Serial Consolidation of Orientation Information Into Visual Short-Term Memory

Taosheng Liu[1,2] and Mark W. Becker[1]
[1]Department of Psychology and [2]Neuroscience Program, Michigan State University

Abstract

Previous research suggests that there is a limit to the rate at which items can be consolidated in visual short-term memory (VSTM). This limit could be due to either a serial or a limited-capacity parallel process. Historically, it has proven difficult to distinguish between these two types of processes. In the present experiment, we took a novel approach that allowed us to do so. Participants viewed two oriented gratings either sequentially or simultaneously and reported one of the gratings' orientation via method of adjustment. Performance was worse for the simultaneous than for the sequential condition. We fit the data with a mixture model that assumes performance is limited by a noisy memory representation plus random guessing. Critically, the serial and limited-capacity parallel processes made distinct predictions regarding the model's guessing and memory-precision parameters. We found strong support for a serial process, which implies that one can consolidate only a single orientation into VSTM at a time.

Adaptive behavior often requires temporary storage of information in a more durable and consciously accessible form. In vision, this more durable store is commonly described as visual short-term memory (VSTM). It is generally accepted that VSTM has a capacity limit of only three to four items, thus imposing a fundamental limit on visual cognition (Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974). In theory, people could minimize the practical impact of this storage limit and still function well if they were able to rapidly consolidate new behaviorally relevant items into the VSTM buffer as needed (Ballard, Hayhoe, & Pelz, 1995; O'Regan, 1992). This view, however, suggests that the ability to rapidly consolidate information into VSTM is another potential limiting factor in visual cognition.

Indeed, there seems to be a limit to the amount of information that can be simultaneously consolidated into VSTM. For example, studies have found that a longer presentation time is necessary to consolidate more items (Jolicœur & Dell'Acqua, 1998; Vogel, Woodman, & Luck, 2006). Furthermore, Zhang and Luck (2008, Experiment 4) suggested that the consolidation is a discrete (all or none) process in which additional time allows more

items to be consolidated. However, these results cannot reveal the nature of this consolidation limit. Specifically, this limit could result either from a strictly serial process, in which a cognitive bottleneck allows only one item to be consolidated at a time, or from a limited-capacity parallel process, in which two items can be consolidated simultaneously but, because of limits in the bandwidth, each with less precision. Determining whether the rate of consolidation is limited because of a strictly serial process or because of a limited-capacity parallel process has important implications for how one conceptualizes the underlying cognitive architecture and for understanding the fundamental limits that it imposes on visual cognition (Logan, 2002; Townsend & Wenger, 2004). Yet discriminating between these two possibilities has proven to be extremely difficult. The commonly used behavioral measures of reaction time and proportion of correct responses

Corresponding Author:
Taosheng Liu, Department of Psychology, Michigan State University, 316 Physics Rd., East Lansing, MI 48824
E-mail: tsliu@msu.edu

are often too coarse to differentiate between the two alternatives because of model mimicry (Townsend, 1990).

In the present experiment, we used a sequential-simultaneous paradigm to investigate the nature of the limit on VSTM consolidation (Duncan, 1980; Hoffman, 1978; Scharff, Palmer, & Moore, 2011a, 2011b; Shiffrin & Gardner, 1972). In the sequential condition, two items were presented one at a time, whereas in the simultaneous condition, the two items were presented at the same time. The per-item presentation duration was the same for both conditions. Worse performance in the simultaneous than in the sequential condition would converge with previous evidence of a limit on the rate of consolidation (Jolicœur & Dell'Acqua, 1998; Vogel et al., 2006). To further examine the nature of the limit, we obtained a continuous measure of the precision of consolidated information in VSTM via a recall procedure. If consolidation is a limited-capacity parallel process, memory representations will be less precise as more items need to be simultaneously consolidated, which should result in worse performance in the simultaneous condition than in the sequential condition. If consolidation is a serial process, worse performance in the simultaneous condition should reflect a mixture of two types of trials: trials on which the item was consolidated, which should have equivalent precision to that in the sequential condition, and trials on which the target was not consolidated and the participant guessed at random. These predictions can be tested with a mixture model that quantifies the precision and guessing rate separately (Zhang & Luck, 2008).

