining our results and individuals are able to consolidate two stimuli in parallel.

Experiment 2

In Experiment 1, we found no difference between sequential and simultaneous presentations when there were two targets, suggest-

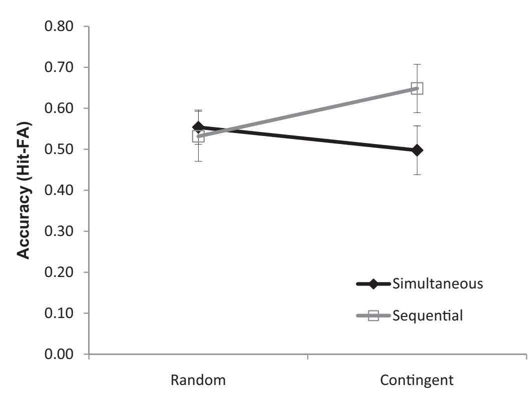

Figure 2. Accuracy (hits-false alarms) in Experiment 1b as a function of presentation condition (simultaneous or sequential) and block type (random or with contingencies). Error bars show the *SEM*.

ing the visual system can consolidate two simple items in parallel without capacity limits. In Experiment 2, we examined the upper limits of this process. To do so we increased the number of targets to three and four. In addition, given that Experiment 1b confirmed that prior knowledge of likely locations positively impacts performance, we increased the positional uncertainty of the stimuli to examine whether parallel consolidation was possible even when one had little a priori knowledge about the location of potential targets. To accomplish this second goal, stimuli were presented in random locations along the circumference of an imaginary circle, instead of at the corners of a fixed square as in Experiment 1. This made it difficult to predict the location of the stimuli during each trial. If all items could be consolidated concurrently, then performance should be similar in the two conditions; however, if only a subset of the items are able to be consolidated concurrently, then sequential performance should surpass that of simultaneous once the capacity limit has been reached.

Experiment 2a

Method.

Participants. Thirteen observers from the Michigan State University took part in this experiment. Two volunteered (including one author, I.M.), and 11 were compensated with course credit for their participation. Two additional observers were tested on the threshold procedure but were not further tested on the main task because of a very high threshold (>400 ms).

Stimuli. Stimuli were circles with a 1° diameter, presented around the circumference of an imaginary circle of 8° diameter centered at fixation. Stimuli could be presented anywhere around the circumference, with the constraint that they would not be presented within 45° angular distance of each other. In addition to red, green, blue, yellow, four more colors (cyan, magenta, orange and purple) were also added as possible stimulus colors. Masks were similar to Experiment 1a, except they were shown in 1° circular apertures.

Task and design. There were two within-subject factors in the experiment: presentation condition (sequential and simultaneous) and set size (2, 3, 4). Simultaneous trials were identical to those of Experiment 1a, except that either 2, 3, or 4 targets were presented and followed by the same number of masks. Sequential trials were also identical to those in Experiment 1a for set size of 2. For set size of 3 and 4, an additional stimulus frame and 700 ms ISI (200 ms mask + 500 ms fixation) was added for each increase in set size.

Procedure. For each subject, we first determined the threshold exposure duration using a similar procedure as in Experiment 1a. Only set size two was used during the thresholding procedure; the duration was again determined using the sequential condition. Participants then ran 10 blocks of 48 trials. Presentation condition (sequential or simultaneous) alternated between blocks and set size was randomly interleaved across trials within each block. Block order was counterbalanced across participants. The first two blocks were considered as practice, and were excluded from analysis.

Results and Discussion. The average presentation duration was 86ms (range: 42-135ms). Mean accuracy (hits-false alarms) is shown in Figure 3a. A two (presentation condition) by three (set-size) repeated-measures ANOVA revealed a main effect of set size, $F(2, 24) = 24.45$, $p < .001$, and set size by presentation condition interaction, $F(2, 24) = 4.80$, $p < .05$, with no main effect of presentation condition, $F(1, 12) = 0.33$. The interaction resulted because performance dropped more precipitously for the simultaneous than the sequential condition. This pattern is consistent with the expectation that consolidation capacity would be exceeded in the simultaneous condition with large set sizes. To determine the minimum set size at which consolidation capacity was exceeded, we compared performance between the simultaneous and sequential trials at each set size, running three paired *t* tests as planned comparisons. These analyses revealed no significant differences at any of the set sizes, with only a trend toward better performance in the sequential condition for set size four, $t(12) = 1.76$, $p = .1$. Thus, the data from this experiment are somewhat inconclusive. The significant interaction in the ANOVA suggests that consolidation capacity was exceeded at larger set sizes in the simultaneous condition, but the paired comparisons for even the largest set size failed to find impaired performance for the simultaneous condition, as one would expect if consolidation capacity had been exceeded.

