Contralateral delay activity tracks information load in visual working memory: Evidence from the multiple object tracking task

Xiaoshu Lin[1,2+], Chaoxiong Ye[1+], Tiina Parviainen[1], Qiang Liu[2,3]*

[1] Department of Psychology, University of Jyvaskyla, Jyvaskyla, Finland.
[2] Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, China.
[3] Center of Brain and Cognitive Neuroscience, Liaoning Normal University, Dalian, China.

Author Note

This work is funded by the National Natural Science Foundation of China (grant number 31970989 to Qiang Liu), and the Academy of Finland (grant number 333649 to Chaoxiong Ye). All authors declared that there are no any financial and personal relationships with other people or organizations that could inappropriately influence (bias) this work. The authors declare no competing financial interests.

+ Xiaoshu Lin and Chaoxiong Ye contributed equally to this work and should be considered as co-first authors.

*Correspondence should be addressed to Qiang Liu, Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, China. email: lq780614@163.com

Abstract

Contralateral delay activity (CDA) is widely used to index the number of representations stored in visual working memory (VWM). However, it was alternatively proposed that CDA reflects the current focus of spatial attention, as evidenced by the CDA observed in a multiple-object tracking (MOT) task. This study examined whether the CDA observed in the MOT task was derived from representational storage, which might have come from the feature (cue) masking of the target items in the tracking stage. Therefore, masks were used to manipulate whether participants needed to remember the tracking items during the tracking stage. Experiment 1 used a VWM-MOT dual task to show that the demand for representational storage existed only in the masked condition. Experiment 2 showed that the CDA existed in both the masked and unmasked conditions, but the amplitude of CDA was only sensitive to the tracking load in the masked condition. These results suggest that the CDA could be related to both sustained attentional activation and representational storage. Importantly, CDA showed the effect of the tracking load only when the process of representational storage was recruited during tracking. Collectively, this work provides strong evidence that CDA indexes only VWM storage instead of the current focus of spatial attention.

Keywords

Contralateral delay activity, multiple-object tracking task, visual working memory, attention

1. Introduction

Visual working memory (VWM) is generally regarded as a workspace that can store a certain amount of information online in the service of advanced cognitive functions (Luck & Vogel, 2013). As a subsystem of working memory (WM), VWM underlies significant cognitive operations, such as the visual comparison of objects over space and time (Pilling & Barrett, 2016), reading comprehension, reasoning, intelligence, and learning abilities, which affect educational outcomes (e.g., Bull et al., 2008; Conway et al., 2003; Daneman. & Carpenter., 1980; Süß et al., 2012; Fukuda et al., 2010). In recent years, researchers have devoted much effort into exploring the neural mechanisms of VWM by adopting electrophysiological and neuroimaging techniques.

In electroencephalography (EEG) studies of WM, contralateral delay activity (CDA), a negative contralateral wave over posterior electrode sites during the delay period, has been widely used to index memory storage. CDA amplitude has been widely reported to increase with increases in the number of items stored in VWM and reaches an asymptote at about 3–4 items (Cowan, 2001; Luck & Vogel, 2013). A recent meta-analysis summarized 11 studies (Luria et al., 2016) and concluded that CDA amplitude could track the changes in mnemonic load and was highly predictive of individual VWM capacity (Kundu et al., 2013; Kang & Woodman, 2014; Balaban & Luria, 2017). Previous studies have proposed that CDA indexes only VWM storage instead of the current focus of spatial attention.

The storage account of CDA was mirrored by previous findings that the mnemonic load affected the CDA signal only in the WM task rather than in the attention task, which did not involve representational storage (Feldmann-Wustefeld et al., 2018; Wang et al., 2019; Hakim et al., 2019). Overall, these findings seem to support the conclusion that the CDA is derived from the storage of mnemonic representations in VWM.

However, researchers have also found evidence to support that the CDA indexes the current focus of spatial attention. For example, CDA amplitude increased according to the actual number of tracked items in the multiple object tracking (MOT) task. During the MOT task, participants are required to follow moving objects constantly. The classic paradigm of the MOT task contains 6 to 8 items, including 1 to 3 target items that share the same feature (cue), and the rest are nontargets. At first, all items are statically displayed for 500 ms, and then the feature (cue) of the target item disappears, leading to no distinction between target items and nontargets, but the position of all items continues to be visible. Next, in the tracking stage, all items move erratically for a chosen time. After stopping, one of the items is cued. The participants are then asked to determine whether the item was the target item. Prior research found that the posterior electrode was able to detect the CDA , which was more negative on the opposite side of the tracking item, and its amplitude increased with an increase in the number of tracking items (Drew & Vogel, 2008). Moreover, the CDA amplitude was an excellent predictor of tracking performance and was able to capture online changes in the number of tracked objects (Drew et al., 2011, 2012). Thus, researchers have proposed an attentional activation account to argue that CDA indexes the current focus of spatial attention at the locations that were encoded into VWM.

However, because the target items are kept visible throughout the MOT task, the participants may only need to pay sustained attention to the targets during tracking without necessarily involving the process of representational storage. Therefore, the CDA observed in the MOT task possibly reflects the focus of spatial attention (Cavanagh & Alvarez, 2005; Storm & Pylyshyn, 1988; Luria et al., 2010; Berggren & Eimer, 2016), thereby supporting the attentional activation account.

Some researchers believe that CDA observed in the MOT task still supports the storage account (Ikkai et al., 2010; Feldmann-Wustefeld et al., 2018). Although the target items are constantly visible throughout the trial, the feature (cue) of the target items is masked in the tracking stage, thereby creating difficulties in individuating the targets from visually identical distractors. To precisely track the targets, participants might attempt to store the target representations by linking the identity information with spatial positions to update the current positions during the tracking period (Tsubomi et al., 2013).

Therefore, the key to distinguishing these two accounts lies in whether the classical MOT task involves VWM, and if so, whether CDA can still be observed and index the number of tracking loads when there is no VWM load in the MOT task. At present, there is no empirical evidence that can rule out one of the hypotheses.

Therefore, the current study sought to test these accounts by attempting to exclude VWM load in the MOT task. Our hypothesis is that WM s exist in the tracking stage of the MOT task and that the storage requirement for representation comes from the masking of cues. The primary strategy was to manipulate the feature (cue) masking of target items in the tracking stage. We designed an MOT task with two conditions that may contain different cognitive progress: the masked condition (invisible cue in the tracking stage) with VWM processing and the unmasked condition (visible cue in the tracking stage) without VWM processing. In Experiment 1, by observing whether an added storage requirement affected the tracking accuracy, we verified that the masked condition involved WM and that the unmasked condition was merely an attention task. In Experiment 2, we found that the CDA amplitude was sensitive to the number of tracked targets only in the masked condition.

