The bandwidth of VWM consolidation varies with the stimulus feature: evidence from event-related potentials

Renning Hao[a], Mark W. Becker[b], Chaoxiong Ye[a,c], Qiang Liu[a], Taosheng Liu[b]

a. Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, Dalian, China, 116029

b. Department of Psychology, Michigan State University, East Lansing, MI 48824

c. Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland, 40014

Running Head: Bandwidth of VWM consolidation

Number of manuscript pages: 19

Number of tables: 0

Number of figures: 7

Correspondence to:

Qiang Liu,
Research Center of Brain and Cognitive Neuroscience,
Liaoning Normal University,
Dalian 116029, China
E-mail: lq780614@163.com
Or
Taosheng Liu,
Department of Psychology,
Michigan State University,
East Lansing, MI 48824,
E-mail: tsliu@msu.edu

Acknowledgements:

This research was supported by grants from the National Natural Science Foundation of China (NSFC31571123) to Q. L., and from the National Institutes of Health (NIH R01EY022727) to T. L.

Abstract

Our previous work suggests that two colors can be consolidated into visual short-term memory (VSTM) in parallel without a loss of memory precision, while consolidation of two orientations is performed in a strictly serial manner. Those experiments compared VSTM performance for simultaneously and sequentially presented stimuli. However, there is still controversy about whether the bandwidth for consolidation is determined by the type of information. To further investigate this issue, here we measured electroencephalography while participants attempted to consolidate one, two or four simultaneously presented colors (Experiment 1) or orientations (Experiment 2) under limited presentation times. We used the contralateral delay activity (CDA) as an electrophysiological marker of the number of items that were consolidated. For colored stimuli, the CDA amplitude increased between set-size one and two but did not further increase for set size four. By contrast, for orientation, the CDA amplitude remained at the set size one amplitude as set size increased to two or four items. Furthermore, in a long exposure duration (300 ms) condition that did not limit the consolidation process, the CDA amplitude pattern indicated that VSTM capacity is limited to about three colored items and about two orientation items in our paradigm. Thus, the CDA effects observed in the short presentation time was not limited by VSTM storage, but rather by consolidation. These results are consistent with our previous behavioral research and suggest that the bandwidth of VSTM consolidation is determined by the stimulus feature.

Keywords: visual short-term memory, bandwidth, consolidation, contralateral delay activity

The public significance of the study

Previous studies on the bandwidth of visual short-term memory consolidation relied on behavioral measures, which were affected by assumptions about the underlying processes, and as a result have produced inconsistent conclusions. We used the contralateral delay activity, an electrophysiological measure to probe the bandwidth of consolidation for the first time. Our results show distinct patterns for consolidating color and orientation information, thus providing converging evidence for the behavioral studies and help resolve previous controversies.

Introduction

Early representation of visual information is fleeting, unprocessed, and subject to masking (Sperling, 1960). In order to further process visual stimuli, one must consolidate the volatile perceptual representation into a relatively stable and durable VSTM representation (Jolicoeur & Dell'Acqua, 1998; Vogel, Woodman, & Luck, 2006). It is generally accepted that visual short-term memory (VSTM) has a capacity of about 3~4 items (Luck & Vogel, 1997; Pashler, 1988). However, in recent years, researchers have begun to focus on the characteristics of this consolidation process, and have found that the capacity of this VSTM consolidation process is also limited (Jolicoeur & Dell'Acqua, 1998; Stevanovski & Jolicoeur, 2007, 2011; Vogel, Woodman, & Luck, 2006; West, Pun, Pratt, & Ferber, 2010; Zhang & Luck, 2008). For example, in a masked change detection task, Vogel et al. (2006) found performance initially improved as the interval between a memory array and masks increased, but then reached a plateau such that additional time did not further improve memory performance. However, such findings cannot reveal whether the limits of consolidation result from a serial process, which allows consolidation of only one item at a time, or from a limited-capacity parallel process, which allows consolidation of multiple items but with limited capacity.

Huang and colleagues suggested that items are consolidated into VSTM serially, as their Boolean map theory predicts that only one feature value per perceptual dimension can be consolidated at a time (Huang, 2010; Huang, 2015; Huang & Pashler, 2007; Huang, Treisman & Pashler, 2007). In their experiment, a change detection task was used with two simple color squares presented sequentially or simultaneously. Results showed that performance was better in the sequential than simultaneous condition, suggesting that it is difficult to consolidate two colors simultaneously (Huang et al., 2007). However, Mance, Becker and Liu (2012) pointed out that there were contingencies in Huang at al.'s study that may have allowed participants to predict the location and colors of items in the sequential condition, potentially leading to superior performance in the sequential condition that was unrelated to consolidation limits. Mance et al. (2012) also presented two color squares either sequentially or simultaneously in a change detection task. Importantly, they removed the contingencies and presented the stimuli for an exposure duration estimated to be just long enough to consolidate a single item. They found no difference in performance between the sequential and simultaneous condition, although a sequential advantage was observed when set size was increased to three or four. These results suggested that the parallel consolidation of colors is possible but the capacity of this parallel consolidation process was limited to two items. In a follow-up study, Miller, Becker & Liu (2014) used a color recall task in conjunction with model fitting to characterize the memory precision and guess rate. The results found no change in memory precision or guess rate between simultaneous and sequential presentation of two colors, suggesting that two colors could be consolidated in parallel without cost.

In another set of studies, the consolidation of orientation information was investigated by presenting two oriented gratings either sequentially or simultaneously. For these orientation stimuli, memory performance was better for sequential than simultaneous presentation, suggesting a severe capacity limit in consolidating orientation information (Becker, Miller & Liu, 2013). To further study the nature of the consolidation limit, Liu & Becker (2013) used a recall task in conjunction with model fitting and found that simultaneous presentation resulted an increase in guess rate with no change in memory precision, suggesting that the consolidation of orientation was a strictly serial process (Liu & Becker, 2013). They proposed that the capacity limit, or bandwidth, of VSTM consolidation depends on the stimulus feature. Orientation might require more processing resource than color, and thus while two colors can be consolidated in parallel, two orientations are consolidated in a strictly serial manner.

However, Rideaux et al. (2015, 2016) have challenged the conclusion regarding orientation, arguing that two orientations can also be consolidated via a limited-capacity parallel mechanism. Their main support for this claim (Rideuax, et al., 2016) is their finding that simultaneous presentation of two orientations produced higher guess rates and decreased precision compared to sequential presentation. However, it is worth noting that this was true only when the location of items was predictable in the sequential condition, which may have allowed covert shifts of attention prior to stimulus presentation that artificially increased memory precision in the sequential condition. When the location of items was unpredictable (their Experiment 2), their results were consistent with Becker, Miller & Liu (2013)'s: the simultaneous presentation of two orientations produced higher guess rates but equivalent memory precision as sequential presentation. While Liu & Becker (2013) have argued that this pattern is strong evidence for a strictly serial mechanism, they suggested that their data are evidence for a parallel consolidation mechanism. Their rationale was that the guess rate for sequential presentation was not 50%. While under an ideal condition a strictly serial mechanism should yield a 50% guess rate, we believe that there are a number of reasons why this precise 50% prediction might not hold. For instance, if the consolidation rate for the first item is very rapid in some trials, in those trials there may be adequate time to serially process the second item, thereby reducing the guess rate. In addition, it is possible that certain pairs of two simultaneously presented orientations support a strategy that allows for both orientations to be consolidated as a single stimulus; for instance, if the two orientation were about 90° apart, participants might encode a single angle of 90°, i.e., they could remember the two orientations as a single spatial configuration.