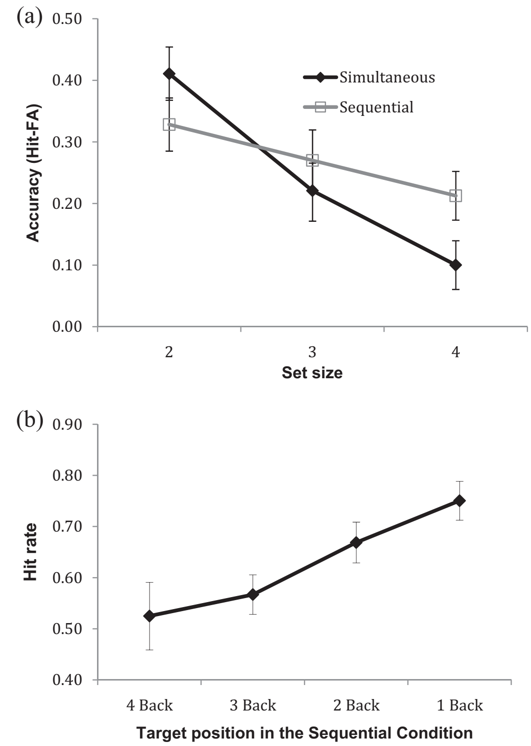

Figure 3. (a) Accuracy (hit-false alarm) as a function of set size for Experiment 2a. (b) Hit rates for the sequential condition as a function of the serial position of the stimulus that was eventually probed. Note this is the average hit rate across set sizes, aligned to the memory probe. Error bars show the *SEM*.

Our failure to find significant differences in the paired comparisons may have been due to too little statistical power, but a second possibility is that differences in the retention interval across conditions may have also contributed to this failure. In this experiment, simultaneous targets were probed after a relatively short retention interval that was identical to the interval for the last item presented in the sequential condition. That is, most sequential targets (except the last one) were probed after longer intervals. Thus, on average, sequential targets had to be retained longer than simultaneous targets and this disparity was especially severe at large set sizes. If representations decay, become noisier, or abruptly terminate over these short retention intervals (Lee & Harris, 1996; Zhang & Luck, 2009), a longer average retention interval for the sequential condition may have artificially lowered performance for the sequential relative to the simultaneous condition. Furthermore, it is possible that greater numbers of stimuli lead to greater interference between the initial and succeeding items (Magnussen & Greenlee, 1997, 1999). Consistent with these concerns, we found evidence for a recency effect in the sequential condition: targets presented later in the sequence were recognized better. A one-way repeated-measures ANOVA on the hit rate with serial position as factor in the sequential condition (Figure 3b) found a significant main effect, $F(3, 36) = 11.03$, $p < .001$.

Given this retention interval effect, the most conservative approach to evaluating the consolidation limit would be to set the retention interval in the simultaneous condition equal to the longest possible retention interval in the comparable sequential condition. This approach would make the average retention interval longer for the simultaneous than sequential condition thereby biasing against the simultaneous condition. If this still produced equivalent performance for simultaneous and sequential presentations one would be fairly confident that the stimuli in the simultaneous condition were being processed in an unlimited capacity, parallel fashion.

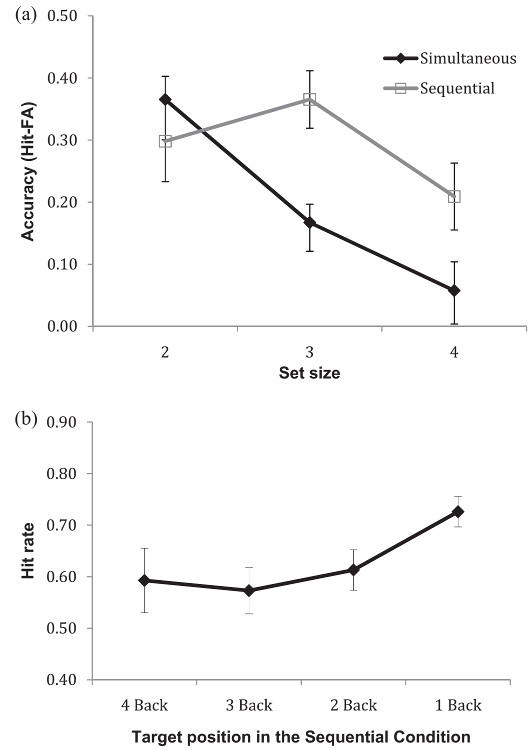

Figure 4. (a) Accuracy (hit-false alarm) as a function of set size for Experiment 2b. (b) Hit rates for the sequential condition as a function of the serial position of the stimulus that was eventually probed. Note this is the average hit rate across set sizes, aligned to the memory probe. Error bars show the *SEM*.