Method

Participants

Twelve graduate and undergraduate students at Michigan State University gave informed consent and were compensated at the rate of $10 per hour for their participation. All experimental protocols were approved by the university's institutional review board.

Stimuli, task, and design

The stimuli were circular sinusoidal gratings followed by noise masks (see the Supplemental Material available online for details of stimulus parameters), both of which were presented at the corners of an imaginary square (eccentricity: 3°). Participants performed an orientation-recall task in three conditions (trial structures for each condition are depicted in Figure 1). In the set-size-one condition, a single grating was presented; in the sequential condition, two gratings were presented in succession in two locations; in the simultaneous condition, two gratings were presented at the same time in two locations.

The locations in the sequential and simultaneous conditions were randomly sampled from the four possible locations on each trial. All gratings were presented for the same duration (150 ms) and subsequently masked for 200 ms. The orientation of each grating was randomly set to 1 of 12 possible values: 10°, 24°, 38°, 52°, 66°, 80°, 100°,114°, 128°, 142°, 156°, 170° (assuming horizontal is 0°).

In all conditions, a location cue (a 1.5° square outline) appeared at the end of each trial in one of the stimuli's location, along with an adjustable probe grating (presented at fixation). Participants adjusted the probe's orientation to match that of the cued grating. Four keys were used to rotate the probe grating (initial orientation was always vertical): two coarse adjustment keys and two fine adjustment keys that rotated the probe by ±4° and ±1° per key press, respectively. Participants were told to make adjustments until they were satisfied, at which point they pressed the space bar to complete the trial. The next trial started about 1 s after their response. In the sequential and simultaneous conditions, because grating orientations were randomly sampled from the 12 possible orientations, the two gratings had identical orientations on a small proportion of trials (~8%). We removed these trials from all analyses.

The three presentation conditions were run in blocks of 50 trials each, with a prompt at the beginning of each block informing participants of the block type. There were two superblocks, each containing a random sequence of the three block types, for a total of six blocks (two blocks per condition). Before the experiment started, participants practiced the orientation-adjustment task in the set-size-one condition for 20 trials.

Data analysis

For each trial, we calculated the offset (error) in the recalled orientation by subtracting the participant's orientation setting from the true orientation of the cued grating. For descriptive data analysis, we computed the arithmetic mean and the variance of the offset for each participant. For model fitting, we fit the offset data with a model that assumes observed performance results from a mixture of two types of trials. On a certain proportion of trials (g), participants hypothetically did not consolidate the stimulus into VSTM and simply guessed the orientation randomly, which should produce a uniform distribution. On the remaining trials, participants hypothetically consolidated the stimulus orientation into VSTM, which conformed to a circular normal distribution with a mean (μ) and standard deviation (σ). The model was fit to the observed offset data (both aggregate data and individual data) using standard maximum-likelihood methods (Myung, 2003). For more details on model-fitting procedures, see the Supplemental Material.
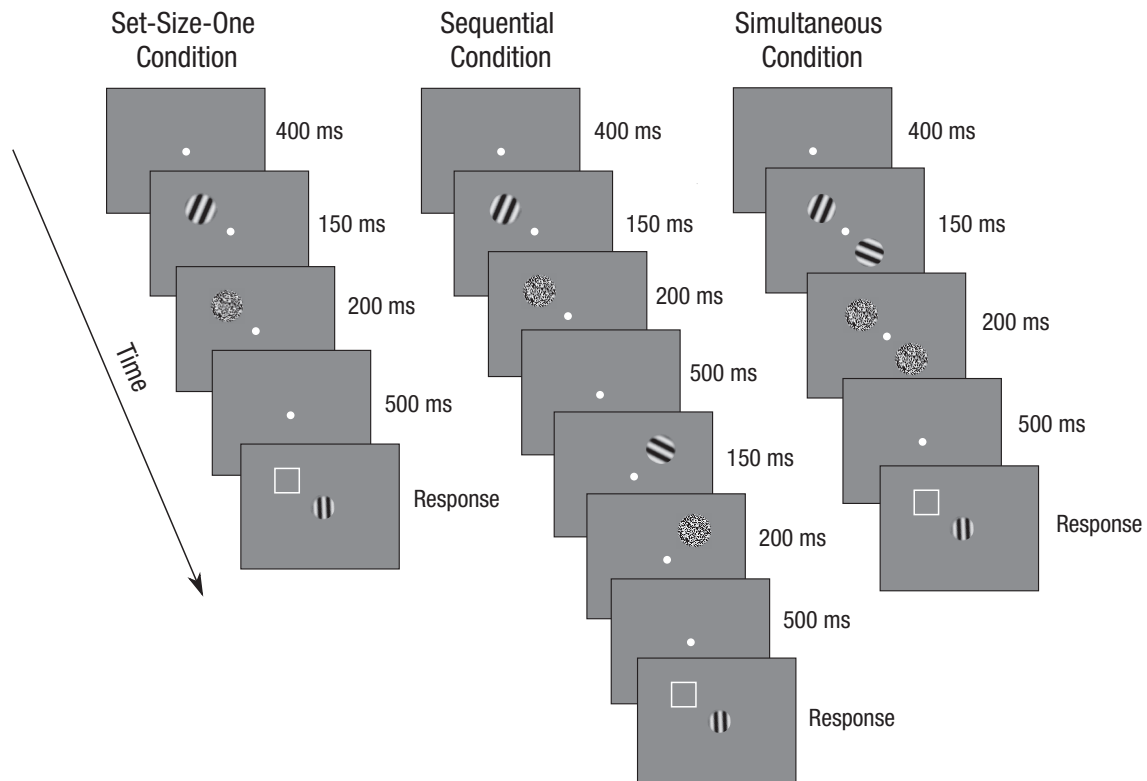
Fig. 1. Schematic of the trial structures in the three conditions. In all conditions, a sinusoidal grating was presented at one of four corners of an imaginary square, followed by a noise mask at the same location. In the set-size-one condition, one grating was presented; in the sequential condition, two gratings were presented one at a time; in the simultaneous condition, two gratings were presented at the same time. In all conditions, stimulus presentation was followed by a cue at one of the stimuli's location (square outline) along with a probe stimulus in the center. Observers adjusted the probe's orientation to match that of the cued grating.

