

# Two Discoveries, One Principle: Using a Two-Stage Bayesian Model to Explain a Dissociated Working Memory Distraction Effect

Ziyi Duan<sup>1</sup>, Yushang Huang<sup>2</sup>, Bohao Shi<sup>3</sup>, and Xiaowei Ding<sup>1</sup>

<sup>1</sup> Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health,

Department of Psychology, Sun Yat-sen University

<sup>2</sup> School of Education Science, Zhoukou Normal University

<sup>3</sup> Department of Psychology and Behavioral Sciences, Zhejiang University

How working memory (WM) resists perceptual distraction with its limited capacity is a fundamental question to understand its mechanism. To address this question, we used a continuous recall paradigm to directly compare the distraction effect during encoding and the delay periods. Across Experiments 1–3, we observed a substantial distraction-related cost on mnemonic fidelity when distractors presented during the delay (delay-distraction condition), but not if they were introduced at encoding (encoding-distraction condition) or across both periods (full-distraction condition). However, the distraction cost revived when we increased the difficulty to distinguish distractors from targets (Experiments 4 and 4S) and when we changed distractors' relevant features during the delay (Experiment 5). We also found that the robust distraction cost in the delay-distraction condition did not occupy extra spatial resources (Experiments 6a and 6b). These results suggested a dissociated distraction effect, which could be related to the dynamic resource allocation across two WM periods. Here, we proposed a Bayesian model and considered the task relevance and visual uncertainty as two main factors that determine the resource allocation principle at two different stages. This model successfully captured the main findings across all behavioral experiments and performed better than other alternative models. Taken together, the current work advanced our understanding of the distraction resistance of WM under the framework of limited resource allocation.

**Keywords:** working memory, distraction, cognitive resource, Bayesian model

**Supplemental materials:** <https://doi.org/10.1037/xge0001358.supp>

Working memory (WM), a system holding mental representations temporarily available for use in thoughts and actions (Baddeley, 2012; Cowan, 2017; D'Esposito et al., 1995; Oberauer et al., 2016, 2018; Smith & Jonides, 1999), is characteristic by its limited capacity (Anderson et al., 1996; Case, 1972; Just & Carpenter, 1980; Ma et al., 2014). The ability to properly allocate resources to goal-relevant information and filter out irrelevant information serves as a potential way to handle the extreme limitations in WM capacity (Arnell & Stubitz, 2010; Awh & Vogel, 2008; Conway et al., 2003; Fukuda & Vogel, 2009, 2011; Jost & Mayr, 2016; Liesefeld et al., 2014, 2020; Lorenc et al., 2021; Vogel & Awh, 2008; Vogel et al., 2005). Therefore, understanding how WM allocates

its limited capacity to resist distraction is a core question that needs to be addressed.

Resource allocation exists both during encoding and the delay periods of WM process. For example, researchers used pre-cues or retro-cues to study external and internal attentional allocations (e.g., Griffin & Nobre, 2003; Vogel et al., 2005). Besides, resource allocation was also studied by presenting perceptual distractors simultaneously with targets at encoding or during the delay (e.g., Feldmann-Wüstefeld & Vogel, 2019; McNab & Dolan, 2014; Tabi et al., 2021). Previous studies indicated different mechanisms during encoding and the delay periods. Selectively encoding relevant information was regarded as filtering ability at encoding (e.g.,

---

This article was published Online First March 16, 2023.

Xiaowei Ding  <https://orcid.org/0000-0002-0583-291X>

This work was supported by grants from the National Natural Science Foundation of China (32271103), Humanities and Social Sciences Foundation of the Ministry of Education of China (19YJC190004), Natural Science Foundation of Guangdong Province (2021A1515011103), and Fundamental Research Funds for Colleges and Universities-Key Training Program for Young Teachers(19wkzd23) awarded to author Xiaowei Ding.

We have no conflicts of interest to disclose.

The datasets supporting this article is publicly available at the Open Science Framework: <https://osf.io/j8b92/>.

Ziyi Duan served as lead for conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing-

original draft, and writing-review and editing. Yushang Huang served as lead for validation and contributed equally to formal analysis, and served in a supporting role for writing-review and editing. Bohao Shi served in a supporting role for formal analysis, validation, and writing-review and editing. Xiaowei ding served as lead for conceptualization, funding acquisition, supervision, writing-original draft, and contributed equally to investigation and writing-review and editing. Ziyi Duan, Bohao Shi, and Xiaowei Ding contributed to methodology equally.

Correspondence concerning this article should be addressed to Xiaowei Ding, Department of Psychology, Sun Yat-sen University, 132 East Waihuan Road, Guangzhou 510006, China. Email: [dingxw3@mail.sysu.edu.cn](mailto:dingxw3@mail.sysu.edu.cn)

Fukuda & Vogel, 2009, 2011; Liesefeld et al., 2020; McNab & Klingberg, 2008). Nevertheless, control processes that inhibit distractor processing and prevent unwanted information from being encoded were emphasized during the delay. Previous work also indicated unique contribution to WM capacity from encoding and delay distractor resistance (McNab & Dolan, 2014). Clinical researchers also found that patients with deficient striatal dopamine had decreased distractor-filtering ability at encoding (Cools et al., 2010; Mehta et al., 2004) but increased ability during the delay (McNab & Klingberg, 2008). However, systematic investigation of how these two stages influence the success of distraction resistance is lacking.