In short, while we believe the simultaneous vs. sequential paradigm is a powerful method for investigating consolidation into VSTM, the interpretation of results depends on the set of assumptions one makes about the task. Thus, it seems that a converging method, and particularly one that does not require comparisons across conditions with different numbers of stimuli per display, would help clarify that results are due to consolidation mechanisms rather than other strategic or low-level perceptual differences that may differ across simultaneous and sequential presentation conditions.

In the present experiments, we used an electrophysiological marker, contralateral delay activity (CDA), to probe VSTM consolidation processes. Importantly, this method does not require comparing a simultaneous to a sequential condition, thereby eliminating the interpretational issues raise above. The CDA is characterized by a negative slow wave that is larger over the contralateral than ipsilateral hemisphere to the memorized visual field (Fukuda & Vogel, 2009; Ikkai, McCollough & Vogel, 2010; Luck & Vogel, 2013; Vogel & Machizawa, 2004; Vogel, McCollough & Machizawa, 2005; Woodman & Vogel, 2008). The amplitude of the CDA scales with the number of items held in memory (Drew & Vogel, 2008; Jost, Bryck, Vogel & Mayr, 2011; Luria & Vogel, 2011, 2014; McCollough, Machizawa & Vogel, 2007; Vogel & Machizawa, 2004; Vogel, McCollough & Machizawa, 2005; Woodman & Vogel, 2008) but not with the resolution or complexity of the items (Balaban & Luria, 2015; Ye, et al.,2014). Given that the CDA scales with the number of items held in VSTM, in the present study, we used the CDA to examine whether multiple colors (Experiment 1) or orientations (Experiment 2) can be consolidate in parallel or are consolidated in a serial manner.

We asked participants to remember one, two or four simultaneously presented items while severely limiting the consolidation time to the duration needed to consolidate a single item. We hypothesize that if the consolidation is a serial process, only a single item should be consolidated regardless of the set size of the display. Thus, there would be no difference among the CDA amplitudes for the different set sizes (Fig. 1a). However, if the items could be consolidated into VSTM in parallel, the CDA amplitude should increase with set size until the bandwidth of consolidation is exhausted (Fig. 1b). Furthermore, we also ran a condition in which consolidation time was not severely limited (300 ms exposure duration). This long exposure duration is used to measure the capacity of VSTM storage to ensure that any limits we find in the minimum time condition can be attributed to limits in the consolidation process rather than limits in storage capacity. In this condition, we would expect CDA amplitude to increase with set size until the total capacity is reached (Fig. 1c).

INSERT FIGURE 1 ABOUT HERE

Experiment 1

Methods

Participants

Twenty students (11 females) from Liaoning Normal University volunteered to participate in this experiment for paid compensation. They reported no history of neurological problems, reported having normal color vision and normal or corrected-to-normal visual acuity. Signed informed consent was provided by each participant prior to participation, and all procedures were in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

We based our sample size on a previous study that compared CDA produced by storing one item verses two items (Experiment 2 of Luria & Vogel, 2011). Based on the method proposed by Thalheimer and Cook (2002) for estimating effect sizes from F statistics, we estimated that their effect size was 1.39 for that comparison. To determine our sample size we assumed that our effect size would only be 50% of their effect size (.7), and calculated the sample size required to yield a power of .85 given $\alpha = .05$ for a paired-sample t-test (Faul, Erdfelder, Lang, & Buchner, 2007). This calculation yielded a sample size of 21.

Stimuli

In Experiment 1, the stimuli were presented with E-prime software on a 21-inch cathode ray tube (CRT) monitor (800 * 600 pixel, 144-Hz refresh rate). Participants were seated in an electrically shielded and sound-attenuated recording chamber at a viewing distance of 60 cm. Each memory item was a colored square with the color randomly chosen without replacement from a set of six highly discriminable colors (RGB values, red: 233,0,0; green: 30,138,18; blue: 26,49,178; orange: 210,85,7; yellow: 231,228,66; purple: 156,0,158). The mask consisted of a 6 × 6 multicolored checkerboard pattern composed of the same six colors as the stimuli; the color of each square in the mask was randomly assigned for each mask presentation. Each colored square and mask subtended 0.65° × 0.65° of visual angle.

Procedure: Main task

Participants performed a color identification task with the trial structures depicted in Fig. 2. Items were presented within 4° × 7.3° rectangular regions bilaterally, centered 3° to the left and right of the middle of the screen. The memory array consisted of 1, 2 or 4 different colored squares which were selected at random in each hemifield with the constraint that the given color could appear no more than once in each hemifield. Stimulus positions were randomized on each trial, with the constraint that the distance between squares within a hemifield was at least 2° (center to center).

Each trial started with the presentation of a fixation point ("+") in the middle of the screen. Then, two arrow-cues were presented for 200 ms above and below fixation, indicating the to-be-attended side on that trial. After a variable delay which ranged from 100 to 200 ms, the memory array was presented for 300 ms or the minimum time (the duration was determined by a threshold procedure described below). A mask was then presented for 100 ms, which was followed by a retention interval (when only the fixation cross was presented) of 900 ms and then the test array. The test array remained visible for 2500 ms or until a response was made. A 1000 ms period preceded the start of the next trial. Participants were required to keep their eyes fixated at the central cross while storing the colors in the hemifield indicated by the cue. The probe array in the cued hemifield was different from the corresponding color in the memory array in 50% of trials; they were identical in the remaining trials. When a colored square changed, a new colored square that was not used in the memory array would be randomly selected from the remaining colors. The task was to indicate whether the test array was identical to the memory array, with accuracy rather than response speed being stressed.

We varied set size at 3 levels (1, 2 and 4) at two exposure durations: 300 ms vs. minimum time, with all six conditions intermixed within blocks. All participants completed at least 12 trials of practice to ensure the participants understood the instructions and a total of ten blocks of 72 trials each, resulting in 120 trials per condition.

INSERT FIGURE 2 ABOUT HERE

Procedure: Thresholding exposure duration

For each participant, we determined the minimum exposure duration required to consolidate a single colored item that was presented alone in its hemifield. Prior to the main task, each participant ran four blocks (64 trials each) of this thresholding task. The method of constant stimuli was used with eight durations: 7 ms, 14 ms, 28 ms, 56 ms, 98 ms, 154 ms, 224 ms and 308 ms. In these blocks, a single color stimulus was presented and masked in a random location within each hemifield (within 4 * 7.3° rectangular regions bilaterally, centered 3° to the left and right of the middle of the screen) (see Fig. 3). Participants indicated whether the test item was identical to the memory item. 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 in the main task.

INSERT FIGURE 3 ABOUT HERE

Electroencephalography recording and analyses

Electroencephalographic (EEG) data was recorded with a QuickAmp amplifier (Brain Products GmbH, Munich, Germany). EEG was recorded from 64 tin electrodes mounted in an elastic cap, using the International 10/20 system. Vertical electrooculogram (VEOG) and horizontal electrooculogram (HEOG) were recorded with two electrodes, one placed below the left eye, and another placed next to the right eye. Impedance at each electrode site was maintained below 5 kΩ. The EEG and EOG were amplified using a 100 Hz low-pass and digitized at a sampling rate of 500 Hz. The EEG was algebraically re-referenced offline to the average of the left and right mastoids during post-recording analyses and segmented into 1000 ms epochs starting from 100 ms before the memory array onset. Trials with remaining artifacts exceeding ±75μV in amplitude were rejected. Participants with trial rejection rates that exceeded 25% were excluded from the analyses. No subject was excluded on this basis.