Results

Raw offset data

We evaluated the bias and variability in orientation-recall performance, indexed by the mean and variance of the offset, respectively. Participants, on average, reproduced the true orientation of the cued grating without systematic bias (Fig. 2a); a one-way repeated measures analysis of variance showed that mean offsets did not differ significantly across the three conditions, $F(2, 22) < 1$, and one-sample t tests showed that none of the mean offsets differed significantly from zero (all $ps > .13$). Response variability, however, differed greatly across presentation conditions (Fig. 2b). Because variance was not normally distributed, we transformed the variance by taking its logarithm (Fig. 2c), which was significantly different across conditions, $F(2, 22) = 90.8$, $p < 10^{-10}$, $\eta_p^2 = .89$. All pairwise comparisons were significant (paired t test, all $ps < 10^{-4}$). Thus, recall of the target orientation became progressively more variable from the set-size-one to the sequential condition and from the sequential to the

simultaneous condition. This pattern of results was highly consistent across participants (Fig. 2d).

Model fit

We used a mixture model to evaluate how simultaneous presentation of two items affected the precision (σ) and guess rate (g) separately. If the performance decline due to simultaneous presentation can be explained solely by a decrease in precision, this would imply a limited-capacity parallel process—two items can be encoded in parallel but with less precision for each item. Conversely, if the performance decline can be explained solely by an increase in the guess rate, this would imply a serial process—increasing the number of simultaneously presented items affects the probability of successful consolidation but not the precision of memory for those items that were consolidated.