Regarding limited resource allocation, we already have a range of widely acknowledged theoretical models, which provided us profound understandings (e.g., slot model: Luck & Vogel, 1997; Zhang & Luck, 2008; resource model: Bays et al., 2009; Bays & Husain, 2008; Ma et al., 2014; van den Berg et al., 2012; Wilken & Ma, 2004; and interference model: Oberauer & Lin, 2017). However, none of these models provided a deliberative account of the distractor-induced interference taking both periods of WM tasks into consideration. Details about the allocation of cognitive resources at different stages remain to be filled in. In the current study, we focused on both the theoretical explanation and computational mechanism underlying how people allocated limited WM resources when facing perceptual distractors.

In general, we regarded the WM task as a Bayesian inference process and the distraction effect as competing limited WM resources between targets and distractors. To capture the Bayesian processes in the resource allocation process, we formulized the continuous information-sampling process by using samples to simulate internal representations. The limited capacity of sampling is an analogy of limited resources. The sampling number of a specific stimuli would be higher if it gained more resources, resulting in a more precise internal representation of it, which can be described by a more concentrated distribution with smaller *SD* in a computational framework.

To build a unified principle of resource allocation between targets and distractors across the encoding and delay periods, we referred to the extrinsic and intrinsic motivation, which drives the information sampling process (Tishby & Polani, 2011). The extrinsic motivation was associated with classical task-directed objectives, denoted as the *task relevance*. The intrinsic motivation was based on inherent cognitive features to reduce momentary uncertainty, denoted as the *visual uncertainty*. Both task relevance and visual uncertainty served as key parameters that influenced the resource allocation principle in our model (Gottlieb et al., 2013).

Specifically, *task relevance* represents the priority of an item required to perform well in the current task, which is determined by both top-down task goals and bottom-up stimulus saliency. This idea is analogical to the biased-competition framework (Bundesen et al., 2011; Desimone & Duncan, 1995) and the priority map theory (Fecteau & Munoz, 2006; Serences & Yantis, 2006; Sprague et al., 2018), which assumes that top-down modulations bias bottom-up saliency signals, resulting in a priority map coding for behavioral relevance. Therefore, task relevance can be determined by the task requirement (e.g., targets vs. distractors). Besides, it is also influenced by stimulus saliency. For example, targets with less prominent features to be separated from distractors may have less task relevance than targets with obvious features. Furthermore, it may differ across different

observers with the same task requirement at different times (de Vries et al., 2020; Stokes, 2015; Wan et al., 2020). For example, when asking participants to remember two items simultaneously and perform two tasks related to each one, participants can flexibly hold the one that is most relevant to the current task and switch to the next when the goal changes (e.g., Rose et al., 2016; van Loon et al., 2018; Yu & Postle, 2018). Therefore, it can be adjusted according to different cognitive goals in real time, capturing the flexibility of WM.

*Visual uncertainty* is defined as the degree of unfamiliarity with each item, which can be calculated by the information entropy or variance of its probabilistic distribution. In probability and decision theory, the only consistent way for modeling and reasoning about uncertainty is provided by the Bayesian theory of probability (Cox, 1964; Jaynes, 2003; Savage, 1972). The hierarchical Bayesian view provides us with a uniform framework to understand the perception and action by assuming the brain is trying to optimize the evidence for its model of the world (Feldman & Friston, 2010; Friston, 2009). Decreasing uncertainty about future states serves as a natural motivation that drives our attention and behaviors. For example, previous research found that uncertainty in WM can promote the resampling of the environment to reduce the uncertainty (Sullivan et al., 2012), similar to the situation when people tend to try harder figuring out something ambiguous, especially when it appears to be interesting. In neuroscience research, sensory novelty is known to enhance neural responses throughout visual, frontal, and temporal areas (Gottlieb et al., 2013; Ranganath & Rainer, 2003) and activate reward-responsive dopaminergic areas (Berlyne, 1950; Düzel et al., 2010; Lehman & Stanley, 2011), serving as an intrinsic motivation to lead attention to uncertain targets. According to single neuron studies, neurons have the novelty tendency that responds more strongly to a new stimulus than a familiar one is known as the novelty preference. To capture the intrinsic motivation of information sampling and predict human behaviors, we adopted visual uncertainty to represent the processing uncertainty to each visual item. Notably, our visual uncertainty is a general term including both spatial and non-spatial features (e.g., location-related or orientation-related uncertainty). The lower uncertainty in the encoding of stimuli into memory representation provides, the better probing performance can be, due to more solid evidence for inference. This is consistent with the Bayesian view that regarding every cognitive process as probabilistic inference (Yuille & Kersten, 2006).

