A severe capacity limit in the consolidation of orientation information into visual short-term memory

Mark W. Becker · James R. Miller · Taosheng Liu

Abstract Previous research has suggested that two color patches can be consolidated into visual short-term memory (VSTM) via an unlimited parallel process. Here we examined whether the same unlimited-capacity parallel process occurs for two oriented grating patches. Participants viewed two gratings that were presented briefly and masked. In blocks of trials, the gratings were presented either simultaneously or sequentially. In Experiments 1 and 2, the presentation of the stimuli was followed by a location cue that indicated the grating on which to base one's response. In Experiment 1, participants responded whether the target grating was oriented clockwise or counterclockwise with respect to vertical. In Experiment 2, participants indicated whether the target grating was oriented along one of the cardinal directions (vertical or horizontal) or was obliquely oriented. Finally, in Experiment 3, the location cue was replaced with a third grating that appeared at fixation, and participants indicated whether either of the two test gratings matched this probe. Despite the fact that these responses required fairly coarse coding of the orientation information, across all methods of responding we found superior performance for sequential over simultaneous presentations. These findings suggest that the consolidation of oriented gratings into VSTM is severely limited in capacity and differs from the consolidation of color information.

M. W. Becker (✉) · J. R. Miller · T. Liu
Department of Psychology, Michigan State University,
316 Physics Road,
East Lansing, MI 48824, USA
e-mail: becker54@msu.edu

T. Liu
e-mail: tsliu@msu.edu

T. Liu
Neuroscience Program, Michigan State University, East Lansing,
MI 48824, USA

When viewing complex scenes, a great deal of visual information impinges upon the retina and is represented in early visual cortex. However, these early sensory representations are fleeting and relatively unprocessed (Sperling, 1960). In order to form visual representations that are somewhat durable and not subject to masking by subsequent visual stimulation, one must consolidate the items into a more stable visual short-term memory (VSTM) representation (Becker, Pashler, & Anstis, 2000; Jolicœur & Dell'Acqua, 1998). The contents of VSTM form the basis for our stable and consciously available visual representations (Chun & Potter, 1995; Phillips, 1974; Rensink, O'Regan, & Clark, 1997). As such, a great deal of research has investigated the storage capacity of this system and how the capacity and resolution of items in VSTM interact (Alvarez & Cavanagh, 2004; Awh, Barton, & Vogel, 2007; Wilken & Ma, 2004; Zhang & Luck, 2008). This research suggests that the VSTM system has a relatively fixed capacity of about three to four items, thereby establishing a fairly fundamental limit on the numbers of items that one can represent at any instant.

Despite this limit, we are able to function well in complex environments, presumably because we can rapidly select new, behaviorally relevant information from the environment and can rapidly consolidate new items into the VSTM buffer (Ballard, Hayhoe, & Pelz, 1995; O'Regan, 1992), albeit at the cost of losing some previous items within that buffer (Becker & Pashler, 2002; Wolfe, Klempen, & Dahlen, 2000). Consequently, the potential impact of the capacity limit can be minimized by efficiently altering the contents of VSTM so that the most relevant objects in the environment for one's current task are available when needed (Ballard et

al., 1995; Droll, Hayhoe, Triesch, & Sullivan, 2005; Triesch, Ballard, Hayhoe, & Sullivan, 2003). This view suggests that two processes are required for the system to function efficiently, despite VSTM capacity limits: One must be able to select relevant information from the scene for consolidation, and one must be able to rapidly consolidate that information into the VSTM system. While a great deal of research has investigated the selection of items from the environment (Desimone & Duncan, 1995; Pratt & Hommel, 2003; Soto, Hodsoll, Rotshtein, & Humphreys, 2008), little is known about the consolidation process itself.

For instance, the extent to which this consolidation process is itself limited is unclear. To investigate this issue, a number of researchers (Jolicœur & Dell'Acqua, 1998; Stevanovski & Jolicœur, 2007, 2011; Vogel, Woodman, & Luck, 2006; West, Pun, Pratt, & Ferber, 2010) have varied the set size of a briefly presented array and investigated how memory for the array is affected by the set size manipulation. In general, these studies have shown that memory decreases as set size increases. The authors have interpreted this as demonstrating that the consolidation process has a limited capacity. While reduced memory performance with larger set sizes is consistent with the interpretation that VSTM consolidation is capacity limited, the fact that the result relies on comparisons between arrays of different set sizes raises the possibility of other interpretations. If, for instance, the decision noise or interference between items increased with larger set sizes, one might find reduced memory performance even if the capacity of consolidation was not exceeded by the displays (Eckstein, Thomas, Palmer, & Shimozaki, 2000; Palmer, Verghese, & Pavel, 2000).

We recently avoided the potential confounds associated with changing memory load by briefly presenting and masking two colored stimuli that were presented either simultaneously or sequentially. Using this procedure, we found equivalent performance regardless of the method of presentation, suggesting that consolidation of two color stimuli could be achieved in parallel (Mance, Becker, & Liu, 2012).

Here we extend this work to determine whether this parallel consolidation process generalizes to another early visual feature, namely orientation. There are a number of reasons for determining whether the parallel consolidation effect generalizes to orientation. First, some of the prior work in which set size manipulations were used has found more severe capacity limits for the consolidation of orientation than of color (Stevanovski & Jolicœur, 2011; Woodman & Vogel, 2008). These prior findings might suggest that it would not be possible to consolidate two orientations in parallel. Alternatively, given the caveat about set size manipulations, the findings of worse memory for orientations as set size increased could result from orientation stimuli producing more decisional noise or interference than

do color stimuli when multiple items need to be represented. A method such as simultaneous versus sequential presentation with the sizes of the memory sets equated should be able to differentiate between these two explanations.