Two pairs of electrode sites at posterior parietal (P7/P8 and PO7/PO8) were chosen for analysis. The contralateral waveforms were computed by averaging the activity recorded at left hemisphere electrode sites when participants were cued to remember the memory array in the right hemifield, and vice versa. The CDA was usually measured by subtracting the ipsilateral activity from the contralateral activity, with a measurement window around 400-1000 ms after the onset of the memory array (Luria & Vogel, 2011; Peterson et al., 2015). In the present study, however, the memory array was followed by a mask array, which is necessary to limit the time for consolidation, but raised the potential problem that responses to the mask may overlap with the early phase of the CDA. For example, the masks could evoke a Pd component associated with termination of attention (Sawaki et al., 2012). Given that the CDA is postulated to reflect sustained neural activity throughout the entire delay period, we averaged the CDA in a time window of 600-1000 ms after memory array onset- a window that should be late enough to minimize the effect of the mask on the CDA. The average CDA waveforms were smoothed by applying a 10 Hz low-pass filter.

Results and Discussion

Behavioral results

The average minimum time across participants was 60 ms (range = 21-147 ms, SD = 32 ms). Percent correct (accuracy) for the minimum time conditions and 300 ms conditions were calculated for each participant. An ANOVA including the set size (1 vs 2 vs 4) and exposure duration (300 ms vs. minimum time) on accuracy yielded main effects of set size ($F_{2,38} = 283.99$, $p < .001$, $\eta_p^2 = .94$) and exposure duration ($F_{1,19} = 93.89$, $p < .001$, $\eta_p^2 = .83$). The interaction between the two factors was also significant. ($F_{2,38} = 5.71$, $p < .01$, $\eta_p^2 = .23$) (Fig. 4a). Post hoc analysis revealed that there were both significant difference in accuracy between 1 and 2 colors (1 color: M = .83, SD = .08; 2 colors: M = .73, SD = .09; $t_{19} = 8.01$, $p < .001$, Cohen's *d* = 1.15), and between 2 and 4 colors (2 colors: M = .73, SD = .09; 4 colors: M = .59, SD = .05; $t_{19} = 8.86$, $p < .001$, Cohen's *d* = 2.02) in minimum time condition. Similarly, those differences were also significant in the 300 ms condition (1 color: M = .95, SD = .02; 2 colors: M = .90, SD = .06; 4 colors: M = .69, SD = .07; 1 vs. 2 colors: $t_{19} = 4.20$, p

< .001, Cohen's *d* = 1.11; 2 vs. 4 colors: t_{19} = 14.88, p < .001, Cohen's *d* = 3.24). Such steady decline in behavioral accuracy at larger set sizes is expected as decision noise or interference also increases with set size (Eckstein et al., 2000; Palmer et al., 2000). Thus, behavioral accuracy in the change detection task is not diagnostic of whether the consolidation process is parallel or serial. Indeed, such considerations prompted previous work to compare sequential and simultaneous presentation conditions at the same set size (e.g., Huang & Pashler, 2007; Mance, Becker & Liu, 2012).

INSERT FIGURE 4 ABOUT HERE

Electrophysiological results

By leveraging the properties of the CDA, we were able to directly assess how many items were consolidated simultaneously without the need to compare to a sequential condition. The grand average subtraction waveforms of CDA in the minimum time condition and 300 ms condition for each set size are shown in Fig.4b and 4c, respectively. The CDA emerged after 400 ms and persisted throughout the retention period. We used a 600-1000 ms window (shaded area) to calculate CDA amplitude, due to potential mask-evoked ERPs (see Methods). The averaged CDA amplitudes are shown in Fig 4d. An ANOVA including set size (1 vs 2 vs 4) and exposure duration (minimum time vs. 300 ms) on mean amplitude yielded main effect of set size ($F_{2,38}$ = 17.05, p < .001, η_p^2 = .47), but no effect for duration conditions ($F_{1,19}$ = .66, p > .05). The interaction between the two factors was significant ($F_{2,38}$ = 5.12, p < .05, η_p^2 = .21).

For the minimum time condition, planned comparisons revealed that there was a significant difference in amplitude between 1 and 2 colors (1 color: M = -.88, SD = .94; 2 colors: M = -1.27, SD = 1.15; t_{19} = 2.61, p < .05, Cohen's *d* = .37), but no significant difference between 2 and 4 colors (2 colors: M = -1.27, SD = 1.15; 4 colors: M = -1.42, SD = 1.38; t_{19} = 1.11, p > .05). The finding that the CDA amplitude increased from set-size one to set-size two, suggests that two items can be consolidated into VSTM in parallel. However, the fact that there was no further increase in CDA between set-size two and four, suggests that there is a limit in the number of items that can be processed in parallel. These results are consistent with our prediction (see Fig. 1b) and suggest that the consolidation bandwidth is at least 2 items. These results are also consistent with previous behavioral studies (Mance, Becker & Liu, 2012; Miller, Becker & Liu, 2014), which suggested that consolidating color information proceeds in a parallel manner with the bandwidth limited to two items.

For the 300 ms condition, planned comparisons revealed that there were both significant difference in amplitude between 1 and 2 colors (1 color: M = -.55, SD = .91; 2 colors: M = -1.05, SD = 1.01; t_{19} = 3.02, p < .01, Cohen's *d* = .52), and between 2 and 4 colors (2 colors: M = -1.05, SD = 1.01; 4 colors: M = -1.74, SD = 1.12; $t_{19} = 3.26$, $p < .01$, Cohen's $d = .64$). The fact that the CDA amplitude increased between 2 and 4 colors when we prolonged the stimulus duration to 300 ms, thereby removing the constraint on consolidation, suggests that at least three colors can be stored within VSTM. This pattern is consistent with our prediction regarding CDA amplitude when consolidation time is relaxed (see Fig. 1c)

In sum, our minimum time results support the notion of parallel consolidation of two simultaneously presented colors, but the bandwidth of consolidation prohibits more than two items from being consolidated. The 300 ms condition provides evidence that this limit is due to consolidation rather than storage capacity limitations. In Experiment 2, we extend this method to the consolidation of oriented gratings in order to investigate whether the process of consolidating orientation information also occurs in parallel or occurs serially.

Experiment 2

Methods

Participants

Twenty-two students (12 females) from Liaoning Normal University volunteered to participate in this experiment for paid compensation. They reported no history of neurological problems, reported having normal orientation vision and normal or corrected-to-normal visual acuity. Signed informed consent was provided by each participant prior to participation, and all procedures were in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Stimuli

The main apparatus was similar to that used in Experiment 1 except visual stimuli were sinusoidal gratings (contrast: 0.7, spatial frequency: 3 cycle/deg) in a circular aperture presented on a gray background. The edge of the aperture was smoothed such that no sharp change in luminance was present between the grating and the background. The mask stimulus was a circular aperture containing pixel noise, with random luminance levels in a uniform distribution. Each orientation stimuli subtended 0.9° × 0.9° of visual angle and mask subtended 1° × 1° of visual angle.

Procedure: Main task

The procedure and main task were similar to that used in Experiment 1 with the trial structures depicted in Fig. 5. The memory array consisted of 1, 2 or 4 different gratings which could be in one of 12 orientations :10°, 24°, 38°, 52°, 66°, 80°, 100°, 114°, 128°, 142°, 156° and 170°. Stimulus positions were randomized on each trial, with the constraint that no two gratings within the same hemifield had the same orientation and that the distance between gratings within a hemifield was at least 2° (center to center). When a grating changed its orientation in the probe array, a new orientation with a 90° difference would be selected.

Similar to Experiment 1, we varied set size at 3 levels (1, 2 and 4) at two exposure durations: 300 ms vs. minimum time, with all six conditions intermixed within blocks. All participants completed at least 12 trials of practice to ensure the participants understood the instructions and a total of ten blocks of 72 trials each, resulting in 120 trials per condition.