We fit the mixture model to the aggregate data (Figs. 3a–3c). Overall, the data were well fit by a mixture model (all $ps > .4$, assessing the divergence between the
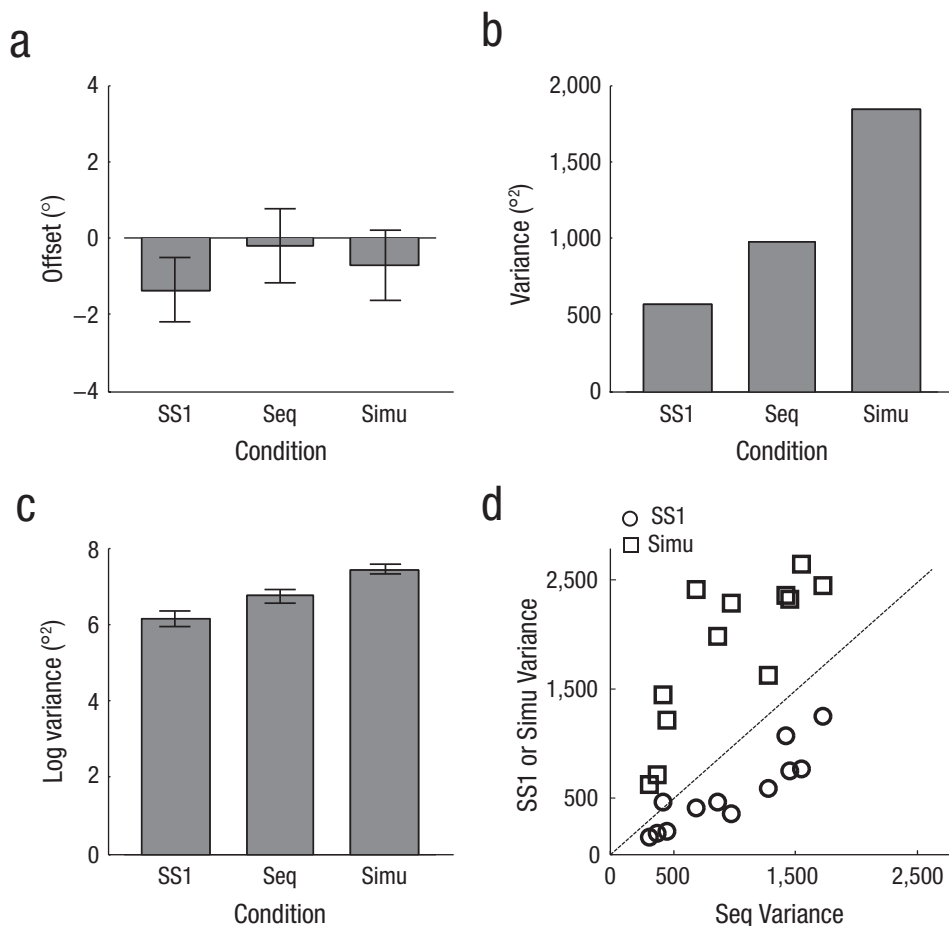
Fig. 2. Performance data. The (a) grand-average mean offset between the recalled orientation and the actual orientation of the cued grating, (b) average variance in offset, and (c) average log variance are shown separately in each graph for the set-size-one ("SS1"), sequential ("Seq"), and simultaneous ("Simu") conditions. Error bars in (a) and (c) indicate ±1 *SEM*. The scatter plot (d) shows individual-participant variance in the set-size-one or simultaneous condition as a function of variance in the sequential condition. For points above the diagonal, variance is greater in the set-size-one or simultaneous condition; for points below the diagonal, variance is greater in the sequential condition. Each symbol represents data from a single participant.

sample data and the model via Kolmogorov-Smirnov tests). The only substantial difference in parameter values among conditions was the guessing parameter (*g*; see insets in Figs. 3a–3c). To evaluate the statistical reliability of these results, we fit the mixture model to individual participant data and performed statistical tests on model parameters (Figs. 3d–3f). The mean parameter (μ) did not differ significantly across conditions, $F(2, 22) < 1$, with no condition significantly different from zero (one-sample *t* test, all *p*s > .27). The precision parameter (σ) also did not significantly differ across conditions, $F(2, 22) < 1$. However, the guess-rate parameter (*g*) showed a significant effect of condition, $F(2, 22) = 21.0, p < 10^{-5}, \eta_p^2 = .66$. All pairwise comparisons were significant (paired *t* test,

all *p*s < .05). Thus, both aggregate and individual data analyses indicated that the difference in performance among the three conditions could be accounted for solely by a change in guess rate.