Experiment 2b

In Experiment 2b we took a conservative approach and matched the retention interval of the simultaneous targets with that of the first item in the sequential targets of the same set size. Doing so meant that the simultaneous targets had a longer retention interval on average than sequential targets, thus providing a conservative estimate of the capacity to concurrently consolidate stimuli into VSTM. If we continued to find equal performance in the two conditions, this would provide strong support that individuals can consolidate multiple items simultaneously.

Method. All methods were the same as in Experiment 2a, with the following modifications.

Participants. Thirteen participants from Michigan State University took part in this experiment; four participants volunteered, nine were compensated with course credit, and two had participated in Experiment 2a (including one author, I.M.).

Task. The sequential trials were identical to those in Experiment 2a. During simultaneous trials, a delay period followed the offset of the stimulus and mask. This period varied by set size to equate the retention interval (time from target offset to probe onset) for the simultaneous trials to the retention interval of the first target in the sequential trial, at the same set size. For example, for a set size of two, this would mean a blank delay equal to 1,200 ms + target exposure duration (determined for individual participants).

Results and Discussion. The average presentation duration for the stimulus was 117ms (range = 42–250 ms). Mean accuracy is shown in Figure 4a. There was a general decline in performance because set size increased with a more pronounced decline in the simultaneous condition. This was confirmed by a two (presentation condition) by three (set size) repeated-measures ANOVA which yielded a main effect of presentation condition, $F(1, 12) = 8.39$, $p < .05$, set size, $F(2, 24) = 21.04$, $p < .0001$, and a significant interaction, $F(2, 24) = 13.51$, $p < .001$. Planned comparisons between the simultaneous and sequential performance at each set size showed no significant difference at set size two, $t(12) = 1.46$, $p > .15$, but significant differences at set size three, $t(12) = 5.18$, $p < .001$, and four, $t(12) = 3.00$, $p < .05$. As is shown in Figure 4b, we again found a recency effect in the sequential condition, one-way ANOVA, $F(3, 36) = 4.94$, $p < .01$.

The finding of equivalent performance between the simultaneous and sequential presentation at set size two provides the strongest evidence yet that at least two items can be consolidated concurrently. At this set size, accuracy was the same in both conditions, despite the fact that on average, simultaneously presented targets had to be retained for a longer period than sequentially presented targets. In addition, the fact that we replicated the same set size two effect using this conservative method suggests that the similar results in our previous experiments were not due to a shorter retention period during simultaneous trials.

Using this conservative method, simultaneous performance was significantly worse than sequential for set sizes three and four. This latter result is important as it shows that our paradigm is sensitive enough to detect deterioration in performance at larger set sizes. Though our findings strongly suggest that processing is subject to capacity limits at larger set sizes, the finding of a recency effect in the sequential condition suggests that some of the decrement in performance might also be attributable to the decay, termination, or interference between representations due to longer retention intervals in the simultaneous condition.

General Discussion

The goal of the current experiments was to examine the amount of information that may be concurrently consolidated in working memory at any one moment. The few previous experiments that investigated this issue (Huang & Pashler, 2007; West et al., 2010) concluded that consolidation was a strictly serial process that could only act upon one item at a time. By contrast, in four experiments with a simultaneous/sequential presentation manipulation, we found that participants can process at least two simple items in an unlimited parallel fashion. In addition, we were able to demonstrate that previous results used to argue for a strictly serial process (Huang & Pashler, 2007) may have been confounded by contingencies that were present in the method. When we included the same contingencies (Experiment 1b) we were able to replicate Huang and Pashler's (2007) results, but once those contingencies were removed, participants were capable of consolidating two simultaneously presented colors as effectively as two sequentially presented colors. This pattern of results was also observed when the location of the stimuli was made unpredictable (Experiment 2a) and the retention interval in the simultaneous condition was on average longer than the retention interval in the sequential condition (Experiment 2b). In short, it is a very robust finding and we have confidence that people can consolidate two simple stimuli in parallel.