INSERT FIGURE 5 ABOUT HERE

Procedure: Thresholding exposure duration

The procedure of thresholding exposure duration was the same as that used in previous experiment except the stimuli were gratings (see Fig. 6).

INSERT FIGURE 6 ABOUT HERE

Electroencephalography recording and analyses

The EEG recording and analysis procedures were identical to those used in Experiment 1. The time range for measuring the CDA was also 600–1000 ms after memory array onset, to avoid contamination of mask-evoked ERPs. Participants with trial rejection rates that exceeded 25% were excluded from the analyses; this led to the exclusion of two participants.

Results and Discussion

Behavioral results

The average minimum time across participants was 57 ms (range = 7-119 ms, SD = 33 ms). An ANOVA including set size (1 vs 2 vs 4) and exposure duration (300 ms vs. minimum time) on accuracy yielded main effects of set size ($F_{2,38}$ = 385.62, p < .001, η_p^2 = .95) and exposure duration ($F_{1,19}$ = 111.06, p < .001, η_p^2 = .85). The interaction between the two factors was also significant ($F_{2,38}$ = 18.61, p < .001, η_p^2 = .49) (Fig. 7a). Post hoc analysis revealed that there were both significant difference between 1 and 2 orientations (1 orientation: M = .84, SD = .08; 2 orientations: M = .69, SD = .08; t_{19} = 11.92, p < .001, Cohen's *d* = 1.88), and between 2 and 4 orientations (2 orientations: M = .69, SD = .08; 4 orientations: M = .56, SD = .04; t_{19} = 7.93, p < .001, Cohen's *d* = 2.01) in minimum time condition. Similarly, those differences were also significant in the 300 ms condition (1 orientation: M = .95, SD = .03; 2 orientations: M = .85, SD = .09; 4 orientations: M = .62, SD = .05; 1 vs. 2 orientations: $t_{19} = 6.13$, $p < .001$, Cohen's *d* = 1.50; 2 vs. 4 orientations: $t_{19} = 17.86$, $p < .001$, Cohen's *d* = 3.32). Similar to Experiment 1, behavioral accuracy also showed a steady decline as set size increased. Again, such decline is expected given associated increase in decision noise or interference when the set size increased and thus cannot reveal the nature of the performance limit.

INSERT FIGURE 7 ABOUT HERE

Electrophysiological results

The CDA emerged at 400 ms and persisted throughout the retention period. The grand average subtraction waveforms of CDA in the minimum time and 300 ms conditions for each set size are shown in Fig. 7b and 7c, respectively. An ANOVA including set size (1 vs 2 vs 4) and exposure duration (minimum time vs. 300 ms) on mean amplitude yielded main effect of set size ($F_{2,38} = 5.18$, $p = .01$, $\eta_p^2 = .21$), but no effect for exposure duration ($F_{1,19} = .057$, $p > .05$). The interaction between the two factors was also significant ($F_{2,38} = 5.08$, $p < .05$, $\eta_p^2 = .21$) (Fig. 7d).

For the minimum time condition, planned comparisons revealed that there were no significant difference in amplitude between 1 and 2 orientations (1 orientation: M = -.75, SD = .89; 2 orientations: M = -.70, SD = .82; $t_{19} = -.37$, $p > .05$) and between 2 and 4 orientations (2 orientations: M = -.70, SD = .82; 4 orientations: M = -.87, SD = .73; $t_{19} = 1.16$, $p > .05$). The fact that the CDA magnitude is constant across the three set size conditions, suggests that only one orientation was consolidated into VSTM in the minimum consolidation time regardless of the set size. This pattern is precisely the prediction of a strictly serial process (see Fig. 1a) and is consistent with previous behavioral studies (Becker, Miller, & Liu, 2013; Liu, & Becker, 2013; Miller, Becker, & Liu, 2014) suggesting serial consolidation of orientation information.

For the 300 ms condition, there was a significant difference in amplitude between 1 and 2 orientations (1 orientation: M = -.39, SD = .78; 2 orientations: M = -1.00, SD = .82; $t_{19} = 3.36$, $p < .01$, Cohen's *d* = .76), but no significant difference between 2 and 4 orientation (2 orientations: M = -1.00, SD = .82; 4 orientations: M = -1.00, SD = .72; $t_{19} = .03$, $p > .05$). This pattern of results suggests that the storage limit for oriented stimuli, while greater than one item, was limited to about two items. Importantly, these results show that when consolidation time was not severely limited people could consolidate more than one item in VSTM, suggesting that the one-item limit we found in the minimum time condition can be ascribed to consolidation rather than storage limitations.

In sum, the minimum time condition suggests that only a single orientation can be consolidated at a time, the hallmark of a strictly serial process. This result is therefore consistent with earlier work suggesting strictly serial consolidation of orientation information (Becker, Miller & Liu, 2013; Liu & Becker, 2013). Finally, the 300 ms condition provided a control which allowed us to demonstrate that once the limit on consolidation time was removed, people could consolidate and store more than one item in VSTM.

Comparison of CDA between Experiments 1 and 2

The CDA results within each experiment conform very well to our predictions, with color showing evidence for parallel consolidation and orientation showing evidence for serial consolidation. To further support the observed difference between features we explicitly compared across the color and orientation experiments. Here we focus on the minimum time condition, because it provided the critical results for testing the consolidation bandwidth. Using data for the minimum time condition from both experiments, we conducted a mixed-factorial ANOVA, with feature (color vs. orientation) as a between-subject factor and set size (1 vs. 2 vs. 4) as a within-subject factor. We found a main effect of set size ($F_{2,76} = 5.14$, $p < .01$, $\eta_p^2 = .12$), but non-significant effect of feature ($F_{1,38} = 2.02$, $p > .05$). The interaction between set size and feature was marginally significant ($F_{2,76} = 2.80$, $p = .067$, $\eta_p^2 = .07$). This marginal interaction is consistent with different CDA profiles for color and orientation established by our within-experiment analyses. While, the interaction did not reach conventional statistical threshold of .05, we note that this is a between-subject comparison, which likely increased variability thereby reducing statistical power. In addition, the analysis included the set size four data, which are not the most diagnostic for distinguishing between parallel and serial consolidation, and could have added additional variability that reduced the power to detect the interaction. To further assess the reliability of this effect, we ran a post-hoc mixed-factorial ANOVA that included only data from set sizes 1 and 2, as these two set-sizes should be the most informative in revealing a different CDA profile (see Fig. 1). This analysis revealed a marginally significant main effect for set size ($F_{1,38} = 3.21$, $p = .08$, $\eta_p^2 = .08$) and a non-significant effect for feature ($F_{1,38} = 1.50$, $p > .05$). Importantly, the interaction between set size and feature was significant ($F_{1,38} = 5.10$, $p < .05$, $\eta_p^2 = .12$). These results thus show that even in a between-subject analysis, there is credible statistical evidence that the CDA amplitude varied differently across set sizes for color and orientation, hence consistent with our within-experiment results.

General Discussion

In the current study, we measured the behavioral accuracy and CDA amplitudes in a masked change detection task to examine the bandwidth of VSTM consolidation. The CDA provides an electrophysiological index of VSTM consolidation that helps resolve the controversy about whether the bandwidth of consolidation depends on the type of information. Specifically, while our previous behavioral work has indicated that the consolidation of color information is a parallel process and the consolidation of orientation information is a serial process, work by others (Rideaux et al., 2015, 2016) has suggested that consolidation is always a parallel process.