Collectively, these results provide evidence that the MOT task involves WM. Although the CDA is directly linked to sustained attention and representational storage of tracked targets, the CDA signal showed an effect on the tracking load only when the process of representational storage was recruited during the tracking period.

2. Experiment 1

2.1 Overview

In Experiment 1, our aim was to verify whether the WM was recruited under both conditions; therefore, we added a color storage memory requirement in both the masked and unmasked conditions. If the unmasked condition task is an attentional task without a storage requirement, the added color storage requirement should have no significant effect on tracking accuracy. Similarly, if the masking condition task has a storage requirement, then the additional color storage requirements will increase the total storage load and should have a significant impact on tracking accuracy.

The task was also adjusted to avoid some possible influencing factors. All adjustments were equally settled in the unmasked condition and masked condition. First, because the cue is constantly visible in the unmasked condition, our design might have elicited a salience effect in the unmasked condition. To avoid this salience effect and the accompanying automatic bottom-up processing, we selected a bar rather than a color as the feature (cue) for the target items. Participants were instructed to track targets with vertical bars among the distractors with horizontal bars, and vice versa. Due to the low distinction between the horizontal and vertical bars, the modified MOT task potentially required participants to allocate sustained attention to the tracked targets, although the feature (cue) of the targets was not masked.

Second, The low distinction and short cue array also made the target items difficult to identify at first, leading to a failure in tracking. We solved this problem by lengthening the cue array and setting an earlier presence for nontargets than for targets.

2.2 Method

2.2.1 Experimental design

Experiment 1 was designed as a 2 (masked and unmasked condition) × 2 (Tracks 1 and 2) × 2 (1 and 2 color squares) within-subjects study. Our goal was to observe how color storage requirements affect tracking accuracy, so we set the color store to one or two to serve as the variable. In addition, we reduced the maximum track load to two to ensure that the memory load in the most difficult conditions was four, as this would avoid exceeding the general memory capacity.

2.2.2 Participants

Twenty-four (4 males; 18.88 ± 0.20 years old) students from Sichuan Normal University with self-reported normal or corrected vision and normal color vision participated in the experiment. All participants provided informed consent and were paid CNY30 for the one-and-a-half-hour experiment. Six subjects were replaced because their accuracy in tracking during the judgment trials or their accuracy in detecting color changes in trials was less than 50% (guess rate) for a given condition.

2.2.3 Stimuli and procedures

Experiment 1 used PsychoPy 3 software to present the stimuli and collect the data. As a MOT-VWM dual task, there were tracking and memory stimuli; all stimuli were presented on a 24-inch LCD monitor (60 Hz refresh rate, 1920 × 1080 pixels).

Participants were seated at a viewing distance of 63 cm in a chamber. Before the start of the experiments, they were asked to stare at the fixation cross (0.6°) and concentrate on the task throughout the trial.

Six squares (0.6° × 0.6°) served as tracking stimuli. They were presented in each hemifield and moved within two invisible rectangular regions (3.25° × 5.9°), with the inner edge of the rectangle laterally offset from the fixation cross by 0.8° on a gray screen. These stimuli moved linearly and erratically at a speed of 0.5°–1°/s during the tracking period, with possible overlaps with each other. A horizontal or vertical bar placed inside the hollow square was selected as the cue to indicate the target. The cue type was randomly varied with a 50% probability for each type.

One or two color squares served as memory stimuli. The memory square (0.6° × 0.6°) was randomly designed to be one of 10 colors (pink, red, fuchsia, purple, blue, cyan, green, lime, yellow, and orange). When the number of memory color squares was one, it randomly appeared 1° above or below the fixation cross. When the number of memory color squares was two, the squares were fixed to display at 1° above and below the fixation cross, respectively.

As shown in Figure 2, each trial started with an arrow and a hollow square with a horizontal or vertical bar for 300 ms. The arrow instructed participants to pay attention to the left or right side, and the square indicating the horizontal or vertical bar would be targets. If the cue was a square with horizontal bar, the targets in the cue array were squares with horizontal bars, and vice versa. Then, a set of six stationary squares that served as targets and nontargets were presented on each side for 1000 ms. There were five nontargets and one target in the Track 1 condition and four nontargets and two targets in the Track 2 condition. During that time, the nontargets were constantly kept visible, whereas the targets appeared only in the later 500 ms. As for the color squares, they appeared only in the first 500 ms with the nontargets, and the participants were guided to memorize the colors of those squares but not the nontargets.

Subsequently, the target and nontarget squares moved independently for 1500 ms. We manipulated whether participants could see the target feature during the tracking period. In the unmasked condition, the bars in the targets and nontargets were kept visible, whereas the bars (target feature) disappeared in the masked condition. Following the tracking stage, the squares stopped moving. One square in each hemifield then depicted a slash inside it, while the others were presented as hollow squares in the stationary test array. In 80% of trials in Experiment 1, the participants were instructed to press “F” if they identified the item with a slash on the attended side as the tracked target; otherwise, they were instructed to press “J.” The item on the attended side was identified as a tracked target or an untracked distractor with equal probability. In these arrays, similar stimuli were presented on both sides to equalize the perceptual inputs in each hemifield.

In addition, in the tracking stage, the probability that the color squares would be shown at the same place as they had in the cue array was 20%. Once the color squares were shown, the subjects were required to judge whether the color had changed but not to determine whether the slashed item was the tracked target. If the color squares were not shown, then the participants were required to judge the tracking items, as described above. The color squares appeared at random in the tracking array at minimum intervals of 500 ms. If the color was consistent with the color shown in the cue array, the participants were instructed to press “F;” otherwise, they were told to press “J.”

Each participant was required to perform all eight conditions (masked condition: Track 1 and memory 1 color, Track 1 and memory 2 color, Track 2 and memory 1 color, and Track 2 and memory 2 color; unmasked condition: Track 1 and memory 1 color, Track 1 and memory 2 color, Track 2 and memory 1 color, and Track 2 and memory 2 color). Each condition had 60 trials, 80% of which required a judgment of the moving targets, and 20% of which required a judgment of the color squares. The order of presentation of all variables was counterbalanced across all participants. Ultimately, after ensuring that the accuracy was higher than the guess rate in all conditions, only the judgment of the tracking items was used for the analysis.

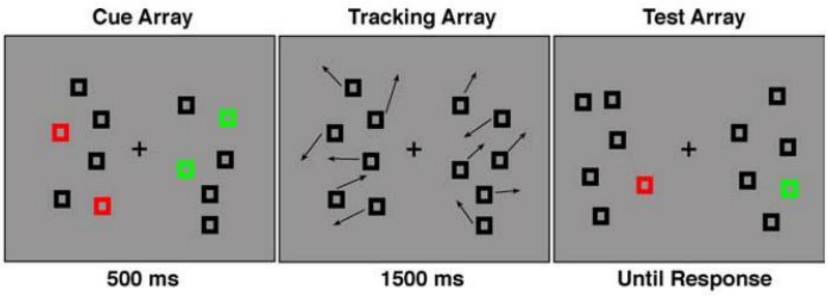


Figure 1. Multiple-object tracking paradigm (a) adopted by Drew and Vogel (2008).