The conclusion that two items can be consolidated in parallel seems at odds with Vogel et al.'s (2006) finding of better change detection performance with brief stimulus-to-mask durations for one colored square than two colored squares. However, in their experiment, as set-size increased, subjects also needed to make more comparisons between the memory array and the test array, thus response uncertainty and decision noise also increased. In other words, better performance with set size one could have been due to less decision noise and response uncertainty, instead of superior consolidation into VSTM.[3] By contrast, we always compared between sequential and simultaneous presentation of equal set sizes. This allowed us to hold decision noise constant across our comparisons. As such, we believe that our method more accurately reflects the number of items that can be consolidated simultaneously.

Our main finding of equivalent performance in sequential and simultaneous presentation conditions contradicts the recently developed Boolean map theory of visual attention (Huang & Pashler, 2007), which predicts that one can only access one feature at a time (e.g., red) in a particular dimension (e.g., color). Thus, the Boolean map theory might need some modification. It seems the theory is flexible enough to accommodate our results, for example, by supposing that different color values (at least when they are highly distinctive) form different perceptual dimensions. Furthermore, it is worth noting that many of the phenomena that the Boolean map theory is argued to account for, are phenomena like texture segregation, figure ground segregation, and the detection of symmetry (Morales & Pashler, 1999). The types of displays used in these tasks typically consist of many elements and are fairly complex. It is possible that the theory is accurate when the visual scene is complex enough to require selective attention that selects some information for additional processing while filtering out other information. Once such a selective attentional mechanism is required it is possible that selection must occur based on a single feature (e.g., select only red objects) at a time. By contrast, when visual displays are so simple that a selection mechanism is not engaged, our results suggest that one is capable of accessing at least two features simultaneously.

In Experiments 2a and 2b, we added set sizes with three or four stimuli to determine the maximum number of items that could be simultaneously consolidated. Although these experiments replicated the finding that at least two items could be processed concurrently and suggest that four items exceed the capacity of concurrent processing, it was less clear whether people could simultaneously consolidate three items or not. In Experiment 2a, set-size three performance was equivalent in the simultaneous and sequential conditions, but in Experiment 2b there was a clear advantage for sequential over simultaneous presentation. The only difference between these experiments was the length of the retention interval in the simultaneous condition. In Experiment 2a it was relatively short and set to the retention interval for the last item in the sequential presentation. In Experiment 2b it was relatively long and set to the retention interval for the first item in the sequential presentation. The difference in the results between these two experiments could be caused by decay (Lee & Harris, 1996), but could also be because of other factors, such as interference (Magnussen & Greenlee, 1997, 1999) or the sudden extinction (Zhang & Luck, 2009) of memory representations. The role of any of these possible factors on participants' performance is consistent with the finding of a recency effect in the sequential conditions of Experiments 2a and 2b.

The finding of these recency and retention interval effects in the sequential condition is concerning for a number of reasons. It makes direct comparisons between the simultaneous and sequential condition difficult, since one cannot perfectly equate retention intervals across these two presentation modes. Given this inherent inequality in retention intervals, it also demonstrates a potential limitation in the simultaneous/sequential method. Finally, the finding highlights the fact that other more complex memory effects can interfere with one's attempts to determine basic cognitive processes like consolidation capacity. While these limitations preclude us from being able to definitely determine the number of items that can concurrently be processed, they do not undermine our conclusion that people can simultaneously process at least two simple items. Experiment 2b supports this conclusion even when all of these extraneous memory factors should have biased performance against the simultaneous condition.

In general, our finding of reduced simultaneous performance with set size four is consistent with claims that consolidation into

[3] Vogel et al. acknowledged that decision noise necessarily increased as they increased set size. Indeed, they provided evidence (Experiment 5) that decision noise alone could not account for the dramatic drop in performance they found as set sizes increased from one to four items. However, their experiments did not directly address whether the modest decrease in observed performance from set size 1 to 2 could have been explained by decision noise.

working memory is a limited capacity process (Jolicoeur & Dell'Acqua, 1998; Vogel et al., 2006). However, our finding of equivalent performance with set size two is inconsistent with claims that this limitation is so severe that it requires people to consolidate a single object at a time (Huang & Pashler, 2007; West et al., 2010). Instead, we find that people can consolidate at least two items in parallel.