More generally, demonstrating that the finding generalizes to other basic visual features will be important for making general claims about the operation of the consolidation process. In addition, there are reasons to believe that color may be a unique visual feature: It is coded by wavelength-selective photoreceptors, making it among the earliest visual features represented, but colors are also associated with very robust conceptual or categorical descriptions (Franklin, Pilling, & Davies, 2005). As such, the representation of color is fairly robust throughout the entire visual processing stream. In addition, behavioral results have suggested that color is a more effective feature than other basic visual features in a number of tasks. For instance, the guidance of attention by a known feature is particularly effective for color relative to orientation (Motter & Belky, 1998; Williams, 1966). Furthermore, a group of patients with hemineglect showed no deficit in finding color-defined targets that appeared in the lesioned hemifield, but they did show significant deficits in finding an orientation-defined target in that field, suggesting that color and orientation are processed differently (Wilkinson, Ko, Milberg, & McGlinchey, 2008). In short, there are a number of reasons to think that color may be a unique visual feature; thus, the finding that one can consolidate two colors into VSTM simultaneously may represent the exception rather than the rule.

In order to investigate whether the simultaneous consolidation of visual features extended to other basic features, we used a simultaneous–sequential procedure with oriented grating stimuli, using methods similar to those from our previous color study (Mance et al., 2012). We chose orientation because it is a basic visual feature that is coded as early as V1 in the visual stream (Hubel & Wiesel, 1968, 1974). Oriented gratings were briefly presented and then masked. In some trials, the two gratings were presented sequentially, and in others they were presented simultaneously (Scharff, Palmer, & Moore, 2011; Shiffrin & Gardner, 1972). People were then probed to report the orientation of one of the gratings. The comparison of performance between simultaneous and sequential presentation allowed us to assess the consolidation of features into VSTM while holding other task demands constant (e.g., memory load and decisional load). If we were to observe results similar to those of our color experiments, this would provide additional information that parallel consolidation of two features into VSTM is a general property. If, however, color and orientation behaved differently, this would suggest that the consolidation process has different capacity limits for different features.

Experiment 1a

Method

Participants Ten observers participated in this experiment. All were graduate or undergraduate students at Michigan State University and were naïve as to the purpose of the experiment. The participants gave informed consent and were compensated at the rate of $10/h. All of the experimental protocols were approved by the Institutional Review Board at Michigan State University.

Visual stimuli Visual stimuli were generated using MGL (http://gru.brain.riken.jp/doku.php?id=mgl:overview), a set of custom OpenGL libraries running in MATLAB (The MathWorks, Natick, MA). The orientation stimuli were sinusoidal gratings (contrast, 0.7; spatial frequency, 2.9 cycles/deg) in a circular aperture (size, 1.1º) presented on a gray background (24.2 cd/m^2). The edge of the aperture was smoothed such that no sharp change in luminance was present between the grating and the background. The grating could be presented in one of four possible locations, located at the corners of an imaginary square (eccentricity, 2.1º). The mask stimulus was a circular aperture (size, 1.3º) containing pixel noise, with random luminance levels in a uniform distribution (0–48.4 cd/m^2). A fixation dot (0.2º) was presented in the center of the screen throughout the experiment. The stimuli were presented on a CRT monitor (1,024 × 768 pixel, 96-Hz refresh rate, with linearized

luminance levels), and observers viewed the display at a distance of 57 cm, with their heads stabilized via a chinrest in a dark room.

Task and design The participants performed an orientation identification task in one of three conditions, with the trial structures depicted in Fig. 1. In the Set Size 1 condition (SS1), a single grating (test) was presented that was followed by a mask for 200 ms. The test grating could be in one of ten orientations: 15º, 30º, 45º, 60º, 75º, 105º, 120º, 135º, 150º, and 165º (assuming that horizontal is 0º and vertical is 90º). After a retention interval of 500 ms, a square outline appeared (size, 1.1º; line thickness, 0.1º) at the grating's location, and participants had to indicate via keypress whether the top of the grating was rotated to the left (counterclockwise) or the right (clockwise) of vertical. The next trial started about 1 s after this response.

The sequential condition (Seq) was similar to the SS1 condition, except that two test gratings were presented in two locations (out of the four possible locations, randomly determined on each trial). Each test grating was presented and followed by a mask (200 ms). A 500-ms retention interval period separated the first mask and the second test grating. Both test gratings were randomly determined to be in one of the ten possible orientations, with the constraint that the two orientations were not identical. A 500-ms retention interval followed the second mask, and then a square outline appeared in one of the two gratings' locations. Participants responded to whether the probed grating
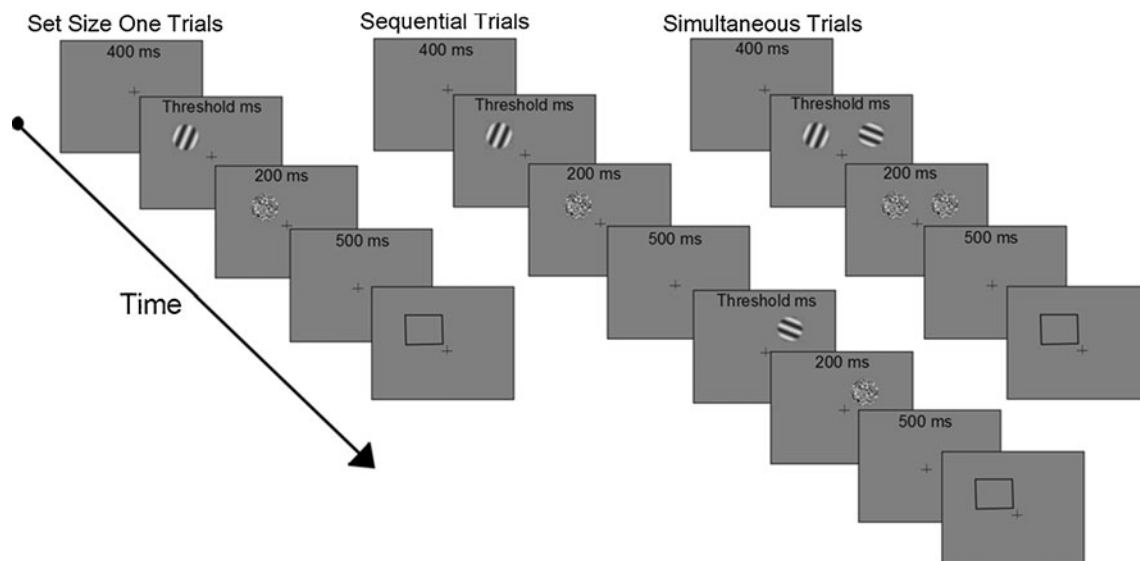