The 300 ms exposure condition in both experiments provided a baseline measure of VSTM *storage capacity* under the current experimental settings. In Experiment 1 with color stimuli, the CDA amplitude increased monotonically as the set size increased from set-size one to two and to four. However, in Experiment 2 with orientation stimuli, the CDA amplitude increased from set size one to two but showed no further increase from set size two to four. Previous research has established that the CDA amplitude reflects the number of items held in the memory and it would increase with set size and reach an asymptotic level when the set size reaches the storage capacity (Drew & Vogel, 2008; Ikkai, McCollough & Vogel, 2010; Jost, Bryck, Vogel, & Mayr, 2011; Luck & Vogel, 2013; Luria & Vogel, 2011, 2014; McCollough, Machizawa & Vogel, 2007; Vogel & Machizawa, 2004; Vogel, McCollough & Machizawa, 2005; Woodman & Vogel, 2008). Given the property of CDA and our results in the 300 ms condition, we can infer that VSTM capacity for color is at least three items and its capacity for orientation is at least two items in our experiments. These measurements regarding the VSTM capacity are necessary to ensure the observed limit in consolidation is not due to limits in the storage capacity.

To assess the consolidation bandwidth, we presented the memory array for the minimum time needed to consolidate a single item and varied the set size. In Experiment 1, we found that the CDA amplitude was significantly larger for two colors than one color, but there was no significant difference between two colors and four colors. This suggests that two colors can be consolidated simultaneously but four colors cannot. In Experiment 2, we observed very different results such that the CDA amplitude for all set sizes was equivalent. This result suggests that at minimum time, only one orientation can be consolidated into VSTM. Importantly, these estimates of consolidation bandwidth are lower than the corresponding estimates of the storage capacity, allowing us to attribute the CDA-set size function to limits in consolidation.

The ERP results in the minimum time conditions thus dovetails nicely with previous studies using purely behavioral measures (Becker, Miller, & Liu, 2013; Liu, & Becker, 2013; Mance, Becker & Liu, 2012; Miller, Becker, & Liu, 2014). Both sets of results support the conclusion that consolidating color information is a parallel process with the bandwidth limited to two items, however, consolidating orientation information is strictly serial. Miller, Becker and Liu (2014) explained this difference by hypothesizing that encoding color requires less information than encoding orientation. For example, processing any area on a uniform color patch will suffice to encode its color, whereas encoding the orientation of a grating requires processing of an extended region. This different encoding requirement could lead to a larger bandwidth for color than orientation. The current CDA results further support this notion by providing an electrophysiological correlate of the consolidation bandwidth. However, the precise neural mechanisms for consolidation is unknown. We speculate the difference in consolidation bandwidth for color and orientation is due to differences in the ability to represent multiple features with neural population codes. Further research is necessary to elucidate the neural basis of consolidation bandwidth.

Given our hypothesis that color requires less consolidation resource than orientation, one might expect the a shorter minimum duration thresholds for color than orientation. However, we observed largely similar threshold durations in Experiments 1 and 2. We note that this is not necessarily unexpected, because consolidation could take the same finite amount of time regardless of how much of its bandwidth is consumed. If, however, the number of items exceeds the bandwidth, the consolidation process would need to be completed multiple times. The minimum threshold duration prevents the consolidation process from completing for multiple times, thereby allowing us to measure the number of items that can be accommodated by the bandwidth for different features. The key point is that thresholding procedure only limits the number of iterations that the consolidation process can complete, and therefore is orthogonal to the bandwidth of a single iteration.

Regarding prior results suggesting that orientation information can be consolidated in parallel (Rideaux et al., 2015, 2016), in the Introduction we discussed possible confounds and alternative explanations that may have been responsible for those results. The present CDA results cast further doubts on the notion that all features are consolidated in parallel as we found differences in CDA patterns between color and orientation: while the color pattern is consistent with parallel consolidation, the orientation pattern is consistent with serial consolidation. Thus the present results provide converging evidence for our original explanation, using a method that does not rely on a simultaneous/sequential comparison, but relies on a completely independent measure of the contents of VSTM – the CDA. Even so, it is worth noting that Rideaux et al., provide behavioral evidence that motion direction can be consolidated in parallel. Motion direction is a feature we have not examined, and it is quite possible that motion direction may be consolidated via a parallel mechanism. Thus, it would be interesting to examine consolidation limits of motion direction in future studies using electrophysiological measures. We conclude from the present findings that the bandwidth for consolidation is determined by the stimulus feature.

References

Balaban, H., & Luria, R. (2015). The number of objects determines visual working memory capacity allocation for complex items. *NeuroImage*, *119*, 54–62. https://doi.org/10.1016/j.neuroimage.2015.06.051

Becker, M. W., Miller, J. R., & Liu, T. (2013). A severe capacity limit in the consolidation of orientation information into visual short-term memory. *Atten Percept Psychophys*, *75*(3), 415–425. https://doi.org/10.3758/s13414-012-0410-0

Consineau, D. (2005). Confidence intervals in within-subject designs:A simpler solution to Loftus and

Masson´s method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45.

Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(16), 4183–4191. https://doi.org/10.1523/JNEUROSCI.0556-08.2008

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*(3), 425–451. https://doi.org/10.3758/BF03212096

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.

Fukuda, K., & Vogel, E. K. (2009). Human Variation in Overriding Attentional Capture. *Journal of Neuroscience*, *29*(27), 8726–8733. https://doi.org/10.1523/JNEUROSCI.2145-09.2009

Huang, L. (2010). What is the unit of visual attention? Object for selection, but Boolean map for access. *Journal of Experimental Psychology. General*, *139*(1), 162–179. https://doi.org/10.1037/a0018034

Huang, L. (2015). Color is processed less efficiently than orientation in change detection but more efficiently in visual search. *Psychological Science*, *26*(5), 646–652. https://doi.org/10.1177/0956797615569577

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, *114*(3), 599–631. https://doi.org/10.1037/0033-295X.114.3.599

Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human visual awareness. *Science (New York, N.Y.)*, *317*(5839), 823–5. https://doi.org/10.1126/science.1143515

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of Neurophysiology*, *103*(4), 1963–1968. https://doi.org/10.1152/jn.00978.2009

Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, *36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Jost, K., Bryck, R. L., Vogel, E. K., & Mayr, U. (2011). Are old adults just like low working memory young adults? Filtering efficiency and age differences in visual working memory. *Cerebral Cortex*, *21*(5), 1147–1154. https://doi.org/10.1093/cercor/bhq185

Liu, T., & Becker, M. W. (2013). Serial consolidation of orientation information into visual short-term memory. *Psychological Science*, *24*(6), 1044–1050. https://doi.org/10.1177/0956797612464381

Luck, S. J., & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

Luck, S. J., & Vogel, E. K. (2013). Visual Working Memory Capacity: From Psychophysics and Neurobiology to individual Differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639. https://doi.org/10.1016/j.neuropsychologia.2010.11.031

Luria, R., & Vogel, E. K. (2014). Come together, right now:Dynamic overwriting of an object's history through common fate. *Psychologist*, *26*(3), 194–198. https://doi.org/10.1162/jocn

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*(2), 429–438. https://doi.org/10.1037/a0023925

MoCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measure of maintaining representations in visual working memory. *Cortex*, *43*, 77–94. https://doi.org/10.1016/S0010-9452(08)70447-7

Miller, J. R., Becker, M. W., & Liu, T. (2014). The bandwidth of consolidation into visual short-term memory depends on the visual feature. *Visual Cognition*, *22*(7), 920–947. https://doi.org/10.1080/13506285.2014.936923

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10-12), 1227–1268. https://doi.org/10.1016/S0042-6989(99)00244-8