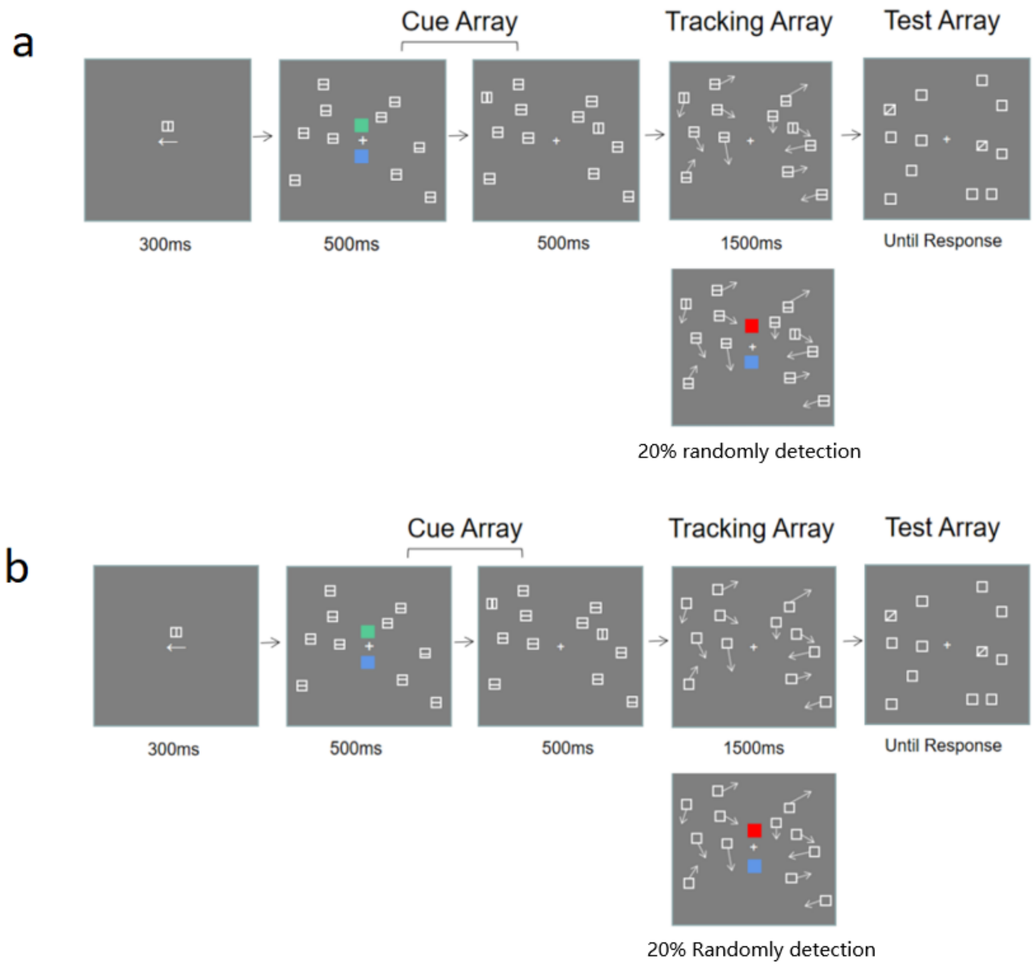


Figure 2. Stimulus and behavioral procedure in Experiment 1 in the (a) unmasked condition and (b) masked condition. Trial events are depicted from left to right. The procedure for the two conditions was the same, except for masking in tracking array.

2.3 Results

We used a 2 (masked and unmasked condition) × 2 (Tracks 1 and 2) × 2 (1 and 2 color squares) repeated measures analyses of variance (ANOVAs) to analyze the accuracy among all conditions. Then, paired t-tests and simple effect analyses were conducted for the follow-up pairwise comparison. The significance level was set as $p < 0.05$. Cohen's d was used as an estimator of the effect size for the t-tests.

As shown in Figure 3, the ANOVA results revealed a significant main effect of condition ($F [1, 23] = 11.474$, $p = 0.003$, $\eta^2_p = 0.333$), tracking load ($F [1, 23] = 14.875$, $p = 0.001$, $\eta^2_p = 0.393$), and color storage load ($F [1, 23] = 5.428$, $p = 0.029$, $\eta^2_p = 0.191$). No significant interactions were detected between condition and tracking load ($F [1, 23] = 3.770$, $p = 0.065$, $\eta^2_p = 0.141$), condition and color storage load ($F [1, 23] = 2.272$, $p = 0.145$, $\eta^2_p = 0.090$), or color storage load and tracking load ($F [1, 23] = 2.188$, $p = 0.153$, $\eta^2_p = 0.087$); however, a significant interaction was observed among condition, tracking load, and color storage load ($F [1, 23] = 4.473$, $p = 0.045$, $\eta^2_p = 0.163$). The paired t-test showed that the accuracy was better in the unmasked than masked condition ($t [95] = 4.836$; $p < 0.001$, Cohen's $d = 0.992$), and the accuracy decreased as the number of tracked targets ($t [95] = 3.702$; $p < 0.001$, Cohen's $d = 0.76$) and the number of color squares ($t [95] = 2.562$; $p < 0.001$, Cohen's $d = 0.526$) increased, merging the unmasked and masked conditions.

Subsequently, in the unmasked condition, the results of simple effect analysis showed a significant main effect of tracking load ($F [1, 23] = 4.874$, $p = 0.038$, $\eta^2_p = 0.175$) but no significant main effect of color storage load ($F [1, 23] = 0.820$, $p = 0.374$, $\eta^2_p = 0.034$) and no significant interaction between tracking load and color storage load ($F [1, 23] = 0.014$, $p = 0.907$, $\eta^2_p = 0.001$). However, in the masked condition, a significant main effect was observed for tracking load ($F [1, 23] = 1.0430$, $p = 0.001$, $\eta^2_p = 0.369$) and for color storage load ($F [1, 23] = 5.424$, $p = 0.029$, $\eta^2_p = 0.191$), but no significant interaction was detected between tracking load and color storage load ($F [1, 23] = 3.983$, $p = 0.058$, $\eta^2_p = 0.148$). These results showed that tracking accuracy was affected both by the number of tracking items and by color storage in the masked condition but not in the unmasked condition.

Collectively, the change in the number of color squares affected only the masked condition. Considering that the change in the number of color items meant a change in the storage load and that the storage load did not affect the accuracy of the attention task and only affected the accuracy of the WM task with a storage load, this indicates that the WM existed in the original MOT task and that our manipulation of the feature (cue) masking of the target items successfully decoupled the storage of the representations.