Fig. 1 Schematic of the three conditions in Experiments 1 and 2. In the Set Size 1 condition (left), a single grating was displayed. In the sequential condition (middle), two gratings appeared, one at a time. In the simultaneous condition (right), two gratings appeared at the same time. The duration that each grating was presented was held constant across conditions and was based on the individual participant's threshold. A noise mask followed each grating. A location probe (square outline) indicated the grating for the responses at the end of the trial. Participants were asked to indicate whether the probed stimulus was rotated clockwise or counterclockwise from vertical (Exp. 1), or whether the probed stimulus was in a cardinal (vertical or horizontal) or oblique orientation (Exp. 2). Experiment 1b was identical, except that the stimuli were moved farther into the periphery

was rotated to the left or the right of vertical. The simultaneous condition (Simu) was identical to the Seq condition, except that two test gratings were presented at the same time (and masked). In both the Seq and Simu conditions, participants were instructed to remember the orientation of *both* gratings to prepare for a memory query. The location of the square outline was randomly determined to match one or the other of the test gratings on each trial.

The three presentation conditions (SS1, Seq, and Simu) were run in blocks of 60 trials, with a prompt at the beginning of each block informing participants of the block type. The blocks were arranged into two super-blocks, each containing a random sequence of the three block types, for a total of six blocks (i.e., two blocks, or 120 trials, per condition).

Thresholding procedure Prior to participating in the main task, all of the participants performed four blocks (two Seq and two Simu) of trials during which the stimulus exposure duration varied according to the method of constant stimuli. We used eight possible durations: 1, 2, 4, 8, 12, 16, 32, and 64 video frames, which corresponded to 10.4 ms at the shortest exposure and 667 ms at the longest exposure (96-Hz refresh rate). Each block consisted of eight trials of each stimulus duration, for a total of 64 trials per block. We used only the data from the two Seq blocks to find individual participant thresholds, and included the two Simu blocks to ensure that the participants had equivalent practice with both types of trials prior to participation in the main task.

We calculated each participant's percentage correct at each duration for these sequential blocks. The resulting data were fit with an exponential function:

$$pc = \delta + \gamma\left(1 - e^{-\beta t}\right),$$

where *pc* is the percentage correct, *t* is exposure duration, and δ, γ, and β are free parameters that control the shape of the psychometric function. In practice, we found that the exponential function fit duration threshold data better than a sigmoidal function, such as the Weibull. The data were fit with standard maximum likelihood methods, and the duration that produced ~80% correct for these sequential trials was used for the stimulus presentation duration for all conditions in the main task.

Results

The average exposure duration in this experiment was 104 ms (range 21–260 ms, *SD* = 72 ms). The percentage correct for each of the three presentation conditions (SS1, Seq, and Simu) was calculated for each participant. A

repeated measures analysis of variance (ANOVA) on these accuracy data (Fig. 2) revealed a main effect of condition [$F(2, 18) = 133$, $p < 10^{-10}$]. Follow-up paired *t* tests demonstrated that performance was significantly better for SS1 than for Seq [$t(9) = 15.4$, $p < 10^{-7}$] and that Seq performance was significantly better than Simu performance [$t(9) = 4.63$, $p = .001$]. This pattern was highly consistent across individual participants (see Fig. 2, top right panel). In this and all subsequent experiments, the reaction time data mirrored the accuracy data: Higher accuracy was associated with faster RTs. We found no indication of a speed–accuracy trade-off. We also investigated whether there was differential performance in the Seq condition for the test grating that appeared first ($M = .68$, $SD = .11$) or second ($M = .68$, $SD = .09$) in the sequence. No indication of such an order effect emerged [$t(9) = 0$, $p > .9$].

The finding of worse performance for the Seq than for the SS1 condition highlights the importance of holding the number of memoranda constant when investigating the consolidation process. In both conditions, a single item was presented at a time, which should have put the same demands on the consolidation process. Thus, the difference in performance between these conditions was likely due to the difference in the numbers of memoranda rather than to a difference in consolidation. When the memory load was held constant at two items, we found that sequential presentation of two gratings resulted in better memory for their orientation than did simultaneous presentation. This finding suggests that the process of consolidating information into VSTM has a very limited capacity: Two oriented gratings cannot be consolidated simultaneously as effectively as sequentially.

While we interpret these results as indicating a limit in the consolidation processes, it is possible that the reduced performance that we observed with simultaneous presentation was due to low-level perceptual effects that degraded the sensory representation of simultaneously presented stimuli prior to consolidation. For instance, some crowding effects between the two simultaneously presented gratings might have degraded the representations. We believe that this is unlikely. The critical spacing for crowding is roughly half of the eccentricity of the stimuli (Bouma, 1970; Pelli, Palomares, & Majaj, 2004), yet the shortest distance separating our two stimuli was 3.0º, and each stimulus was presented at 2.1º of eccentricity. Thus, our spacing was more than twice the critical spacing (1.05º) for crowding. In short, the two stimuli were likely outside of each other's integration fields. However, to experimentally rule out crowding (and other potential low-level perceptual effects), in Experiment 1b we attempted to replicate the finding with more peripherally presented stimuli. Presenting stimuli farther out into the periphery should minimize any low-level grouping or interference effects.