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10.3758/BF03210419

Peterson, D. J., Gözenman, F., Arciniega, H., & Berryhill, M. E. (2015). Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Attention, Perception & Psychophysics*, 2270–2283. https://doi.org/10.3758/s13414-015-0929-y

Rideaux, R., Apthorp, D., & Edwards, M. (2015). Evidence for parallel consolidation of motion direction and orientation into visual short-term memory. *Journal of Vision*, *15*(2), 17–17. https://doi.org/10.1167/15.2.17

Rideaux, R., & Edwards, M. (2016). The cost of parallel consolidation into visual working memory. *Journal of Vision*, *16(6)*(1), 1–14. https://doi.org/10.1167/16.6.1.doi

Sawaki, R., Geng, J. J., & Luck, S. J. (2012). A common neural mechanism for preventing and terminating the allocation of attention.*Journal of Neuroscience*, *32*(31), 10725-10736. https://doi:10.1523/JNEUROSCI.1864-12.2012.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*. https://doi.org/10.1017/CBO9781107415324.004

Stevanovski, B., & Jolicœur, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition*, *15*(5), 532–563. https://doi.org/10.1080/13506280600871917

Stevanovski, B., & Jolicœur, P. (2011). Consolidation of multifeature items in visual working memory: central capacity requirements for visual consolidation. *Attention, Perception & Psychophysics*, *73*(4), 1108–19. https://doi.org/10.3758/s13414-011-0099-5

Thalheimer, W., & Cook, S. (2002, August). How to calculate effect sizes from published research articles: A simplified methodology. Retrieved January 26, 2017 from http://work-learning.com/effect_sizes.htm.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(April), 748–751. https://doi.org/10.1038/nature02447

Vogel, E. K., MoCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. https://doi.org/Doi 10.1038/Nature04171

Vogel, E., Woodman, G., & Luck, S. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1436–1451. https://doi.org/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. https://doi.org/10.1167/10.2.14

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics*, *63*(8), 1293–1313. https://doi.org/10.3758/BF03194544

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin & Review*, *15*(1), 223–229. https://doi.org/10.3758/PBR.15.1.223

Ye, C., Zhang, L., Liu, T., Li, H., & Liu, Q. (2014). Visual working memory capacity for color is independent of representation resolution. *PLoS ONE*, *9*(3), 1–6. https://doi.org/10.1371/journal.pone.0091681

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–U13. https://doi.org/Doi 10.1038/Nature06860

Figure captions

Figure 1 *The left two panels are the predicted patterns for the CDA amplitude for the minimum time if consolidation is strictly serial (a) or parallel for two items (b). The diagnostic comparison is whether adding an additional item to the display results in no increase in CDA amplitude (a) or an increase in CDA amplitude (b). The former would indicate only a single item is consolidated, whereas the latter would indicate two items are consolidated. The right panel (c) is the predicted pattern for the 300ms presentation duration where processing time is not limited, thus allowing the consolidation procedure to complete for multiple times. Here the number of items in memory should increase (as indicated by increasing CDA amplitude) until the storage capacity (K) is reached, at which point the pattern should plateau. A second critical diagnostic comparison is that the plateau in panel b should occur earlier than the* *plateau in panel c. If not, the plateau in panel b could be attributed to limits in the storage capacity rather than limits in the consolidation bandwidth.*

Figure 2 *Experimental procedure used in Main task in Experiment 1.*

Figure 3 *Experimental procedure used in thresholding task in Experiment 1.*

Figure 4 *The results in Experiment 1. (Error bars are estimates of within-subject standard errors following the method of Cousineau (2005).) The behavioral results of minimum time condition and 300 ms condition(a). Difference waves of CDA for arrays of 1, 2 and 4colors of minimum time (b) and 300 ms condition(c) for an averaged two pairs of electrode sites. The mean amplitudes of CDA for minimum time and 300 ms conditions (d).*

Figure 5 *Experimental procedure used in Main task in Experiment 2.*

Figure 6 *Experimental procedure used in thresholding task in Experiment 2.*

Figure 7 *The results in Experiment 2. (Error bars are estimates of within-subject standard errors following the method of Cousineau (2005).) The behavioral results of minimum time condition and 300 ms condition(a). Difference waves of CDA for arrays of 1, 2 and 4colors of minimum time (b) and 300 ms condition(c) for an averaged two pairs of electrode sites. The mean amplitudes of CDA for minimum time and 300 ms conditions (d).*