Combining these two components with the information-sampling process, we could make explicit expectations about how distractors would influence the WM performance across two periods. The key factor was the overall competitive power formed by both task relevance and visual uncertainty. At encoding, the primary driving factor was task relevance since both targets and distractors had equal high visual uncertainty at the beginning. As long as the relative task relevance for targets over distractors was high, observers could successfully resist distractions. During the delay, the visual uncertainty factor began to play its role. If distractors were first introduced during the delay or their features have changed during the delay compared to the encoding period, the visual uncertainty of distractors would be much higher than targets. Therefore, a larger amount of limited sampling would be given to distractors, instead of being used for the maintenance of targets. Reduced samplings naturally lead to lower memory precision for targets. By the above-mentioned speculation, we proposed a new perspective to

understand the allocation of limited WM resources when facing distractors, which was fully implementable in a computational model. The full model description would be done after reporting experiment results as a part of the discussion section.

## The Current Experiments

In the current study, we used a continuous recall paradigm to directly compare the distraction-related effect during encoding and the delay periods. Specifically, participants were asked to remember orientations of three teardrops. The continuous report error of targets was recorded on behalf of memory performance. Importantly, we introduced another three teardrops (distractors) at encoding, during the delay, or across both periods (Experiments 1–3). We showed a robust distraction effect when distractors were presented during the delay (delay-distraction condition) but not at encoding (encoding-distraction condition). More interestingly, participants could also resist distractors' interference when distractors were presented across both periods (full-distraction condition). In the following experiments, we further investigated the underlying mechanism of the dissociated distraction effect. In Experiments 4 and 4S, we decreased the distinguishability between targets and distractors by using precues (same-color conditions) and found distractors significantly impaired WM performance. Even the encoding-distraction and the full-distraction conditions had significant distraction effect in this circumstance. Moreover, we found that changing features of distractors during the delay also revived the distraction-related interference period in the full-distraction condition (Experiment 5). Finally, we revealed that the robust distraction cost in the delay-distraction condition did not need to occupy additional spatial resources (Experiments 6a and 6b).

## Experiment 1

To investigate the interference effect on mnemonic fidelity of WM during different periods, we introduced distractors at encoding or during the delay when participants performed the continuous recall task (Bays et al., 2009; Bays & Husain, 2008; Wilken & Ma, 2004; Zhang & Luck, 2008). We used different colors (red or green) to distinguish targets and distractors. Participants needed to remember three teardrops' orientation at each time while ignoring the distractors no matter when they appeared. We introduced three conditions (no-distraction, encoding-distraction, and delay-distraction) corresponding to without distractors, distractors appearing simultaneously with targets at encoding, and after targets disappeared during the maintenance period.

## Method

### Participants

Twenty-four volunteers (ages 18–25, 7 male, 17 female) participated in Experiment 1, which was exploratory with respect to the sample size. Sensitivity power analysis (G\*Power 3.1) estimated a minimum detectable effect (i.e., the smallest true effect, which would be statistically significant with  $\alpha = 0.05$ , and power  $\beta = 0.80$ ) of  $\eta_p^2 = 0.07$  (equals to  $f = 0.27$ ) for the main effect in a three-repeated measurement within-subjects design with a sample size of 24. One participant quitted the experiment midway, resulting in 23 participants in total. The actual effect size of this main

effect in Experiment 1 was used to evaluate the appropriateness of the sample size to obtain a similar effect size in the following experiments. All participants had normal or corrected-to-normal vision as well as normal color vision, gave informed written consent, and were remunerated ¥60–70 (approximately US\$8–10) for their time. All experimental procedures in our current study were approved by the Sun Yat-Sen University Institutional Review Board.

### Stimuli

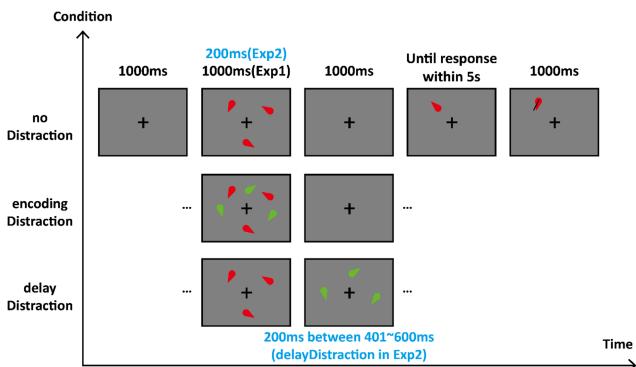
We used teardrops as stimuli ( $1.0^\circ \times 0.5^\circ$  visual angle) for both targets and distractors, whose orientation was selected from one of 12 orientations from  $0^\circ$  to  $360^\circ$  in steps of  $30^\circ$ . To avoid participants using memory strategies, we randomly jittered the stimuli  $1^\circ\text{--}5^\circ$ . Therefore, no overlap existed between any two stimuli, including both targets and distractors. Targets and distractors were distinguished by colors (red [RGB: 255, 0, 0] or green [RGB: 0, 255, 0]). Participants sat in a dark room 57 cm in front of a 17-in. CRT monitor, whose refresh rate was 100 Hz and screen resolution was  $1,024 \times 768$  pixels. The stimuli were displayed on a gray (RGB: 128, 128, 128) background, where each teardrop was distributed within an invisible rectangle ( $4.0^\circ \times 6.0^\circ$ ) with the constraint that every two teardrops should be spaced more than  $1.6^\circ$  apart. The task was programmed using Matlab (Version 9.2.0) and Psychtoolbox extensions (Brainard, 1997).