We attribute differential performance between the sequential and simultaneous conditions to limits in the consolidation process rather than to differences in retention interval. To assess the impact of retention interval on performance, we compared performance for the first and second stimulus in the sequential condition and found them to be nearly identical (see Fig. S1 in the Supplemental Material). Thus, effects due to memory decay or interference probably did not contribute significantly to performance in our task (see also Scharff et al., 2011b).
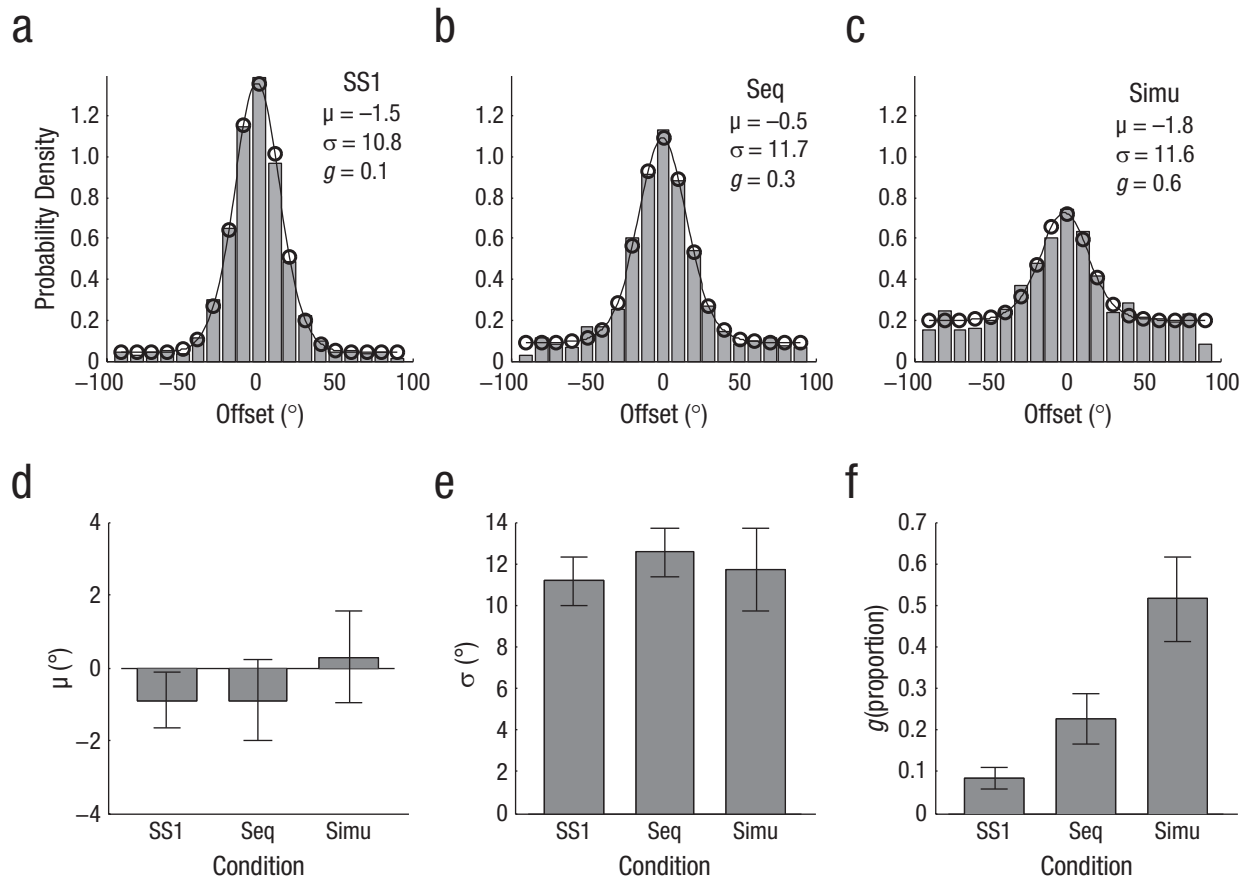
Fig. 3. Model-fit results. The histograms in the top row show probability-density functions for the offset of observed responses in the (a) set-size-one ("SS1") condition, (b) sequential ("Seq") condition, and (c) simultaneous ("Simu") condition. Model fits are illustrated by the continuous lines with open circles. The mean offset (μ), circular standard deviation (σ), and guess rate (g) are also shown for each condition. The graphs in the bottom row (d–f) show average values for these parameters from individual-level data fits in each of the three conditions. Error bars indicate ±1 *SEM*.
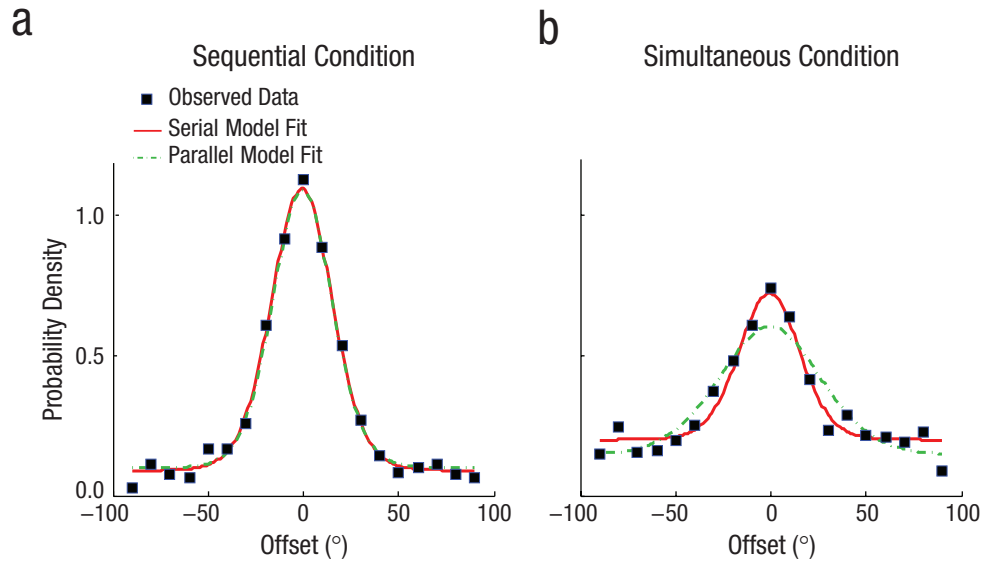
Model comparison