Fig1

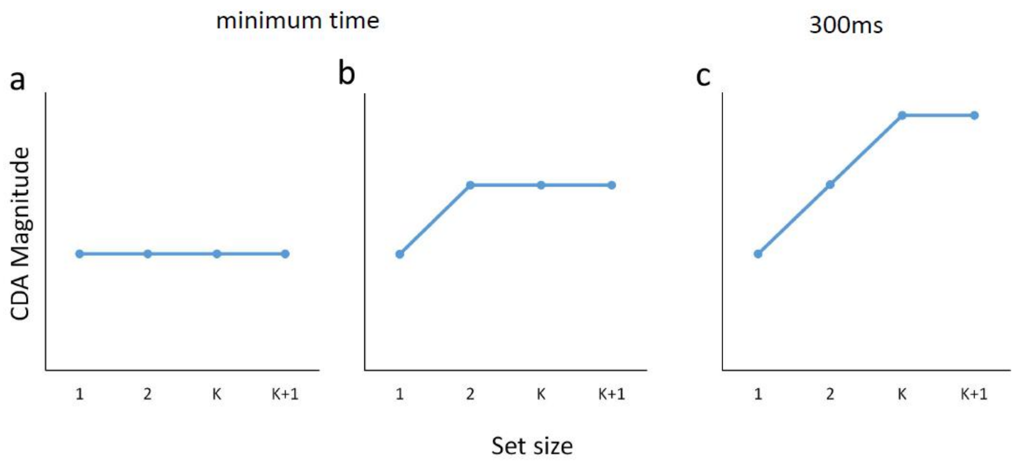


Fig2

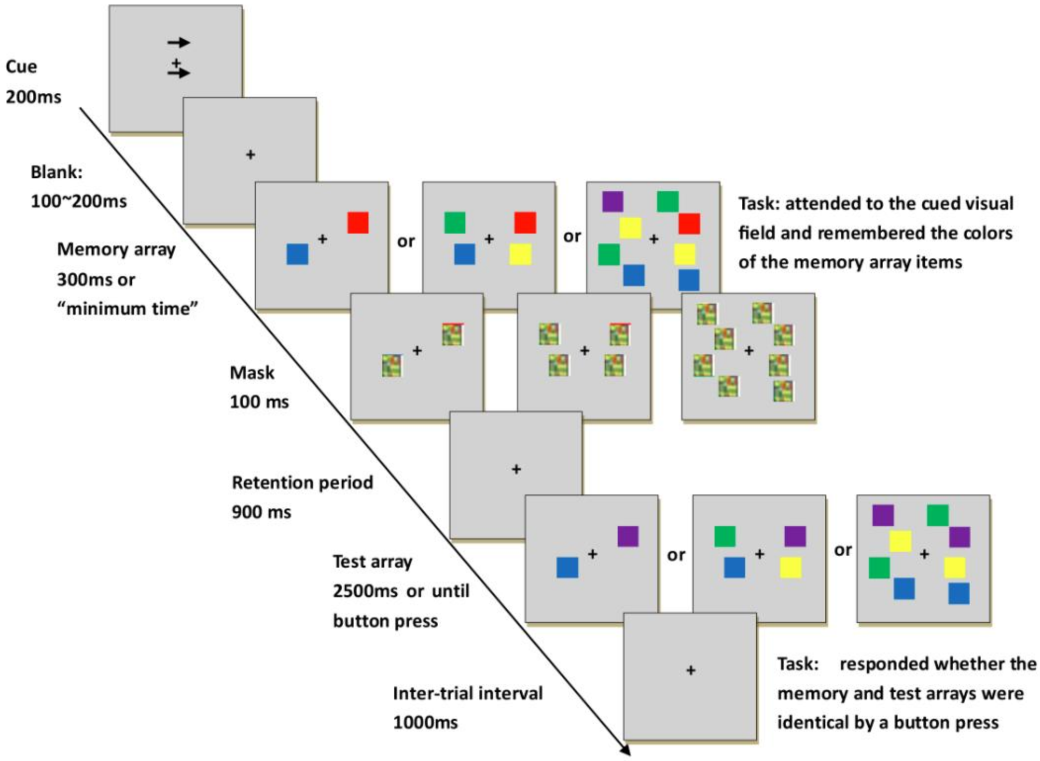

Cue
200ms
Blank:
100~200ms
Memory array
300ms or
"minimum time"
or
or
Task: attended to the cued visual
field and remembered the colors
of the memory array items
Mask
100 ms
Retention period
900 ms
Test array
2500ms or until
button press
or
or
Inter-trial interval
1000ms
Task: responded whether the
memory and test arrays were
identical by a button press

Fig3

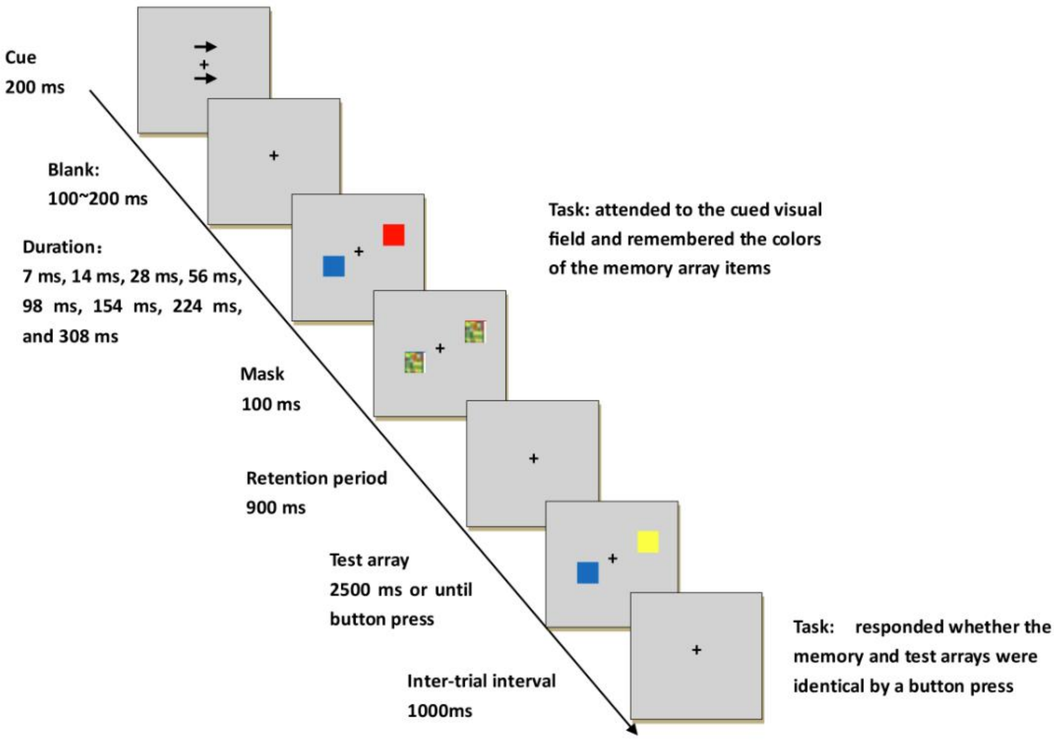

Cue
200 ms
Blank:
100~200 ms
Duration:
7 ms, 14 ms, 28 ms, 56 ms,
98 ms, 154 ms, 224 ms,
and 308 ms
Task: attended to the cued visual
field and remembered the colors
of the memory array items
Mask
100 ms
Retention period
900 ms
Test array
2500 ms or until
button press
Inter-trial interval
1000ms
Task: responded whether the
memory and test arrays were
identical by a button press

Fig4

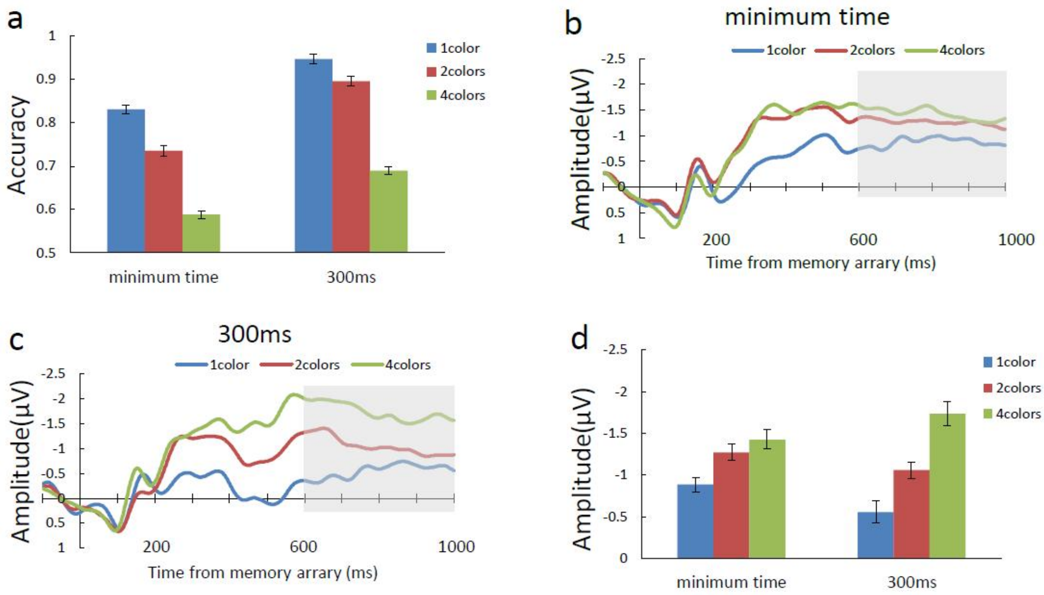

a
Accuracy
1
0.9
0.8
0.7
0.6
0.5
1color
2colors
4colors
minimum time
300ms
b
minimum time
1color
2colors
4colors
Amplitude(μV)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
200
600
1000
Time from memory array (ms)
c
300ms
1color
2colors
4colors
Amplitude(μV)
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
200
600
1000
Time from memory array (ms)
d
Amplitude(μV)
-2.5
-2
-1.5
-1
-0.5
0
1color
2colors
4colors
minimum time
300ms