### Design and Procedure

To explore how distractors presented at encoding or during the delay period influence WM fidelity, we distinguished three different

**Figure 1**

Schematic Illustration of the No-Distraction, Encoding-Distraction, and Delay-Distraction Conditions in Experiments 1 and 2



*Note.* In Experiment 1, participants were required to remember all three target teardrops' orientations (red ones in this example) and ignore the distractors (green ones) when present. After a 1,000-ms delay, one of the target stimuli was tested by asking participants to recall its orientation with the mouse. After their confirmation or 5 s no response, the feedback was provided. The only difference in Experiment 2 was the time duration for the distractors, which was reduced to 200 ms for both (blue color). Notably, the delay period remained 1,000 ms in Experiment 2, while distractors were only presented for 200 ms in the middle. See the online article for the color version of this figure.

conditions (no-distraction condition, encoding-distraction condition, and delay-distraction condition) (see Figure 1). On the baseline condition, each trial began with a central fixation ( $1.5^\circ \times 1.5^\circ$ ) presented for 1,000 ms. Then, three red or green teardrops (balanced between subjects) with different orientations were presented for 1,000 ms, and the participants were asked to remember all the orientations. After another 1,000 ms blank interval, one of the target stimuli would reappear with a random orientation. Participants were asked to adjust the teardrop's orientation to its initial ones by moving the mouse within 5 s. After their confirmation (by pressing the space key), the correct orientation would be presented by using a black line segment. If participants could not respond within the time range, the feedback would be displayed automatically. The intertrial interval was randomly selected from a uniform distribution between 2 and 3 s.

As for the encoding-distraction condition, everything was identical except for the stimuli onset period, where the three target teardrops were presented with three distractor teardrops in a different color. Participants were instructed only to remember the target teardrops (red in half of participants) and ignore the distractors (green in half of participants) before the experiment. Similarly, in the delay-distraction condition, the only difference from the baseline trials was during the 1,000-ms blank interval, which was replaced by three distractors presented with different locations compared to the targets for 1,000 ms.

This experiment took a within-subject design with a single factor (no-distraction, encoding-distraction, and delay-distraction). Each condition had 150 trials, which were divided into nine blocks randomly. A 3-min break was added between every two blocks, and the total duration was around 90 min. Before the formal trials, 30 practice trials were required to ensure every participant understood the whole procedure.

### Data Analysis

For each condition of interest, distribution response errors were obtained by calculating the difference between target and response (reported orientation minus probed target orientation). Response error distributions were centered around  $0^\circ$  of error in a circular normal space (e.g.,  $-180^\circ$  is the same as  $180^\circ$  of error). We estimated relevant characteristics from these response distributions by fitting a von Mises function (the circular analog of the Gaussian distribution), which describes the data in terms of the mean ( $\mu$ ) and circular variance ( $SD$ ). The  $SD$  parameter is inversely proportional to the precision of the memory item.

We also fitted the data with the prominent mixture-model analysis, which assumed that the distribution of response errors should reflect a mixture of a von Mises distribution centered on the true feature and a uniform distribution (Zhang & Luck, 2008). The model has two parameters: guess rate ( $g$ ) and standard deviation ( $SD$ ). Since this model requires extra assumptions and the experimental effects were consistent between  $g$  and  $SD$  across all experiments, we did not report it here. Detailed results were provided in the online supplemental material.

Data were fitted using the MemToolbox (Suchow et al., 2013). Parameter estimates were derived using Markov Chain Monte Carlo (MCMC) to find the maximum a posteriori values given the errors on the task. Notably, MRVL is positively related to memory quality while  $SD$  is negatively related to it. That is, the higher

MRVL, the lower  $SD$ , and the better memory precision for the item's orientation. The model was fit separately for each subject and task condition. To determine whether different conditions affected the quality and quantity of memory, parameter estimates were compared using a one-way repeated-measures analysis of variance (ANOVA) to determine whether  $SD$  differed between conditions. Data in all experiments have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/j8b92/>.