Fig. 2 Accuracy for Experiments 1a and 1b. The left panels show the mean proportions of correct responses for each presentation condition. Error bars indicate the standard errors of the means. The right panels show individual participant data comparing performance on the sequential (Seq) condition (abscissa) to performance in the other two conditions (ordinate). Most participants performed better in the Set Size 1 (SS1) than in the Seq condition (majority of circles above the unity line), as well as better in Seq than in the simultaneous (Simu) condition (majority of asterisks below the unity line). Note that the scale of the right graph for Experiment 1b goes down to .4, in order to allow us to plot one participant who performed just below chance on the simultaneous condition
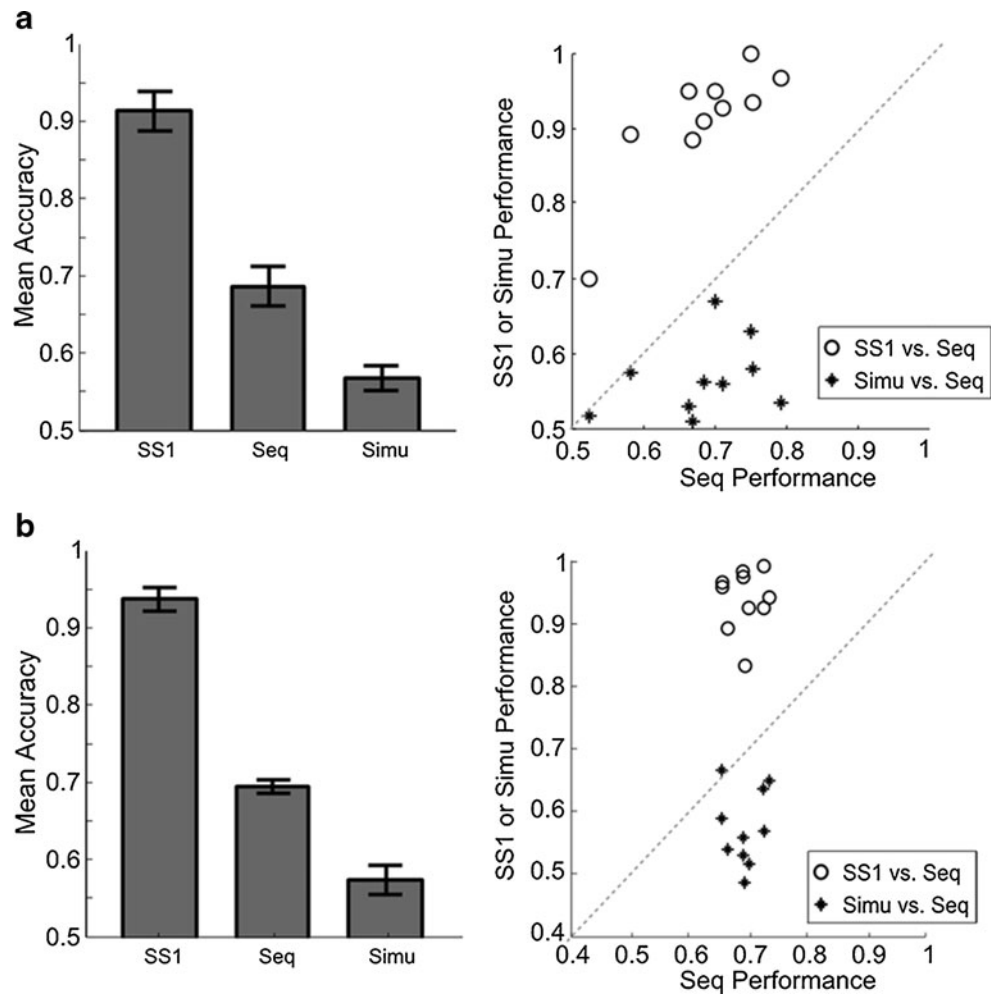
Experiment 1b

In Experiment 1b, the grating stimuli were moved farther into the periphery, at an eccentricity of 6º. This should have minimized any low-level perceptual interactions between two simultaneously presented visual stimuli. We also scaled the stimulus size to 2º and the spatial frequency to 1.5 cycles/deg, to compensate for reduced acuity in the periphery. All other aspects of the method were identical to those of Experiment 1a. Ten additional participants completed this experiment, none of whom had participated in Experiment 1a.

Results

The results mirrored those of Experiment 1a. The average exposure duration in this experiment was 127 ms (range 42–250 ms, $SD = 65$ ms). The percentage correct for each of the three presentation conditions (SS1, Seq, and Simu) was calculated for each participant (see Fig. 2). A repeated measures ANOVA on the accuracy data revealed a main effect of condition [$F(2, 18) = 212$, $p < 10^{-12}$]. Follow-up paired t tests demonstrated that performance was significantly better for the SS1 than for the Seq condition [$t(9) = 13.9$, $p < 10^{-6}$]. More importantly for the present experiment, Seq performance was still significantly better than Simu performance [$t(9) = 5.91$, $p < .001$]. Again, this pattern was highly consistent across participants (Fig. 2, bottom right panel), and we found no significant difference in Seq performance [$t(9) = 1.72$, $p > .1$] as a function of whether the test grating appeared first ($M = .71$, $SD = .05$) or second ($M = .68$, $SD = .04$).