Fig5

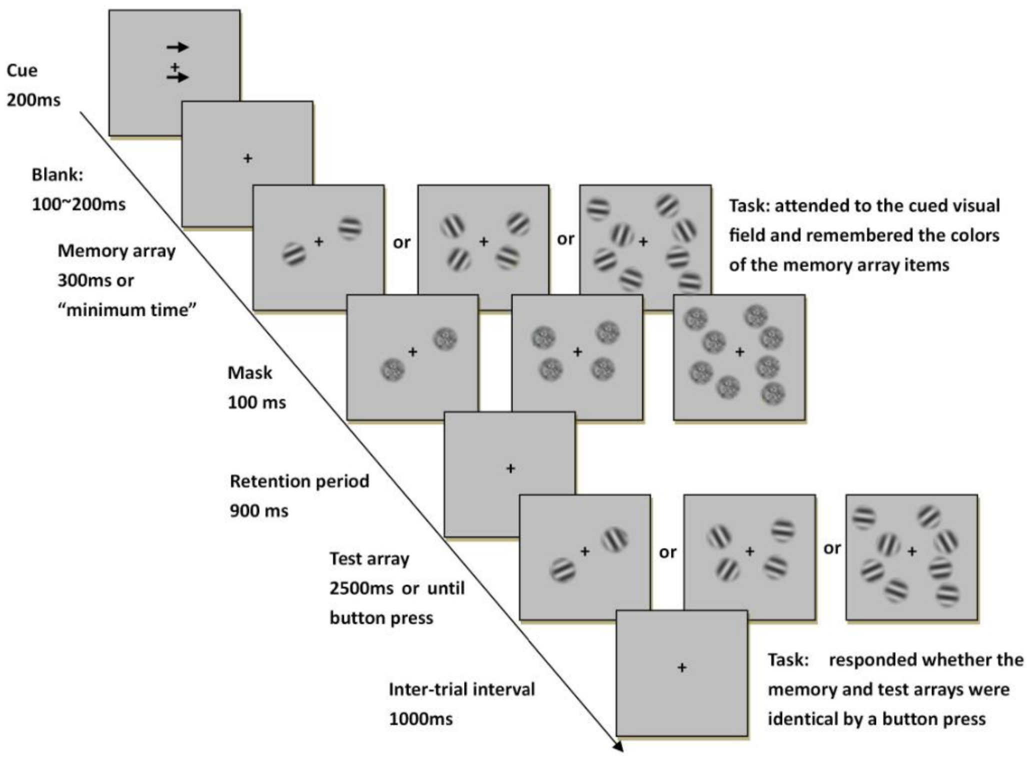
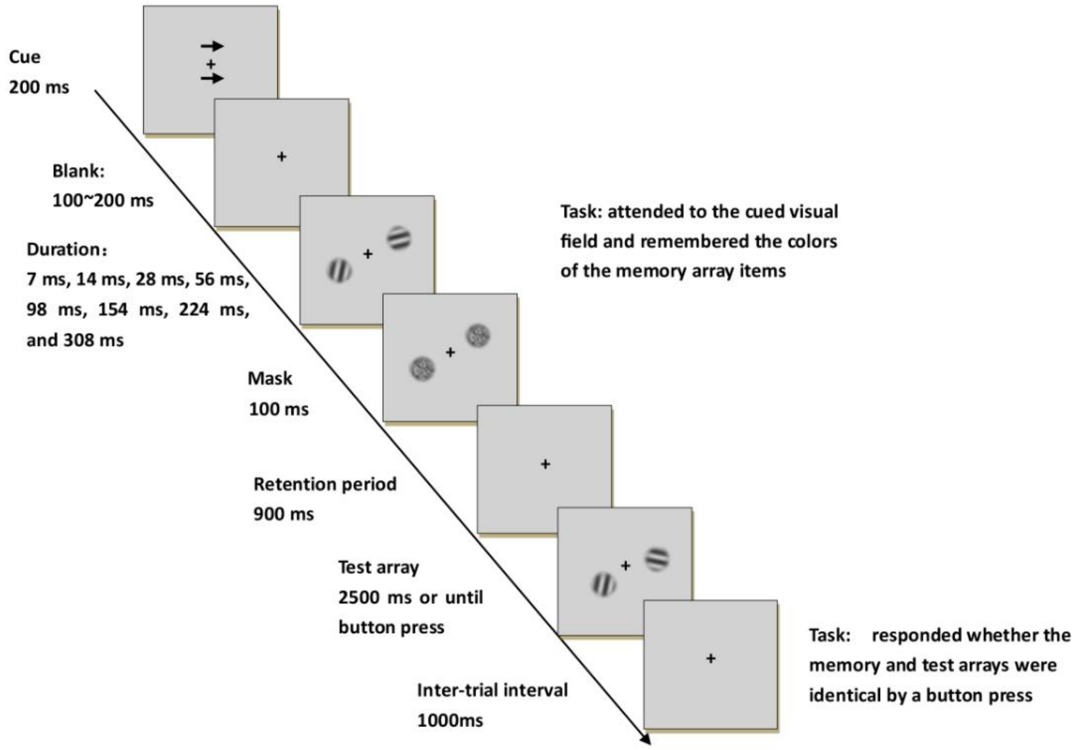
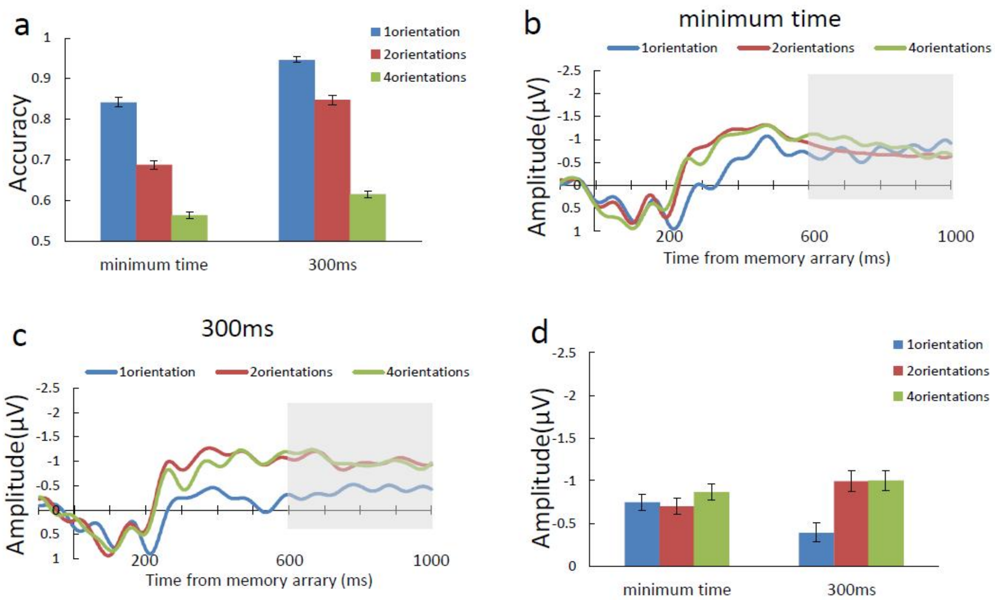

Cue
200ms
Blank:
100~200ms
Memory array
300ms or
"minimum time"
or
or
Task: attended to the cued visual
field and remembered the colors
of the memory array items
Mask
100 ms
Retention period
900 ms
Test array
2500ms or until
button press
or
or
Inter-trial interval
1000ms
Task: responded whether the
memory and test arrays were
identical by a button press

Fig6



Cue
200 ms
Blank:
100~200 ms
Duration:
7 ms, 14 ms, 28 ms, 56 ms,
98 ms, 154 ms, 224 ms,
and 308 ms
Mask
100 ms
Retention period
900 ms
Test array
2500 ms or until
button press
Inter-trial interval
1000ms
Task: attended to the cued visual
field and remembered the colors
of the memory array items
Task: responded whether the
memory and test arrays were
identical by a button press

Fig7



a
Accuracy
1orientation
2orientations
4orientations
minimum time
300ms
b
minimum time
Amplitude(μV)
Time from memory arrary (ms)
c
300ms
Amplitude(μV)
Time from memory arrary (ms)
d
Amplitude(μV)
minimum time
300ms