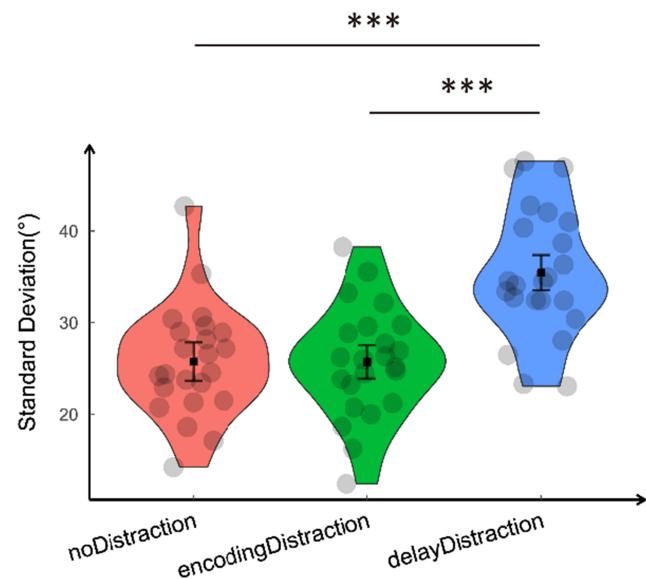
### Results

We tested whether the distractors presented at encoding or during maintenance period would influence the mnemonic fidelity in WM. In general, the one-way repeated-measures ANOVA revealed a significant main effect of conditions on  $SD$ ,  $F(2, 44) = 35.60$ ,  $p < .001$ ,  $\eta_p^2 = 0.62$ . Post hoc  $t$  tests with Bonferroni correction revealed that  $SD$  was significantly higher in delay-distraction condition compared with the other two conditions,  $t_{(\text{no-delay})} (22) = -6.83$ ,  $p < .001$ , Cohen's  $d = 2.15$ ;  $t_{(\text{encoding-delay})} (22) = -8.00$ ,  $p < .001$ , Cohen's  $d = -2.16$ . However, no difference was found between encoding-distraction condition and no-distraction condition in  $SD$ ,  $t_{(\text{no-encoding})} (22) = 0.04$ ,  $p = 1.00$ , Cohen's  $d = 0.01$  (see Figure 2).

### Experiment 2

To rule out the possible influence of processing time on mnemonic fidelity, we shortened the presentation time of both targets

**Figure 2**  
Results for Experiment 1



*Note.* Standard deviation ( $SD$ ) for no-distraction (red), encoding-distraction (green), and delay-distraction (blue) conditions. Grey dots represent single participant's data. Black dots indicate condition means and error bars denote 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

\*\*\* $p < .001$ .

and distractors to 200 ms in Experiment 2, comparable with the duration used in many WM studies (Park et al., 2017).

## Method

### Participants

Twenty-four healthy adults with normal or corrected-to-normal vision as well as normal color vision (ages 18–25, 9 male, 15 female) participated as paid volunteers. We kept the same sample size of Experiment 1 ( $N = 24$ ), which was larger than the required sample size ( $N = 3$ ) to obtain an effect size of the main effect, comparable to Experiment 1 ( $\eta_p^2 = 0.62$ ,  $\alpha = 0.05$ ,  $\beta = 0.80$ ; G\*power 3.1). Having the same sample size, the sensitivity of this test for three-way interaction (i.e., the minimum detectable effect size) was equal to Experiment 1 ( $\eta_p^2 = 0.07$ ,  $\alpha = 0.05$ ,  $\beta = 0.80$ ; G\*power 3.1). Thus, we kept this sample size for the rest of our experiments.

### Procedure

The procedure was the same as Experiment 1, except that the encoding time of both targets and distractors was shortened. Specifically, in all three conditions, targets were presented on the screen for 200 ms at encoding, and the probe tested after a 1,000 ms interval. In the no-distraction and encoding-distraction condition, the interval was a blank screen with the central fixation. In the delay-distraction condition, however, the distractors appeared in the middle 200 ms of the whole interval (i.e., from 401 to 600 ms) (see Figure 1).

## Results and Discussion

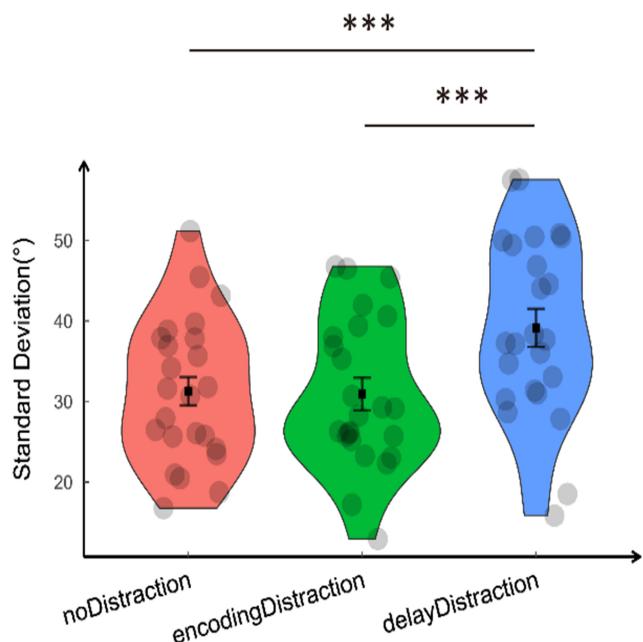
In Experiment 2, the same results were found. One-way repeated-measures ANOVA revealed a significant main effect of conditions on  $SD$ ,  $F(2, 46) = 21.44$ ,  $p < .001$ ,  $\eta_p^2 = 0.48$ . Post hoc  $t$  tests with Bonferroni correction revealed that  $SD$  was significantly higher in delay-distraction condition compared with the other two conditions,  $t_{(\text{no-delay})} (23) = -5.43$ ,  $p < .001$ , Cohen's  $d = -1.60$ ;  $t_{(\text{encoding-delay})} (23) = -5.13$ ,  $p < .001$ , Cohen's  $d = -1.67$ . However, no difference was found between encoding-distraction condition and no-distraction condition in  $SD$ ,  $t_{(\text{no-encoding})} (23) = 0.32$ ,  $p = 1.00$ , Cohen's  $d = 0.08$  (see Figure 3).