The results of the experiment replicated the main finding of Experiment 1a. As we discussed before, it is important to hold the number of memoranda constant; when this was achieved, the sequential presentation of two gratings resulted in better memory for their orientation than did simultaneous presentation. Given that the stimuli were presented far from one another, our finding eliminates a low-level perceptual interference account of the observed decrement in the simultaneous condition. Thus, these results provide additional evidence that the process of consolidating information into VSTM has a very limited capacity.

This decrement for two simultaneously presented gratings was apparent even though the task did not require one to have a very precise representation of the gratings' orientations, but only to report whether the grating was rotated to the right or the left of vertical. We chose this type of task because it was similar to the task we had used in the color experiments (Mance et al., 2012). The stimuli used in that experiment were highly discriminable colors (red, blue, yellow, and green), and thus correct report in that experiment did not require very precise encoding of the specific hue. Despite the fact that both the color task and the orientation task required fairly coarse coding of the stimuli, we found different results. In the color experiments, people were capable of encoding two colors in parallel, whereas here they were unable to encode two oriented gratings in parallel. To investigate whether this difference in performance represented a difference in how the two types of features were being encoded, we further reduced the complexity of the orientation task in Experiment 2 to make it more similar to the color task.

Experiment 2

In our previous color task, only four colors were possible during the entire experiment; by contrast, Experiments 1a and 1b included ten possible orientations. Perhaps this difference in the number of potential targets made the simultaneous condition with gratings especially difficult. That is, one might be able to monitor two channels of information better when the number of possible targets is reduced. To explore this possibility, in this experiment we reduced the number of possible gratings to four, thereby equating the number of possible targets to that in our previous color study.

Method

Experiment 2 was identical to Experiment 1 except for two notable changes. First, the number of possible orientations was reduced to just four (−45º from vertical, vertical, +45º from vertical, and horizontal). On each trial of the Seq and Simu conditions, two orientations were randomly selected, with the constraint that they could not be identical. In addition, we altered the task from one of making a right–left judgment to judging whether the cued grating had been oriented in a cardinal (horizontal or vertical) or an oblique (±45º) orientation. This design allowed us to maximize the difference between the orientations of the stimuli and provided a clear categorical judgment about the orientation of the grating. Ten subjects participated in this experiment; seven of them had participated in Experiment 1a.

Results

The average exposure duration in this experiment was 50 ms (range 10–83 ms, SD = 25 ms). A repeated measures ANOVA on identification accuracy (Fig. 3) revealed a main effect of condition [$F(2, 18)$ = 10.0, p = .001]. Follow-up paired t tests demonstrated that performance was significantly better for the SS1 than for the Seq condition [$t(9)$ = 2.47, p < .05], and also significantly better for Seq than for the Simu condition [$t(9)$ = 2.52, p <.05]. This pattern was highly consistent across individual participants (Fig. 3, right
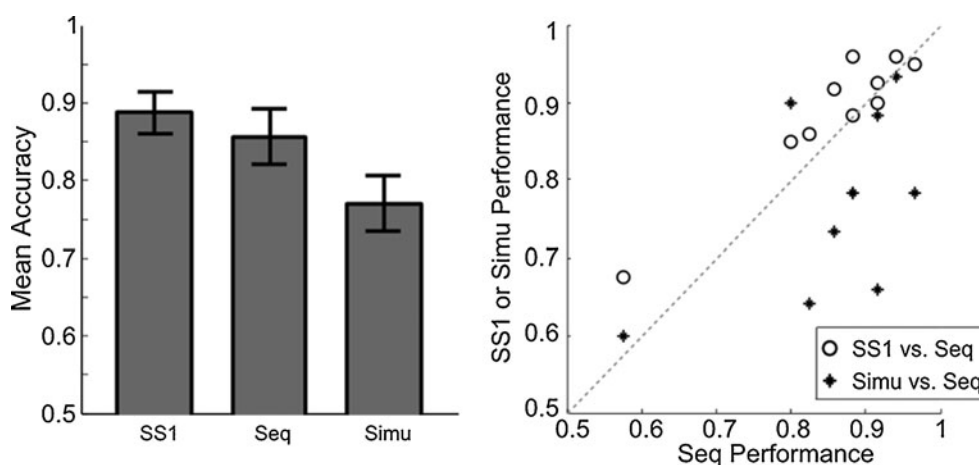

Fig. 3 Accuracy for Experiment 2. The left panel shows the mean proportion of correct responses for each presentation condition. Error bars indicate the standard errors of the means. The right panel shows individual participant data comparing performance on the sequential (Seq) condition (abscissa) to performance in the other two conditions (ordinate). Most participants performed better in the Set Size 1 (SS1) than in the Seq condition (majority of circles above the unity line), as well as better in Seq than in the simultaneous (Simu) condition (majority of asterisks below the unity line)

panel). Within the Seq condition, we again found no indication that performance differed for target gratings that appeared first ($M = .84$, $SD = .10$) or second ($M = .87$, $SD = .13$) in the sequence [$t(9) = 1.05$, $p > .3$].

In short, while overall performance in both the simultaneous and sequential conditions was higher in this experiment than in Experiment 1, the pattern was the same as in that experiment: Participants were better able to encode the orientation of two sequentially presented than of two simultaneously presented gratings. As in Experiment 1, these results suggested that the encoding of orientation is a limited-capacity process.

These results showing an inability to encode two orientations in parallel again differed from our results showing that two colors can be encoded in parallel. While these disparate findings may indicate that the encodings of color and orientation are processed differently, one further distinction between our present methods and the previous color methods might have been responsible for the discrepancy. Specifically, in our previous work, the presentation of the test stimuli was followed by a probe color, and participants had to indicate whether the probe matched the color of either of the test stimuli. As such, the procedure in those color experiments did not require binding the test colors to their particular locations. By contrast, Experiments 1 and 2 utilized a location probe and asked participants to report the orientation of the grating at that location. Thus, these tasks required binding the orientation feature to a particular spatial location. In Experiment 3, we investigated whether a task that required no such binding would still produce a severe capacity limit for grating stimuli.