Our finding that consolidation can proceed in parallel for a limited number of items may be related to how information held in VSTM is utilized during cognitive operations. Recent work on the focus of attention in working memory suggests that attention needs to be able to simultaneously select and access multiple stored representations for many higher-level cognitive processes (Oberauer & Bialkova, 2009). For example, this joint access is required when individuals must perform an arithmetic task such as addition or subtraction, during which multiple items (numbers) must be accessed simultaneously in order to perform the mathematical operation; similar processing has been implicated for many other cognitive task such as language comprehension, solving spatial relations among objects, inductive reasoning, and so forth which rely heavily on working memory (Oberauer & Bialkova, 2009). Though many of these operations are not necessarily performed on visual representations, given the role of working memory as the interface for complex thought (Baddeley, 2003), it may very well be that visual information also needs to be accessed in a similar fashion. That is, if visual working memory is used as a way to maintain and compare objects in our environment (Hyun, Woodman, Vogel, Hollingworth, & Luck, 2009), simultaneous consolidation and access to multiple items may make the comparison process much more efficient.

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.

Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science, 16,* 637–643.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18,* 622–628.

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 834–846.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews, 4,* 829–839.

Delvenne, J. F. (2005). The capacity of visual short-term memory within and between hemifields. *Cognition, 96,* B79–B83.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

Duncan, J. (2006). Brain mechanisms of attention. *The Quarterly Journal of Experimental Psychology, 59*(1), 2–27.

Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 388–403.

Hollingworth, A., & Luck, S. J. (2008). Visual memory systems. S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 3–8). New York, NY: Oxford University Press.

Huang, L. (2010). What is the unit of visual attention? Object for selection, but Boolean map for access. *Journal of Experimental Psychology: General, 139,* 162–179.

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychology Review, 114,* 599–631.

Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human visual awareness. *Science, 317,* 823–825.

Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1140–1160.

Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36,* 138–202.

Kastner, S., DeWeerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology, 86,* 1398–1411.

Lee, B., & Harris, J. (1996). Contrast transfer characteristics of visual short-term memory. *Vision Research, 36,* 2159–2166.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281.

Magnussen, S., & Greenlee, M. W. (1997). Competition and sharing of processing resources in visual discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1603–1616.

Magnussen, S., & Greenlee, M. W. (1999). The psychophysics of perceptual memory. *Psychological Research, 62,* 81–92.

Mitroff, S. R., Simons, D. J., & Levin, D. T. (2004). Nothing compares 2 views: Change blindness can occur despite preserved access to the changed information. *Perception and Psychophysics, 66,* 1268–1281.

Morales, D., & Pashler, H. (1999). No role for colour in symmetry perception. *Nature, 399,* 115–116.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.

Oberauer, K., & Bialkova, S. (2009). Accessing Information in Working Memory: Can the focus of attention grasp two elements at the same time? *Journal of Experimental Psychology: General, 138,* 67–87.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics, 16,* 283–290.

Reardon, K. M., Kelly, J. G., & Mathews, N. (2009). Bilateral attention advantage on elementary visual tasks. *Vision Research, 49,* 691–701.

Scharff, A., & Palmer, J. (2008). Distinguishing serial and parallel models using variations of the simultaneous-sequential paradigm. *Journal of Vision, 8,* 981.

Scharff, A., Palmer, J., & Moore, C. M. (in press). Measuring perceptual capacity for features and words using an extended simultaneous-sequential paradigm. *Journal of Experimental Psychology: Human Perception and Performance.*

Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception and Psychophysics, 64,* 754–763.

Sereno, A. B., & Kosslyn, S. M. (1991). Discrimination within and between hemifields: A new constraint on theories of attention. *Neuropsychologia, 29,* 659–675.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Science, 9,* 16–20.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74*(11 Whole No. 498), 1–29.

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science, 1,* 46–54.

Umemoto, A., Drew, T., Ester, E. F., & Awh, E. (2010). A bilateral advantage for storage in visual working memory. *Cognition, 117,* 69–79.

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.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1436–1451.

West, G. L., Pun, C., Pratt, J., & Ferber, S. (2010). Capacity limits during perceptual encoding. *Journal of Vision, 10,* 1–14.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics, 63,* 1293–1313.

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin & Review, 15,* 223–229.

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20,* 423–428.

Received February 4, 2011
Revision received April 4, 2011
Accepted April 11, 2011 ■