We also compared fits of a serial model and a parallel model to the observed data. The serial model had a single precision parameter (σ) and separate guessing parameters for the simultaneous and sequential conditions ($g_{sequential}$, $g_{simultaneous}$). The parallel model had a single guessing parameter (g) and separate precision parameters for the simultaneous and sequential conditions ($\sigma_{sequential}$, $\sigma_{simultaneous}$). Although both models fit data from the sequential condition well, the serial model fit data from the simultaneous condition noticeably better (Fig. 4). This pattern was confirmed by Kolmogorov-Smirnov tests indicating that both models fit the sequential data well ($p > .35$ for both models), whereas only the serial model ($p > .60$) and not the parallel model ($p < .05$) fit the simultaneous data well. We further used the Bayesian information criterion (BIC) to compare the relative likelihood of the two models (Raftery, 1995; Wagenmakers,

2007). For the observed aggregate data from the simultaneous and sequential conditions, the serial model was approximately 2×10^8 times more likely to fit the data than the parallel model (a change in BIC score of 38.8 in favor of the serial model). We also fit individual participant data and found that the serial model was more likely to fit 9 out of 12 participants' data. These model comparisons confirmed that the data were better accounted for by a serial model than by a parallel model.

Discussion

Although previous studies have shown that the rate at which information is consolidated into VSTM is limited (Chun & Potter, 1995; Dell'Acqua & Jolicœur, 2000; Jolicœur & Dell'Acqua, 1998; Vogel et al., 2006), the nature of this limit was unknown. Our study identifies this limit as a strictly serial bottleneck. We were able to distinguish this serial process from a limited-capacity

a

Sequential Condition

■ Observed Data
— Serial Model Fit
-·- Parallel Model Fit

b

Simultaneous Condition

Fig. 4. Results of the model comparisons. The graphs show the serial model's and parallel model's predicted probability-density functions for the offset of responses in (a) the sequential condition and (b) the simultaneous condition. Also shown are observed data, binned as in Figure 3. The parameters for the serial model were as follows: $\mu = -0.8°$, $\sigma = 11.7°$, $g_{sequential} = 0.3$, $g_{simultaneous} = 0.6$; the parameters for the parallel model were as follows: $\mu = -0.6°$, $g = 0.3$, $\sigma_{sequential} = 11.3°$, $\sigma_{simultaneous} = 20.5°$.

parallel process by obtaining a continuous measure of memory precision in a sequential-simultaneous paradigm and utilizing a mixture model to evaluate theoretical predictions (Zhang & Luck, 2008). We found that the decrease in performance for simultaneous compared with sequential presentation can be accounted for by higher guess rates with no loss of precision for consolidated items. This finding provides strong evidence that the consolidation limit results from a cognitive bottleneck that allows only one item to be consolidated at a time, rather than from a limited-capacity parallel process that would allow multiple items to be consolidated simultaneously, each with less precision.