Experiment 3

There is evidence that features are bound to the locations in which they appear during the initial consolidation of information into VSTM. However, this binding between features and locations seems to dissipate as more abstract VSTM memory representations are formed (Logie, Brockmole, & Jaswal, 2011; Woodman, Vogel, & Luck, 2012). Thus, after consolidation the memory for features may not be bound to their locations. It is possible that requiring observers to maintain the binding of features to locations beyond this early stage would require a serial or limited-capacity system, but that the encoding of features unbound to spatial locations could be done in parallel (Treisman & Gelade, 1980; Treisman & Zhang, 2006). Experiment 3 was designed to test this location binding explanation of the disparate color and orientation results by using an orientation matching task that did not require location binding.

Method

Experiment 3 was identical to Experiment 2, except that instead of a square outline a grating appeared at fixation, and participants had to indicate whether the probe orientation matched the test grating orientation (SS1) or whether it matched *either* of the two test gratings' orientations (Seq and Simu). On half of the trials, the probe matched the test grating (SS1) or one of the test gratings (Seq and Simu). In the latter case, the probe matched either of the test gratings with 50% probability. In the nonmatch trials, the probe orientation was randomly sampled from the three (SS1) or two (Seq and Simu) nonpresented orientations on that trial. This task did not require observers to bind a given orientation with a particular location and was essentially identical to the task we had previously used with color patches of stimuli. Twelve observers participated in this experiment. Of these, one had participated in Experiment 1a, one had participated in Experiment 2, and four had participated in both experiments.

Results

The average exposure duration in this experiment was 55 ms (range 10–125 ms, $SD = 34$ ms). A repeated measures ANOVA on matching accuracy (Fig. 4) revealed a significant main effect of presentation [$F(2, 22) = 32.8$, $p < 10^{-6}$], and follow-up paired t tests confirmed that performance was better for the SS1 than for the Seq condition [$t(11) = 6.75$, $p < 10^{-4}$], and that the Seq condition once again yielded better performance than the Simu condition [$t(11) = 3.95$, $p = .002$]. Again, this pattern was highly consistent across individual participants (Fig. 4, right panel). In short, this pattern of results again replicated the findings from Experiments 1 and 2. However, unlike those prior experiments, within the Seq condition, performance was significantly better [$t(11) = 2.7$, $p < .05$] when the target grating appeared second ($M = .86$, $SD = .09$) rather than first ($M = .75$, $SD = .09$). Note that these values are hit rates, as on nonmatch trials the first versus second target was undefined, such that separate false alarm rates could not be calculated. The order effect suggests that performance declined as the retention interval increased, but only for the matching task, which required a more precise representation of orientation than did the more categorical judgments in Experiments 1 and 2. Importantly, the presence of this order effect cannot explain the worse performance for the Simu than for the Seq condition, as the retention interval in the Simu condition was equal to the shorter of the two retention intervals in the Seq condition.
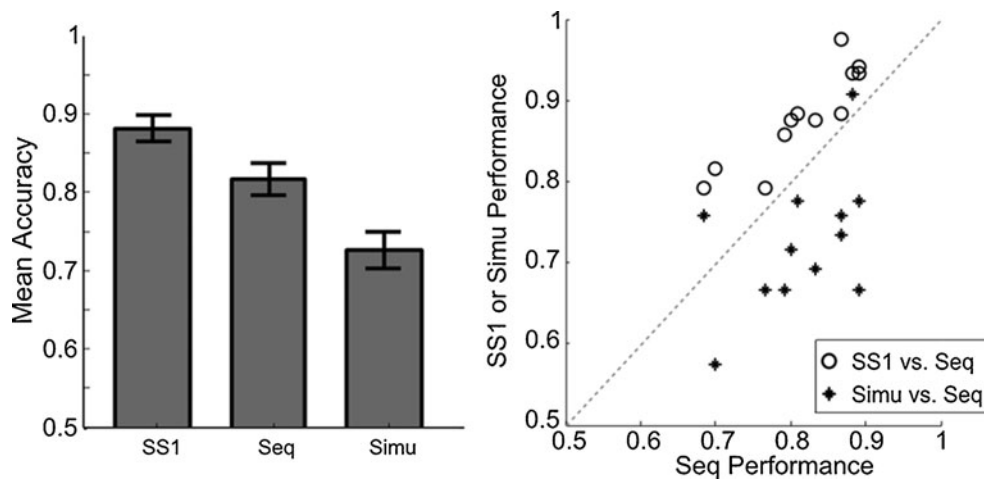
Fig. 4 Accuracy for Experiment 3. The left panel presents the overall proportions correct. Error bars indicate the standard errors of the means. The right panels show individual participant data comparing performance on the sequential (Seq) condition (abscissa) to performance in the other two conditions (ordinate). Most participants performed better in the Set Size 1 (SS1) than in the Seq condition (majority of circles above the unity line), as well as better in Seq than in the simultaneous (Simu) condition (majority of asterisks below the unity line)

General discussion

In this study, we investigated the capacity limit of the process for consolidating orientation information into VSTM. Across three experiments, we found a very consistent pattern of results at both the group mean level and the individual participant level: Performance was better for trials with a single grating rather than two sequentially presented gratings, and performance was better in the latter condition than with two simultaneously presented gratings. The finding of better performance in the sequential than the simultaneous conditions occurred across a number of methods of responding and with varying numbers of possible stimulus orientations. These results thus reflect the general properties of the consolidation of orientation information, rather than being a special case resulting from particular experimental methodologies.