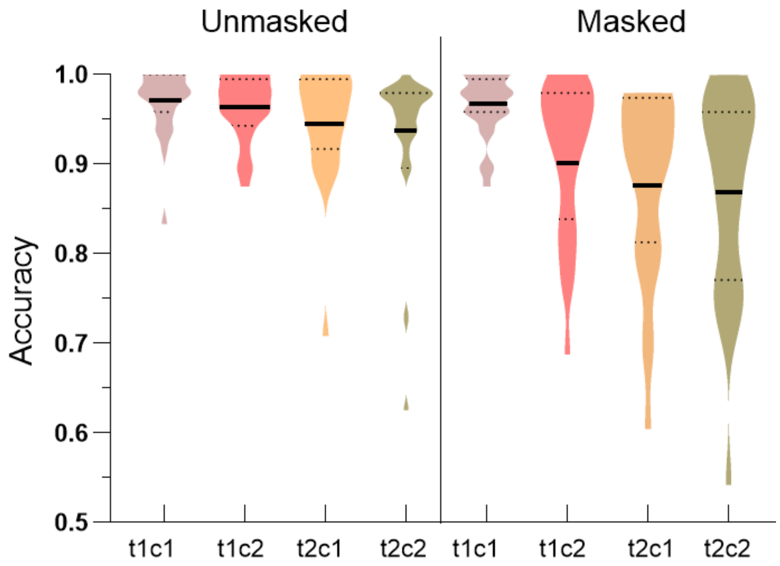


Figure 3. Accuracy for the four combinations of tracking load and number of color squares under the unmasked and masked conditions (t1c1: tracking 1 target and memory 1 color, t1c2: tracking 1 target and memory 2 colors, t2c1: tracking 2 targets and memory 1 color, and t2c2: tracking 2 targets and memory 2 colors). The black line represents the average performance, and the dotted line represents the quartile.

2.4. Discussion

In Experiment 1, we attempted to use a dual task to investigate whether the storage representation had different effects on the two conditions. We found that the color memory representation in the unmasked condition did not affect the tracking accuracy. In contrast, the same representation affected tracking accuracy in the masked condition. This result confirms our hypothesis: The original MOT task (with masks) recruits WM representations to maintain continuously refreshing targets. That is, the masked condition recruits WM, whereas the unmasked condition does not.

3. Experiment 2

3.1 Overview

In Experiment 1, we found different patterns of recruitment of representational storage under two conditions. Because our manipulation of the masking cue was set in the tracking stage and the defining feature of CDA is to show the sensitivity of the changes of load in the delayed stage, which is consistent with the tracking stage in the MOT task, we present the following hypothesis: if the CDA indexes representational storage, then tracking load would affect the CDA when the feature of the targets is masked but not unmasked. Conversely, the attention activation account of CDA predicts that the tracking load effect would be observed, regardless of the masking of the target feature. Therefore, in Experiment 2, we determined whether the CDA is sensitive to the number of changes in tracking items under the masked and unmasked conditions.

3.2 Method

3.2.1 Experimental design

Experiment 2 was designed as a 2 (masked and unmasked condition) $\times$3 (Tracks 1, 2, and 3) within-subjects study. If masking of the target feature recruits memory representations, tracking load would have an effect on the CDA in the masked but not unmasked condition. In contrast, CDA predicts that the tracking load effect would be observed regardless of whether the feature of targets was masked.

3.2.2 Participants

Thirty-seven participants (12 males; 21.08 $\pm$ 0.32 years old) were recruited from Sichuan Normal University. They self-reported normal color vision and normal or corrected-to-normal visual acuity and provided informed consent before participation. Each participant was compensated CNY 100 for their participation.

Some data were excluded due to technical errors and excessive EEG artifacts, so 24 participants (9 males; 20.88 ± 0.39 years old) were ultimately included in the final statistical analysis. This number was sufficient to detect reliable variations in CDA amplitude according to previous MOT tasks on the CDA (Drew & Vogel, 2008; Drew et al., 2011).

3.2.3 Stimuli and procedures

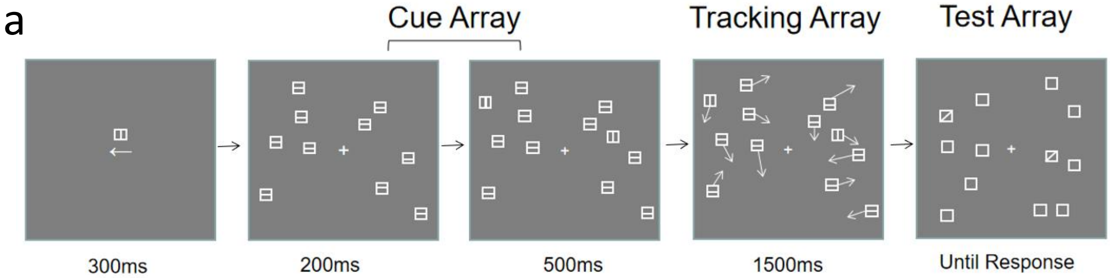
Experiment 2 also used PsychoPy 3 software to present stimuli and collect data, and it was modified based on the program used in Experiment 1. The changes were as follows: 1) The variable of the color squares was eliminated to reduce the impact of unnecessary variables on CDA; therefore, participants were asked to determine the tracking items in all trials; 2) Since the participants now did not need to remember color blocks, the time at which the cue array was presented was reduced to 700 ms. During that time, the nontargets were constantly kept visible, but the targets appeared only in the later 500 ms, which was sufficient to mark the targets; and 3) Because the tracking load served as a variable to observe the sensitivity of the CDA to the number of memory representations, we set the tracking load at 1, 2, or 3 to better observe the effects of tracking load on CDA.

Each participant was required to perform trials in both the masked and unmasked conditions in two separate sessions in the same periods of time on different days to provide enough trials for CDA analysis. They needed to complete 30 blocks of 24 trials each per condition for a total of 1440 trials. Each condition had three tracking loads (Tracks 1, 2, and 3), which were separated by self-paced breaks. Prior to the initiation of the trials, the participants were informed of the tracking load and then completed 10 blocks for each tracking load. The order of the condition and tracking loads was counterbalanced across participants.

3.2.4 EEG recording and analysis

Event-related potentials (ERPs) were recorded in the experiment using our standard recording and analysis procedures. Data were recorded using 64 tin electrodes mounted in an elastic cap (Electrocap International) using the International 10/20 System. All sites were recorded with an FCz reference, and the data were re-referenced offline to the algebraic average of the left and right mastoids. Data from a horizontal electrooculogram (EOG) was recorded from electrodes placed approximately 1 cm from the external canthi of the right eye to measure horizontal eye movements. The EEG and EOG were amplified with a bandpass of 0.01–80 Hz and recorded with a 500 Hz sampling rate. Impedance values were kept below 5 kΩ. We performed independent component analysis (ICA) to decompose the data, and we removed any trials containing either blinks (larger than 50 μV in the drift magnitude of FPz) or eye movement (larger than 30 μV in the drift magnitude of horizontal EOG activity).