Discoveries from Experiments 1 and 2 demonstrated that during both long and short intervals, distractors presented during maintenance interfere with the fidelity of WM. However, this distraction effect disappeared when distractors appeared at encoding, which indicated a dissociable influence of visual distraction on the fidelity of WM at different stages.

## Experiment 3

According to Experiments 1 and 2, we observed a dissociable effect of distractors on WM performance. Then, we wondered whether the distraction effect in delay-distraction condition was mainly caused by the appearance of distractors during the delay, or the absence of processing distractors at encoding? To answer this question, we added a full-distraction condition in Experiment 3, where distractors were presented throughout both the encoding and delay period. There are two possible predictions. First, if WM performance in the full-distraction condition has no cost on memory

**Figure 3**  
Results for Experiment 2



Note. Standard deviation ( $SD$ ) for no-distraction (red), encoding-distraction (green), and delay-distraction (blue) conditions. Grey dots represent single participant's data. Black dots indicate condition means and error bars denote 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

\*\*\* $p < .001$ .

performance, it implies that the distraction interference is mainly caused by the absence of processing beforehand. Second, if the full-distraction condition also causes a significant decrease on mnemonic fidelity, the distraction effect should be explained by the distractors' appearance during the delay.

## Method

### Participants

Twenty-four volunteers (ages 18–25, 15 male, 9 female) participated in Experiment 3 (the same as previous studies). All participants had a normal or corrected-to-normal vision as well as normal color vision, gave informed written consent, and were remunerated ¥60–70 (approximately US\$8–10) for their time.

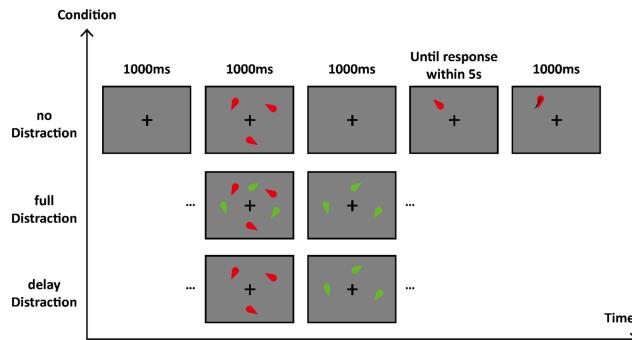
### Procedure

The no-distraction condition and the delay-distraction condition were the same as those in Experiment 1. The full-distraction condition was accomplished by presenting the distractors from the beginning of the encoding period and to the end of the delay period (see Figure 4).

## Results and Discussion

We tested whether the distraction effect would disappear when distractors were presented throughout the two stages (i.e., full-distraction condition). The one-way repeated-measures ANOVA revealed a significant main effect of conditions on  $SD$ ,  $F(2, 46) =$

**Figure 4**  
*Schematic Illustration of the No-Distraction, Full-Distraction, and Delay-Distraction Conditions in Experiment 3*

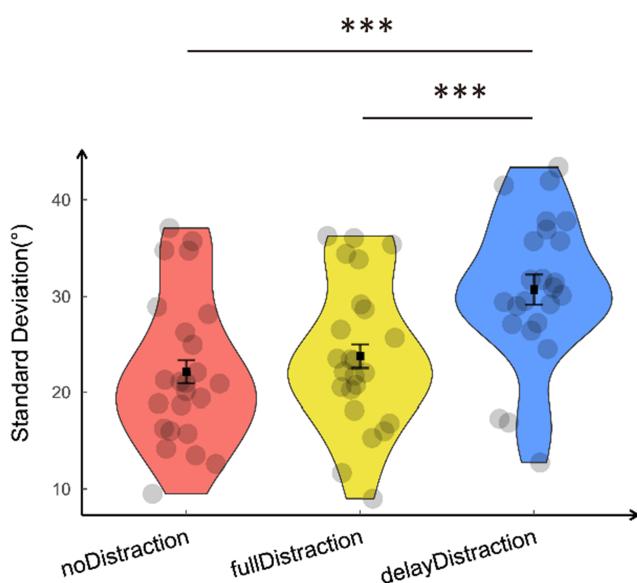


Note. The procedure was similar to Experiment 1 except for the full-distraction condition, where distractors were presented both at encoding and during the delay period. See the online article for the color version of this figure.

49.32,  $p < .001$ ,  $\eta_p^2 = 0.68$ . Post hoc  $t$  tests with Bonferroni correction found that  $SD$  was significantly higher in delay-distraction condition compared with the other two conditions,  $t_{(no-delay)}(23) = -8.81$ ,  $p < .001$ , Cohen's  $d = -2.70$ ;  $t_{(full-delay)}(23) = -6.97$ ,  $p < .001$ , Cohen's  $d = -2.19$ . However, no difference was found between full-distraction condition and no-distraction condition,  $t_{(no-full)}(23) = -2.13$ ,  $p = .13$ , Cohen's  $d = -0.51$  (see Figure 5).

To summarize, we found that as long as distractors were processed by participants at encoding, the distraction effect can be eliminated.