We note that the poorer performance in the simultaneous condition occurred despite the fact that both parts of the simultaneous–sequential method had the same memory load and number of possible responses. The importance of equating these factors is highlighted by the finding of better performance with a single grating than with two sequential gratings. The consolidation processes for both the Set Size 1 and the sequential conditions should have been equivalent, since in both conditions a single grating was presented at a time. The only difference was that participants needed to store two items in the sequential condition, but needed only store a single item in the Set Size 1 condition. Despite the fact that two items is well below the capacity of VSTM (Luck & Vogel, 1998; Pashler, 1988), performance was worse with a higher memory load. This result could have been due to lower memory precision with higher memory load (Bays & Husain, 2008; Zhang & Luck, 2008). However,

given that we used relatively categorical responses that required only coarse orientation encoding, we believe that the difference between the Set Size 1 and sequential conditions was probably not due to memory precision. Instead, we believe that this difference was more likely due to increased decision noise at the response stage (Eckstein et al., 2000; Palmer et al., 2000). Regardless of the source of this difference in performance between these two conditions, the difference highlights the need to equate such factors as memory load in order to make strong claims about the consolidation process. The method that we used here did equate these factors for the sequential and simultaneous conditions. Thus, the consistent finding of worse performance in the simultaneous condition provides us with confidence that this difference is due to a limited ability to adequately consolidate both items in the simultaneous condition, rather than to other factors. Our results thus provide strong evidence that the consolidation of orientation information into VSTM is severely capacity limited.

Interestingly, this finding of a capacity limit for consolidation of orientations differs from our previous findings with colored patches as stimuli (Mance et al., 2012). In that series of experiments, we used a similar simultaneous–sequential procedure and found that performance was equivalent when two color patches were presented sequentially or simultaneously, suggesting that observers could consolidate two color patches as efficiently as one. The progression of the experiments here was designed to make the orientation discriminations more and more similar to the discriminations used in those color experiments. Despite multiple attempts to make the orientation task more categorical and less dependent on a precise representation of the stimuli (Exps. 1 and 2), we continued to find the severe capacity limit in the consolidation of orientations.

In Experiment 3, we used a matching task with a probe stimulus to be consistent with our previous color experiments, but again found a severe capacity limit. This experiment was basically identical to Experiment 1 in Mance et al. (2012), except for the stimuli. Thus, we can directly compare the results from these two experiments, which are plotted in Fig. 5. We ran a 2 (condition: Sim vs. Seq) × 2 (experiment: present Exp. 3 vs. Mance Exp. 1) mixed-factor ANOVA, which revealed no main effect of experiment [$F(1, 28) = 2.04$, $p = .17$], but a significant effect of condition [$F(1, 28) = 11.09$, $p = .002$], as well as a significant interaction [$F(1, 28) = 9.83$, $p = .004$]. The interaction was due to equivalent performance for the simultaneous and sequential conditions in the Mance et al. (2012) experiment, but degraded performance in the simultaneous relative to the sequential condition in the present Experiment 3. Importantly, the overall performance levels across experiments were similar, hence this cannot explain the differential results between orientation and color. Thus, we conclude that a genuine difference exists in the capacity to consolidate color and orientation information.

Others have also reported a difference in consolidation between orientation and color. For instance, Woodman and Vogel (2008) and Stevanovski and Jolicœur (2011) varied set size in a change detection paradigm and found that memory for orientation was more severely affected by increases in set size than was color. However, as we mentioned in the introduction, these set size manipulations did not equate memory load and decision noise, so the changes in performance with increased set size might have been due to factors other than the initial encoding. Our method equated memory load and decision noise between the Seq and Simu conditions, and thus the observed differences between
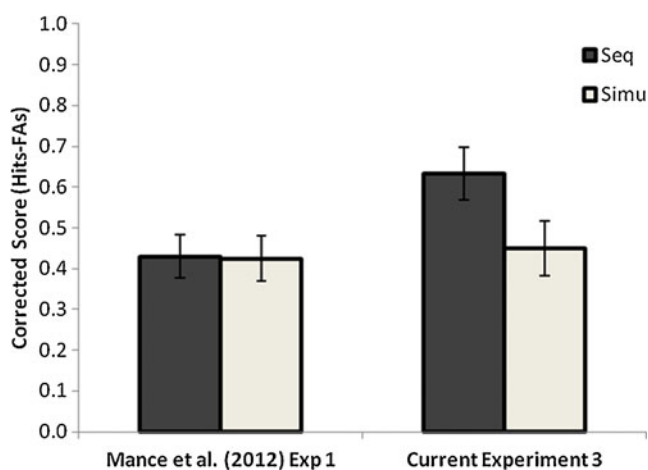

Fig. 5 Comparison of Experiment 3 to Mance et al.'s (2012) Experiment 1. The panel presents the mean corrected scores (hits – false alarms) for Mance et al.'s experiment, which used color stimuli, and the present Experiment 3, which used the same method with orientation stimuli. Error bars represent the standard errors of the means

color and orientation can be attributed to differences in the initial encodings of these features.

Why might the consolidation of two color patches proceed simultaneously, while the consolidation of two oriented gratings is capacity limited? One possibility is that color may allow rapid activation of preexisting verbal codes (at least for highly discriminable, prototypical hues), whereas orientation necessitates the creation of a visual representation; consolidation might be easier for multiple verbal than for multiple visual codes. We find this explanation unlikely, for two reasons. First, across our experiments, we manipulated the suitability of a verbal code to encode orientation. The categorical response of Experiments 1 and 2 presumably allowed for verbal codes (e.g., "left," "right," "vertical," "horizontal," or "tilted"), while the matching task in Experiment 3 likely required participants to use a visual code. This manipulation, however, had no impact on the pattern of results: We consistently found better performance for the sequential than the simultaneous condition. Second, it is not necessarily the case that verbal codes are easier to consolidate. For example, Stevanovski and Jolicœur (2007) investigated working memory consolidation using colored stimuli in a psychological refractory period paradigm. They found that activating a verbal code for color was more, rather than less, capacity-demanding in terms of the use of a central mechanism. Thus, we do not see strong reason to believe that verbal coding contributed to the differences between the consolidation of color and orientation.