EEG activity was calculated using a baseline from 200 ms to 0 ms before the onset of the stimulus array. ERPs were calculated by averaging baseline activity at each electrode across all accurate trials within each condition (unmasked: set size 1, 2, or 3; masked: set size 1, 2, or 3). We averaged the response from a set of five electrode pairs in the posterior-occipital, posterior-parietal, and temporal sites: O1/2, PO3/4, PO7/8, P3/4, and P7/8. The CDA was analyzed in the 1300–1700 ms time windows (i.e., 300–700 ms after the motion onset; Drew and Vogel (2008)). The contralateral and ipsilateral waveforms were defined in terms of the side attended to in a given trial. The difference wave was calculated by subtracting the ipsilateral activity from the contralateral activity. Clear illustrations in the figures are provided by the low-pass filtering of the data (14 Hz).



b

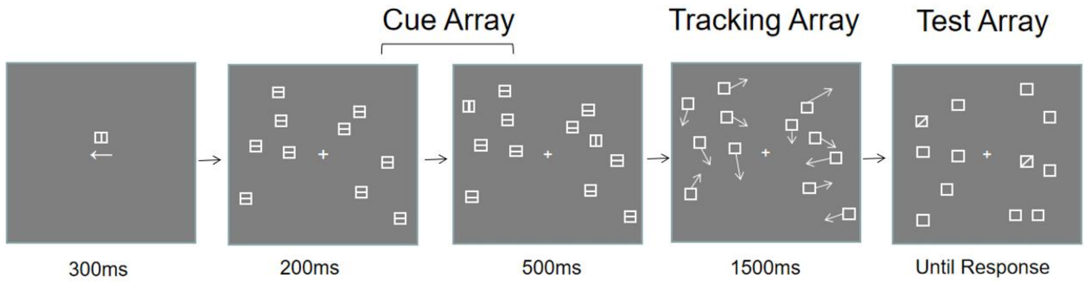


Figure 4. Stimulus and behavioral procedure in Experiment 2 in the (a) unmasked condition and (b) masked condition. Trial events are depicted from left to right. At the start of each trial, an arrow and a square with a horizontal or vertical bar (50% each) appeared for 300 ms, indicating that the participants should pay attention to one side of the screen and track the target(s) with the same bar. In the cue array, the nontargets appeared for 200 ms, and then the targets were presented for a later 500 ms with nontargets. Six squares (including nontargets) in total were consistently present, while the number of targets (1, 2, or 3) varied across blocks. All squares then moved independently, lasting 1500 ms. The bars in these squares remained visible in the unmasked condition, but they were removed in the masked condition. Following the motion, six stationary hollow squares were presented on each hemifield, and the participants were asked to identify whether the square with a slash on the attended side was (one of) the tracked target(s).

3.3 Results

3.3.1 Statistical analysis

We first used repeated measures ANOVAs to analyze the accuracy and the CDA amplitude, and paired t-tests and simple effect analyses were conducted for the follow-up pairwise comparison. For accuracy, We also computed the number of accurately tracked objects using a formula developed by Scholl et al. (2001), $M = n \times (2 \times P - 1)$, where M is an estimate of the number of objects accurately tracked, n is the total number of tracking objects, and P is the observed accuracy. The significance level was set as $p < 0.05$ across all tests, but a marginally significant ($p < 0.1$) result was also reported.

3.3.2 Behavioral results

We conducted a 2 (masked and unmasked condition) × 3 (Tracks 1, 2, and 3) repeated measures ANOVA to analyze accuracy. As Figure 5a shows, the results revealed a significant main effect of condition (F [1, 23] = 27.326, p < 0.001, η^2_p = 0.543) and tracking load (F [2, 22] = 13.737, p < 0.001, η^2_p = 0.555), with a marginal significant interaction between condition and tracking load (F [2, 22] = 3.392, p = 0.052, η^2_p = 0.236).

The paired t-test results showed significance for the accuracy between Tracks 1 and 2 (t [47] = 3.202, p = 0.002, Cohen's d = 0.423), Tracks 1 and 3 (t [47] = 6.1238, p < 0.001, Cohen's d = 1.791), and Tracks 2 and 3 (t [47] = 4.116, p < 0.001, Cohen's d = 1.201), merging the unmasked and masked conditions. Basically, the accuracy decreased as the number of tracked targets increased.

The paired t-test also showed that accuracy was better in the unmasked condition (M = 0.942, standard deviation (SD) = 0.050) than in the masked condition (M = 0.897, SD = 0.072), (t [71] = 6.232, p < 0.001, Cohen's d = 1.479). Critically, all tracking loads showed better accuracy in the unmasked than masked condition (Track 1: unmasked condition > masked condition; t [23] = 2.684, p = 0.013, Cohen's d = 1.119; Track 2: unmasked condition > masked condition; t [23] = 3.330, p = 0.003, Cohen's d = 1.3389; and Track 3: unmasked condition > masked condition; t [23] = 4.789, p < 0.0011, Cohen's d = 1.4479).

we also calculated M score to estimate the number of objects that were accurately tracked. The average *M* scores in Tracks 1, 2, and 3 are shown in Figure 5b for the masked and unmasked conditions. The modulation of the tracking load could effectively index individual tracking performance. *M* scores showed that there was no contamination from the ceiling effect.

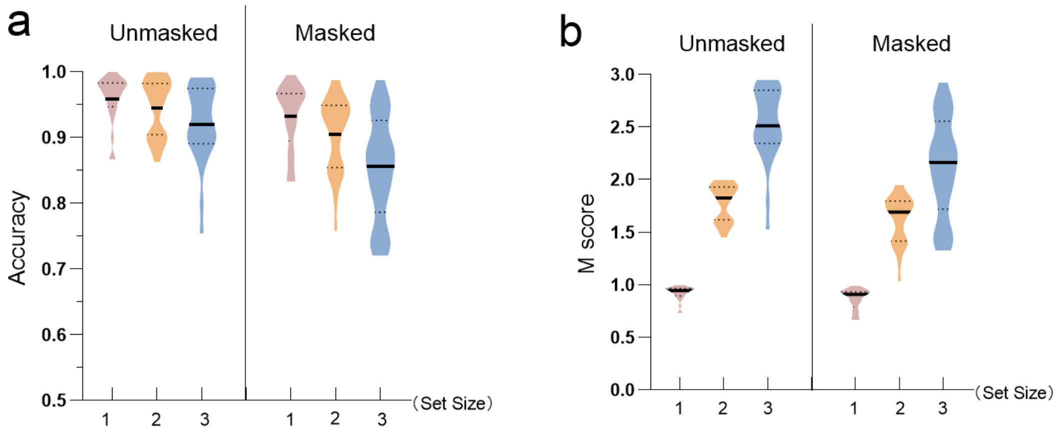


Figure 5. (a) Accuracy and (b) the number of accurately tracked targets across the three types of tracking loads under the unmasked and masked conditions. The black line represents the average performance, and the dotted line represents the quartile. The plot represents the distribution of the M scores for all participants.