In comparing results from the simultaneous with the sequential condition, we held other task factors (e.g., overall memory and decisional load) constant across tasks, thereby providing an ideal comparison. However, we also note that results from the set-size-one condition provided additional evidence against a parallel model. A parallel model would predict a decrease in precision between the set-size-one and the simultaneous condition, but we observed no loss of precision across these conditions. This finding of no change in the precision of the memory representation across any of our presentation conditions is strong support that consolidation of information into VSTM is a serial process. The increase in guess rate from the set-size-one to the sequential condition is likely due to decisional noise or higher memory load in the sequential condition (one item vs. two items).

Our results cannot be due to differential low-level factors such as masking, as any early perceptual factors were equated in the sequential and simultaneous conditions. Similarly, our results cannot be explained by a limit in storage capacity, as two items are well below the typical estimates of three to four items of VSTM storage limit (Luck & Vogel, 1997; Pashler, 1988), and the storage demand was identical (two items) in both the sequential and simultaneous conditions. Finally, a high-level decisional account might posit that participants sometimes confused the locations of the two stimuli and recalled the orientation of the uncued stimulus, and furthermore, this occurred more frequently in the simultaneous than in the sequential condition. If true, we would have expected some clustering of recalled orientation around the uncued stimulus. We reanalyzed our data in terms of the offset from the uncued orientation but did not find any evidence for such clustering (see Fig. S2 in the Supplemental Material). Thus, the performance decrement was not the result of confusion about which stimulus to recall.

We believe that our results reflect a bottleneck in the transfer of information between early perceptual processes and late memory and decisional processes—an inability to consolidate multiple VSTM representations at the same time. Given the importance of prefrontal and posterior sensory areas in working memory (Postle, 2006; Ranganath, 2006), we speculate that the observed serial-consolidation process might reflect a limit in bandwidth in the communication from sensory areas to prefrontal areas. It is also worth noting that we previously reported equivalent performance for two colors that were presented sequentially or simultaneously (Mance, Becker, &

Liu, 2012). Thus, it may be that some basic features are subject to the strict serial bottleneck we report here, whereas others are able to bypass this bottleneck. Further research is needed to systematically characterize VSTM consolidation for different feature dimensions and the underlying cognitive and neural mechanisms.

In summary, we have demonstrated that consolidation of orientation information into VSTM is subject to a strictly serial bottleneck. The ability to distinguish between serial and limited-capacity parallel processes has proven to be extremely difficult because of model mimicry (Townsend, 1990). Current experimental approaches involve complex factorial design and sophisticated analysis of reaction-time data (Townsend & Wenger, 2004). However, our novel approach was conceptually simple and easy to implement, and it potentially can be applied to other domains to distinguish these processes.

Acknowledgments

Declaration of Conflicting Interests

Funding

Supplemental Material

Additional supporting information may be found at http://pss .sagepub.com/content/by/supplemental-data

References

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

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.

Dell'Acqua, R., & Jolicœur, P. (2000). Visual encoding of patterns is subject to dual-task interference. *Memory & Cognition, 28*, 184–191.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87*, 272–300.

Hoffman, J. E. (1978). Search through a sequentially presented visual display. *Perception & Psychophysics, 23*, 1–11.

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

Logan, G. D. (2002). Parallel and serial processing. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology: Methodology in experimental psychology* (3rd ed., Vol. 4, pp. 271–300). New York, NY: Wiley.

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

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.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology, 47*, 90–100.

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.

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

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

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*, 23–38.

Raftery, A. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163.

Ranganath, C. (2006). Working memory for visual objects: Complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience, 139*, 277–289.

Scharff, A., Palmer, J., & Moore, C. M. (2011a). Evidence of fixed capacity in visual object categorization. *Psychonomic Bulletin & Review, 18*, 713–721.

Scharff, A., Palmer, J., & Moore, C. M. (2011b). 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*, 72–82.

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.

Townsend, J. T., & Wenger, M. J. (2004). The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review, 11*, 391–418.

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.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review, 14*, 779–804.

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