**Figure 5**  
*Results for Experiment 3*



Note. Standard deviation ( $SD$ ) for no-distraction (red), full-distraction (yellow), and delay-distraction (blue) conditions. Grey dots represent single participant's data. Black dots indicate condition means and error bars denote 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

\*\*\* $p < .001$ .

These results suggested that the distraction effect during the delay was mainly caused by absence of processing at encoding.

## Experiment 4

In this experiment, we aimed to investigate whether the success of distractor resistance was related to the efficiency of suppression distractors. That is, whether decreasing the distinguishability between targets and distractors at the beginning would lead to resistance failure. To test this possibility, we used a pre-cue instead of a prominent color difference to distinguish the targets and distractors. Then, we compared the memory performance of same-color conditions with the performance of original different-color conditions. Therefore, we had seven conditions in total: no-distraction, encode-distraction (different-color), delay-distraction (different-color), full-distraction (different-color), encode-distraction (same-color), delay-distraction (same-color), and full-distraction (same-color) conditions. Moreover, we also replicated these findings with the full-distraction (different-color) and full-distraction (same-color) conditions in Experiment 4S. See the online supplemental material for details.

## Method

### Participants

Twenty-four volunteers (ages 18–25, 12 male, 12 female) participated in Experiment 4 (the same as previous studies). All participants had a normal or corrected-to-normal vision as well as normal color vision, gave informed written consent, and were remunerated ¥60–70 (approximately US\$8–10) for their time.

## Procedure

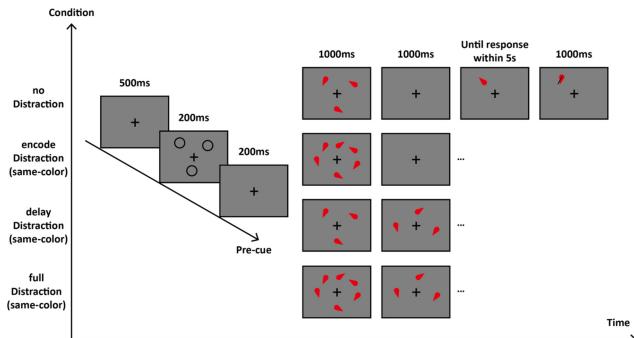
To diminish the efficiency of distinguishing targets from distractors, we changed the distractors to match the same color as the targets and instead used a pre-cue to distinguish them. In the new full-distraction (same color) condition, the participants first saw three circles ( $2.1^\circ \times 2.1^\circ$ ) briefly (200 ms) presented on the screen to indicate the targets' location. Then, six teardrops with the same color (red or green) appeared on the screen for 1,000 ms. Importantly, the three teardrops sharing the same location with the three circles were targets that participants needed to memorize and were tested on at the end of the trial (see Figure 6). To control the differences among conditions, we also added a pre-cue in the no-distraction, encode-distraction (different-color), delay-distraction (different-color), and full-distraction (different-color) conditions. Literally, the pre-cue in these conditions was redundant, since participants could use color to distinguish them.

## Results and Discussion

We tested whether the distraction effect would revive when it is more difficult to distinguish the targets from distractors. The one-way repeated-measures ANOVA revealed a significant main effect of conditions on  $SD$ ,  $F(6, 23) = 38.91$ ,  $p < .001$ ,  $\eta_p^2 = 0.91$ . With post hoc  $t$  tests and Bonferroni correction, we first replicated the dissociated distraction effect found in previous experiments. That is, we found  $SD$  was significantly higher in delay-distraction (different color) condition compared to no-distraction, encode-distraction (different color), and full-distraction (different color) conditions,  $t_{(no-delay\_diff)}(23) = -$

**Figure 6**

*Schematic Illustration of the No-Distraction, Encode-Distraction (Same-Color), Delay-Distraction (Same-Color), and Full-Distraction (Same Color) Conditions in Experiments 4*



**Note.** We replaced the prominent color difference with a pre-cue to indicate the targets. To control the consistency among all conditions, the pre-cue was also added to the no-distraction, encode-distraction (different-color), delay-distraction (different-color), and full-distraction (different-color) conditions (not shown in this figure). See the online article for the color version of this figure.

8.50,  $p < .001$ , Cohen's  $d = -1.21$ ;  $t_{(encode\_diff-delay\_diff)}$  (23) = −11.00,  $p < .001$ , Cohen's  $d = -1.27$ ;  $t_{(full\_diff-delay\_diff)}$  (23) = −11.19,  $p < .001$ , Cohen's  $d = -1.30$ . No difference was found between no-distraction, encode-distraction (different color), and full-distraction (different color) conditions,  $t_{(no\_encode\_diff)}$  (23) = 0.72,  $p = 1.00$ , Cohen's  $d = 0.07$ ;  $t_{(no\_full\_diff)}$  (23) = 0.91,  $p = 1.00$ , Cohen's  $d = 0.09$ ;  $t_{(encode\_full\_diff)}$  (23) = 0.23,  $p = 1.00$ , Cohen's  $d = 0.03$ . However, when we used pre-cues to distinguish between targets and distractors, all same-color conditions had significant higher  $SD$  than the no-distraction condition,  $t_{(no\_encode\_same)}$  (23) = −7.77,  $p < .001$ , Cohen's  $d = -1.64$ ;  $t_{(no\_delay\_same)}$  (23) = −10.00,  $p < .001$ , Cohen's  $d = -2.39$ ;  $t_{(no\_full\_same)}$  (23) = −8.09,  $p < .001$ , Cohen's  $d = -1.98$ . Moreover, we also found the same-color conditions had significant higher  $SD$  than those corresponding different-color conditions,  $t_{(encode\_diff-encode\_same)}$  (23) = −8.33,  $p < .001$ , Cohen's  $d = -1.70$ ;  $t_{(delay\_diff-delay\_same)}$  (23) = −5.08,  $p < .001$ , Cohen's  $d = -1.18$ ;  $t_{(full\_diff-full\_same)}$  (23) = −8.19,  $p < .001$ , Cohen's  $d = -2.07$  (see Figure 7).