3.3.2 Electrophysiological results

We conducted a 2 (masked and unmasked condition) × 3 (Tracks 1, 2, and 3) repeated measures ANOVA of the average CDA . As shown in Figure 6, a significant main effect was detected for condition (F [1, 23] = 8.692, $p = 0.007$, $\eta^2_p = 0.274$) and for tracking load (F [2, 22] = 11.082, $p < 0.001$, $\eta^2_p = 0.502$). Importantly, a significant interaction was apparent between the two factors (F [2, 22] = 4.246, $p = 0.028$, $\eta^2_p = 0.279$), driven primarily by the larger effect of the tracking load in the masked condition.

The first important finding is that the CDA amplitude was more negative in the masked than unmasked condition (t [71] = 3.764, $p < 0.001$, Cohen's d = 0.893).

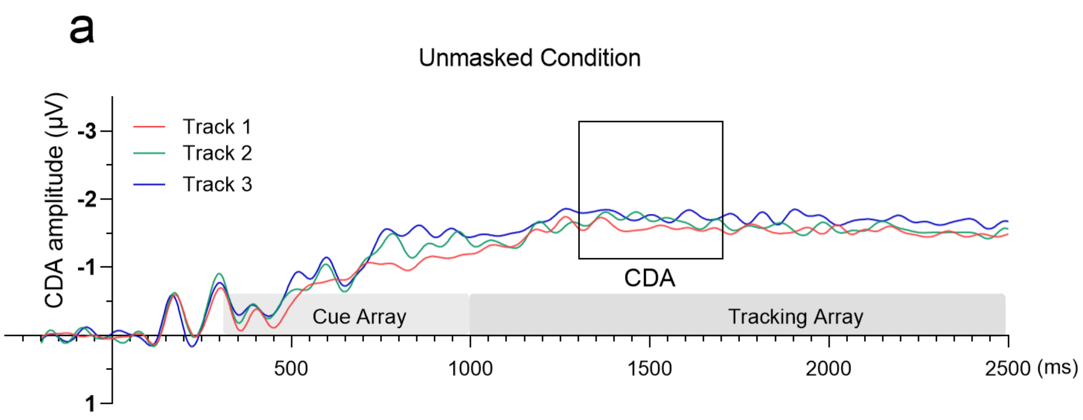
Critically, this significant finding is mainly driven by tracking three targets (masked condition > unmasked condition; t [23] = 2.684, $p = 0.013$, Cohen's d = 1.119) but not in Track 1 (t [23] = 1.329, $p = 0.197$, Cohen's d = 0.554) and Track 2 (t [23] = 1.565, $p = 0.131$, Cohen's d = 0.653).

Most importantly, there were different modes in the CDA amplitude between the unmasked and masked conditions. In the unmasked condition, the Tracks 1, 2, and 3 conditions yielded an equivalent CDA amplitude (F [2,22] = 0.783, p = 0.469, η^2_p = 0.066). The follow-up paired t-test did not show a difference between Tracks 1 and 2 (t [23] = 0.766, p = 0.452, Cohen's d = 0.319), Tracks 1 and 3 (t [23] = 1.276, p = 0.215, Cohen's d = 0.532), and Tracks 2 and 3 (t [23] = 0.460, p = 0.650, Cohen's d = 0.192).

However, CDA amplitude was significantly increased with the tracking load increase in the masked condition (F [2, 22] = 17.712, p < 0.001, η^2_p = 0.617). Specifically, even though the CDA amplitude did not show significance in Tracks 1 and 2 (t [23] = 1.480, p = 0.152, Cohen's d = 0.617) or Tracks 2 and 3 (t [23] = 1.756, p = 0.092, Cohen's d = 0.732), the CDA amplitude in Tracks 1 and 3 (t [23] = 6.065, p < 0.001, Cohen's d = 2.53) showed that there was a tendency toward an increasing CDA amplitude when the tracking load was added in the masked condition.

These results indicate that the CDA failed to show its sensitivity to the number of tracked targets if the targets were constantly visible during the tracking period. However, masking the tracked targets may play an important role when examining the potential covariance between CDA and the target load in the tracking stage.

Overall, the sensitivity of the CDA to the number of tracked targets was detected in the masked condition, where cognitively representational storage was necessary but not in the unmasked condition. These results support the interpretation that CDA observed in the MOT task was indeed derived from the representational storage of spatial information rather than the current focus of spatial attention.