A second possibility, which we favor, is based on the observation that color seems to occupy a richer perceptual space than does orientation. Whereas colors can be experienced in terms of hue, saturation, and brightness, orientation is a unidimensional quantity. In essence, a richer/bigger perceptual space might reduce interference among competing features, thus allowing consolidation to proceed with more efficiency. Incidentally, a bigger perceptual space can also afford categorical perceptual qualities (however, these qualities are not necessarily tied to verbal/semantic representations). To make this explanation more concrete, one can consider its potential neural implementations.

A possible neural mechanism of VSTM consolidation is to establish synchronized firing in the neuronal assemblies responsible for coding the features of an object (Luck & Vogel, 1998; Raffone & Wolters, 2001). It is possible that the neuronal assemblies that represent two distinct colors are relatively nonoverlapping, whereas those representing two orientations would have significant overlap. In the latter case, it might be impossible to establish two different synchronized assemblies because of interference, and one might need to resort to a serial mechanism to establish one assembly at a time. This neural explanation is rather speculative at this point, and more research would be needed to evaluate its validity. However, what is clear is that the processes for consolidating

information into VSTM differ for color and orientation information, suggesting that not all visual features are the same, and it may be unwarranted to generalize findings using one type of visual feature to other visual features.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106–111. doi:10.1111/j.0963-7214.2004.01502006.x

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. doi:10.1111/j.1467-9280.2007.01949.x

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7,* 66–80.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321,* 851–854. doi:10.1126/science.1158023

Becker, M. W., & Pashler, H. (2002). Volatile visual representations: Failing to detect changes in recently processed information. *Psychonomic Bulletin & Review, 9,* 744–750. doi:10.3758/BF03196330

Becker, M. W., Pashler, H., & Anstis, S. M. (2000). The role of iconic memory in change-detection tasks. *Perception, 29,* 273–286.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226,* 177–178.

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. doi:10.1037/0096-1523.21.1.109

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222. doi:10.1146/annurev.ne.18.030195.001205

Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1416–1438.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics, 62,* 425–451. doi:10.3758/BF03212096

Franklin, A., Pilling, M., & Davies, I. (2005). The nature of infant color categorization: Evidence from eye movements on a target detection task. *Journal of Experimental Child Psychology, 91,* 227–248.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology, 195,* 215–243.

Hubel, D. H., & Wiesel, T. N. (1974). Sequence regularity and geometry of orientation columns in the monkey striate cortex. *Journal of Computational Neurology, 158,* 267–293.

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36,* 138–202. doi:10.1006/cogp.1998.0684

Logie, R. H., Brockmole, J. R., & Jaswal, S. (2011). Feature binding in visual short-term memory is unaffected by task-irrelevant changes of location, shape, and color. *Memory & Cognition, 39,* 24–36. doi:10.3758/s13421-010-0001-z

Luck, S. J., & Vogel, E. K. (1998). Response from Luck and Vogel. *Trends in Cognitive Sciences, 2,* 78–80.

Mance, I., Becker, M. W., & Liu, T. (2012). Parallel consolidation of simple features into visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 429–438.

Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research, 38,* 1805–1815. doi:10.1016/S0042-6989(97)00349-0

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461–488. doi:10.1037/h0084327

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40,* 1227–1268.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44,* 369–378. doi:10.3758/BF03210419

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12):12, 1136–1169. doi:10.1167/4.12.12

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16,* 283–290. doi:10.3758/BF03203943

Pratt, J., & Hommel, B. (2003). Symbolic control of visual attention: The role of working memory and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 835–845. doi:10.1037/0096-1523.29.5.835

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13,* 766–785.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373. doi:10.1111/j.1467-9280.1997.tb00427.x

Scharff, A., Palmer, J., & Moore, C. M. (2011). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 813–833.

Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology, 93,* 72–83. doi:10.1037/h0032453

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences, 12,* 342–348. doi:10.1016/j.tics.2008.05.007

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

Stevanovski, B., & Jolicœur, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition, 15,* 532–563.

Stevanovski, B., & Jolicœur, P. (2011). Consolidation of multifeature items in visual working memory: Central capacity requirements for visual consolidation. *Attention, Perception, & Psychophysics, 73,* 1108–1119. doi:10.3758/s13414-011-0099-5

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136. doi:10.1016/0010-0285(80)90005-5

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition, 34,* 1704–1719.

Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision, 3*(1):9, 86–94. doi:10.1167/3.1.9

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. doi:10.1037/0096-1523.32.6.1436

West, G. L., Pun, C., Pratt, J., & Ferber, S. (2010). Capacity limits during perceptual encoding. *Journal of Vision, 10*(2):14, 1–12. doi:10.1167/10.2.14

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12):11, 1120–1135. doi:10.1167/4.12.11

Wilkinson, D., Ko, P., Milberg, W., & McGlinchey, R. (2008). Impaired search for orientation but not color in hemi-spatial neglect. *Cortex, 44,* 68–78.

Williams, L. G. (1966). The effect of target specification on objects fixated during visual search. *Perception & Psychophysics, 1,* 315–318.

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 693–716. doi:10.1037/0096-1523.26.2.693

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. doi:10.3758/PBR.15.1.223

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2012). Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognition, 20,* 1–28.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453,* 233–235. doi:10.1038/nature06860