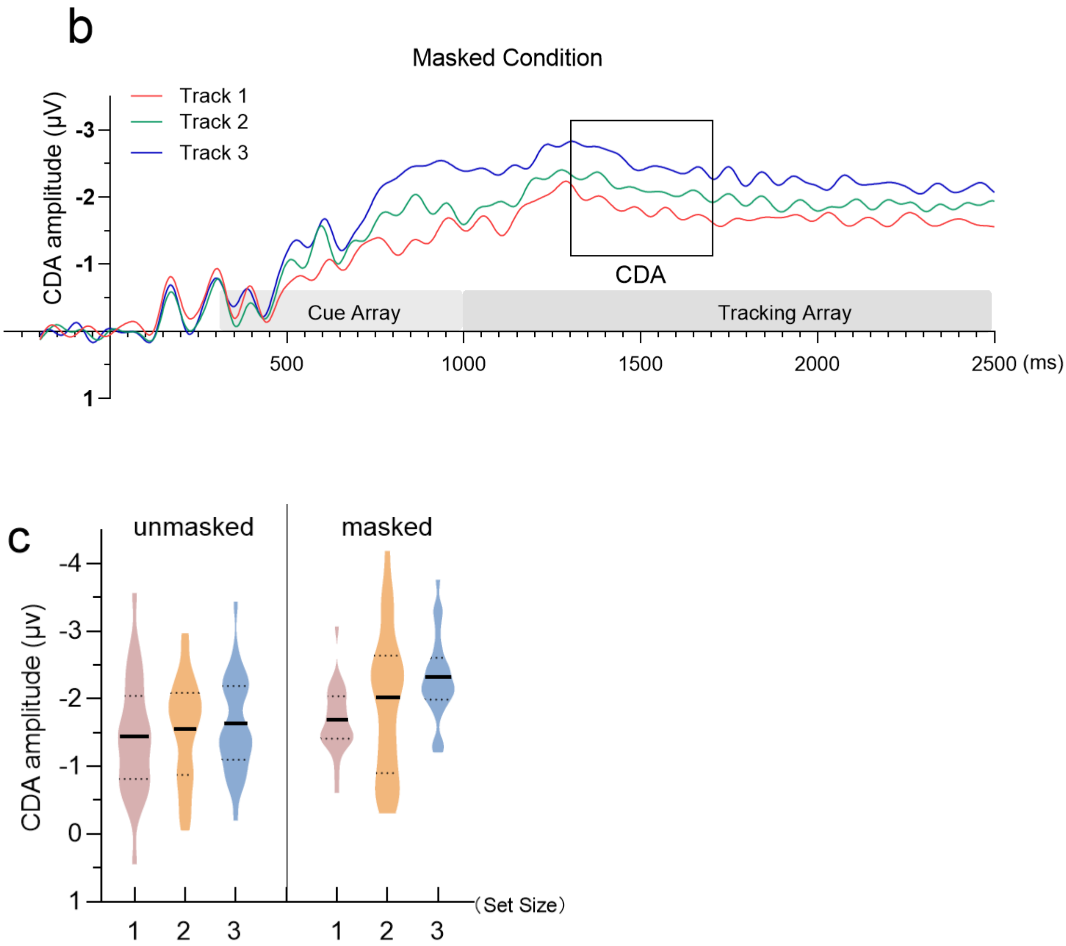


Figure 6. The amplitude of contralateral delay activity (CDA) over time in the (a) unmasked condition and (b) masked condition. Time 0 marks the onset of the arrow and the target cue, and Time 2500 ms marks the onset of the test array. (c) The average CDA amplitude is depicted in each condition during the 1,300–1,700 ms time window; the distributions of the CDA amplitudes for all participants are presented as violin plots. The black line represents the average, and the dotted line represents the quartile.

3.3 Discussion

In Experiment 2, we compared the CDA elicited by tracked targets across the masked and unmasked conditions that invoked cognitively distinct processes. Sustained spatial attention was recruited in both conditions, but representational storage was required only in the masked condition. These results suggest that the

CDA was highly sensitive to the number of tracked targets in the masked condition, whereas CDA showed no effect on the tracking load in the unmasked condition. Overall, a possible conclusion is that the CDA observed in the MOT tasks was directly linked to both sustained attentional activation and representational storage. Importantly, the effects of the tracking load on the CDA were specifically related to the necessity of representational storage for the tracked targets during the tracking period.

4. General Discussion

This study investigated whether the CDA observed in MOT tasks was derived from sustained spatial attention or representational storage. In Experiment 1, we performed a dual task to explore whether WM was recruited in an MOT task and to clarify the role played by WM. We found that the storage representations affected accuracy in the masked conditions but not in the unmasked conditions. In Experiment 2, we found that the unmasked and masked conditions elicited different CDA . These results suggest that the CDA was highly sensitive to the number of tracked targets in the masked condition, whereas CDA showed no effect on the tracking load in the unmasked condition.

Overall, one possible conclusion is that the CDA observed in the MOT tasks was directly linked to both sustained attentional activation and representational storage. Importantly, the effects of the tracking load on the CDA were specifically related to the necessity of representational storage for the tracked targets during the tracking period. The current results are broadly consistent with the overall results from previous MOT and VWM tasks (Drew & Vogel, 2008; Drew et al., 2011; Hakim et al., 2019).

Although both conditions involved item tracking, the possibility of distinguishing targets from nontargets merely by attention was the key factor in recruiting WM. In the unmasked condition, the act of allocating sustained spatial attention to the tracked targets was sufficient to update the current position of the tracked targets. This conclusion was confirmed by Experiment 2 in the unmasked condition. This result suggests that the CDA includes the reflection of sustained attention activation, which could account for the absence of the effect of the tracking load on the CDA . In contrast, given that the features of the targets were not visible during the tracking period and the position of the target items was constantly changing, the representations of the tracked targets should be plausibly maintained to update their current positions. This conclusion was proven by Experiment 2 in the masked condition. Thus, the CDA observed in the masked condition primarily indexed representational storage and tracked the variation in the tracking load, and it was also accompanied by the sustained attentional activation to the targets.

With respect to the effectiveness of the experimental design for the MOT tasks used in this study, the bar, rather than the salient color, was selected as the cue for the target items to avoid contamination from the salient effect and bottom-up process. This selection potentially encouraged participants to track the target items attentively, even in the unmasked condition. Consequently, the accuracy and electrophysiological signals observed in the unmasked condition could genuinely index the cognitive process of attentively tracking targets, and the two conditions involved cognitively identical processes, except for the necessity of representational storage for the tracked targets in the masked condition. In the meantime, we set all the conditions presented in different blocks to avoid possible confounders so that the participants would remember all targets and nontargets in case they lost any possible targets when they were hard to distinguish in the masked condition.

4.1 CDA reflection in the MOT task

In the current tracking task, the CDA amplitude was substantially larger in the masked than unmasked condition. The increased CDA amplitude in the masked condition might have been caused by the process of representational storage, which was not involved in the unmasked condition. Intriguingly, Drew et al. (2011) showed that the CDA was larger in the MOT task than in the VWM task. This result was interpreted as indicating that the increased CDA amplitude observed in the MOT task was attributable to the shifts in spatial attention or the process of updating target positions that were not involved in the VWM task. Drew et al. (2011) had also previously proven that attending purely to the motion itself was insufficient to generate the large tracking activity observed in the MOT task. In conjunction with the current results, cognitively complex processes might be conjectured to generate a larger CDA compared to relatively simple processes.

5. Conclusion

With the controversy regarding the index of the CDA , the current study conclusively demonstrated that both sustained attention activation and representational storage could similarly evoke the CDA . However, the defining feature, namely that CDA was sensitive to the number of tracked items, is derived from the process of representational storage. The current observations challenge the hypothesis by showing that CDA could reflect the storage of a set of tracked items and the attentional activation to them, but the CDA amplitude is only sensitive to VWM storage.

References

Adam, K. C., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The contribution of attentional lapses to individual differences in visual working memory capacity. *J Cogn Neurosci, 27*(8), 1601-1616. doi:10.1162/jocn_a_00811

Adam, K. C. S., Robison, M. K., & Vogel, E. K. (2018). Contralateral Delay Activity Tracks Fluctuations in Working Memory Performance. *J Cogn Neurosci, 30*(9), 1229-1240. doi:10.1162/jocn_a_01233

Balaban, H., & Luria, R. (2017). Neural and Behavioral Evidence for an Online Resetting Process in Visual Working Memory. *J Neurosci, 37*(5), 1225-1239. doi:10.1523/JNEUROSCI.2789-16.2016

Berggren, N., & Eimer, M. (2016). Does Contralateral Delay Activity Reflect Working Memory Storage or the Current Focus of Spatial Attention within Visual Working Memory? *J Cogn Neurosci, 28*(12), 2003-2020. doi:10.1162/jocn_a_01019

Berggren, N., & Eimer, M. (2020). Attentional Access to Multiple Target Objects in Visual Search. *J Cogn Neurosci, 32*(2), 283-300. doi:10.1162/jocn_a_01476

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev Neuropsychol, 33*(3), 205-228. doi:10.1080/87565640801982312

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends Cogn Sci, 9*(7), 349-354. doi:10.1016/j.tics.2005.05.009

Cohen, M. A., Pinto, Y., Howe, P. D., & Horowitz, T. S. (2011). The what-where trade-off in multiple-identity tracking. *Atten Percept Psychophys, 73*(5), 1422-1434. doi:10.3758/s13414-011-0089-7

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends Cogn Sci, 7*(12), 547-552. doi:10.1016/j.tics.2003.10.005

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci, 24*(1), 87-114; discussion 114-185. doi:10.1017/s0140525x01003922

Daneman., M., & Carpenter., P. (1980). Individual Differences in Working Memory and Reading *JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR*.

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2011). Delineating the neural signatures of tracking spatial position and working memory during attentive tracking. *J Neurosci, 31*(2), 659-668. doi:10.1523/JNEUROSCI.1339-10.2011

Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *J Neurosci, 28*(16), 4183-4191. doi:10.1523/JNEUROSCI.0556-08.2008

Drew., T., Horowitz., T. S., Wolfe., J. M., & Vogel., E. K. (2012). Neural Measures of Dynamic Changes in Attentive Tracking Load. *Journal of Cognitive Neuroscience*, 10.

Feldmann-Wustefeld, T., Vogel, E. K., & Awh, E. (2018). Contralateral Delay Activity Indexes Working Memory Storage, Not the Current Focus of Spatial Attention. *J Cogn Neurosci, 30*(8), 1185-1196. doi:10.1162/jocn_a_01271

Fougnie, D., & Marois, R. J. P. S. (2006). Distinct Capacity Limits for Attention and Working Memory : Evidence From Attentive Tracking and Visual Working Memory Paradigms. *17*(6), 526-534.

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review, 17*(5), 673-679. doi:10.3758/17.5.673

Gao, Z., Yin, J., Xu, H., Shui, R., & Shen, M. (2011). Tracking object number or information load in visual working memory: revisiting the cognitive implication of contralateral delay activity. *Biol Psychol, 87*(2), 296-302. doi:10.1016/j.biopsycho.2011.03.013

Hakim, N., Adam, K. C. S., Gunseli, E., Awh, E., & Vogel, E. K. (2019). Dissecting the Neural Focus of Attention Reveals Distinct Processes for Spatial Attention and Object-Based Storage in Visual Working Memory. *Psychol Sci, 30*(4), 526-540. doi:10.1177/0956797619830384

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. doi:10.1152/jn.00978.2009

Jarrod A. Lewis-Peacock, A. T. D., Klaus Oberauer %J Journal of Cognitive Neuroscience. (2012). Neural Evidence for a Distinction between Short-term Memory and the Focus of Attention. *24*(1), 61-79.

Kang, M. S., & Woodman, G. F. (2014). The neurophysiological index of visual working memory maintenance is not due to load dependent eye movements. *Neuropsychologia, 56*, 63-72. doi:10.1016/j.neuropsychologia.2013.12.028

Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cogn Psychol, 52*(4), 346-368. doi:10.1016/j.cogpsych.2005.12.001

Kundu, B., Sutterer, D. W., Emrich, S. M., & Postle, B. R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *J Neurosci, 33*(20), 8705-8715. doi:10.1523/JNEUROSCI.5565-12.2013

Larocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Front Hum Neurosci, 8*, 5. doi:10.3389/fnhum.2014.00005

Li, J., Oksama, L., & Hyona, J. (2019). Model of Multiple Identity Tracking (MOMIT) 2.0: Resolving the serial vs. parallel controversy in tracking. *Cognition, 182*, 260-274. doi:10.1016/j.cognition.2018.10.016

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn Sci, 17*(8), 391-400. doi:10.1016/j.tics.2013.06.006

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neurosci Biobehav Rev, 62*, 100-108. doi:10.1016/j.neubiorev.2016.01.003

Luria, R., Sessa, P., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. J. C. N. J. o. (2010). Visual short-term memory capacity for simple and complex objects. *22*(3), 496-512.

Oksama, L., & Hyona, J. (2008). Dynamic binding of identity and location information: a serial model of multiple identity tracking. *Cogn Psychol, 56*(4), 237-283. doi:10.1016/j.cogpsych.2007.03.001

Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher - order cognition? An individual difference approach. *Visual Cognition, 11*(5), 631-671. doi:10.1080/13506280344000473

Pilling, M., & Barrett, D. J. (2016). Dimension-based attention in visual short-term memory. *Mem Cognit, 44*(5), 740-749. doi:10.3758/s13421-016-0599-6

Scholl, B., and, Pylyshyn, Z., and, & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*.

Storm, R. W., & Pylyshyn, Z. W. J. S. V. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *3*(3), 179-197.

Süß., H.-M., Oberauer., K., Wittmann., W. W., Wilhelm., O., & Schulze., R. (2012). Workingmemory capacity explains reasoning ability—and a little bit more. *Intelligence*, 27.

Tombu, M., & Seiffert, A. E. (2008). Attentional costs in multiple-object tracking. *Cognition, 108*(1), 1-25. doi:10.1016/j.cognition.2007.12.014

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural limits to representing objects still within view. *J Neurosci, 33*(19), 8257-8263. doi:10.1523/JNEUROSCI.5348-12.2013

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2015). Working memory delay activity predicts individual differences in cognitive abilities. *J Cogn Neurosci, 27*(5), 853-865. doi:10.1162/jocn_a_00765

Vergauwe, E., Hardman, K. O., Rouder, J. N., Roemer, E., McAllaster, S., & Cowan, N. (2016). Searching for serial refreshing in working memory: Using response times to track the content of the focus of attention over time. *Psychon Bull Rev, 23*(6), 1818-1824. doi:10.3758/s13423-016-1038-1

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(7067), 500-503. doi:10.1038/nature04171

Wang, S., Rajsic, J., & Woodman, G. F. (2019). The Contralateral Delay Activity Tracks the Sequential Loading of Objects into Visual Working Memory, Unlike Lateralized Alpha Oscillations. *J Cogn Neurosci, 31*(11), 1689-1698. doi:10.1162/jocn_a_01446

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychon Bull Rev, 15*(1), 223-229. doi:10.3758/pbr.15.1.223