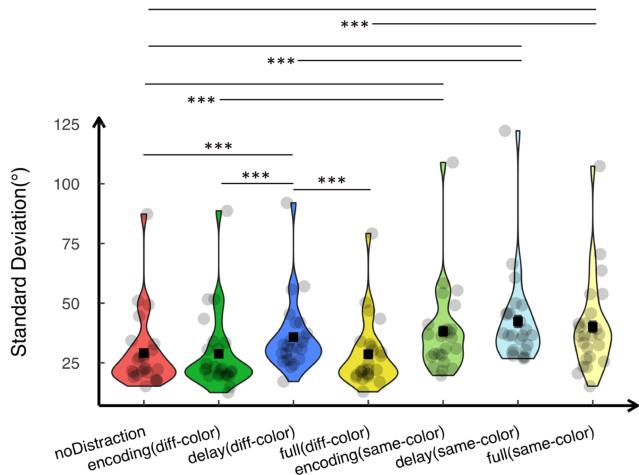
These results indicated the effect of successful suppression on distractors across both encoding and delay periods. Combined with results in Experiment 3, we concluded that the efficiency of differentiating task relevance among multiple items influenced the limited resource allocation in WM.

## Experiment 5

In this experiment, we aimed to investigate whether the success of distractor resistance in full-distraction condition was because of the suppression of distractors' locations, or orientations, or both. We manipulated the orientations or locations of distractors during the delay and tested whether they would cause interference. These manipulations lead to three more conditions: full-distraction (orientation change), full-distraction (location change), full-distraction (both change) conditions. There were three possible predictions. First, if the distractor resistance was only caused by suppression of

**Figure 7**

*Results for Experiment 4*



**Note.** Standard deviation ( $SD$ ) for no-distraction (red), encode-distraction (different-color) (green), delay-distraction (different color) (blue), full-distraction (different color) (yellow), encode-distraction (same-color) (light green), delay-distraction (same color) (light blue), and full-distraction (same color) (light yellow) conditions. Grey dots represent single participant's data. Black dots indicate condition means and error bars denote 95% CI within-subjects (Cousineau, 2005). See the online article for the color version of this figure.

\*\*\* $p < .001$ .

orientations, the full-distraction (location change) condition won't cause interference. Second, if the distractor resistance was only caused by suppression of locations, the full-distraction (orientation change) condition won't cause interference. Last, if the distractor resistance was related to both changes of orientations and locations, all three conditions would lead to interference, where full-distraction (both change) condition would have the biggest impact.

## Method

### Participants

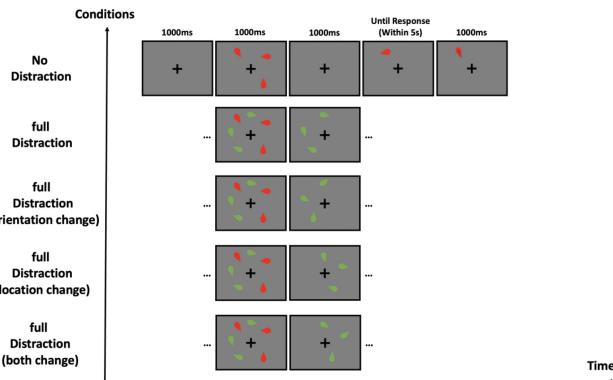
Twenty-eight different volunteers (ages 18–25) participated in Experiment 5 (9 male, 19 female). All participants had a normal or corrected-to-normal vision as well as normal color vision, gave informed written consent, and were remunerated ¥60–70 (approximately US\$8–10) for their time.

### Procedure

In this experiment, both the no-distraction and full-distraction conditions remained the same as Experiment 3. In addition, we added another three conditions to manipulate the orientations and locations of distractors during the delay in the full-distraction condition. For the full-distraction (orientation change) condition, only the orientations of distractors would change randomly when the delay period started. For the full-distraction (location change) condition, only the location of distractors would change randomly while keeping the orientations the same as before. For the full-distraction (both change) condition, both the orientations and locations would change randomly (see Figure 8).

**Figure 8**

*Schematic Illustration of the No-Distraction, Full-Distraction, Full-Distraction (Orientation Change), Full-Distraction (Location Change), and Full-Distraction (Both Change) Conditions in Experiment 5*



Note. Three added conditions were similar to the full-distraction condition, except for the orientation or location changed for distractors during the delay. See the online article for the color version of this figure.

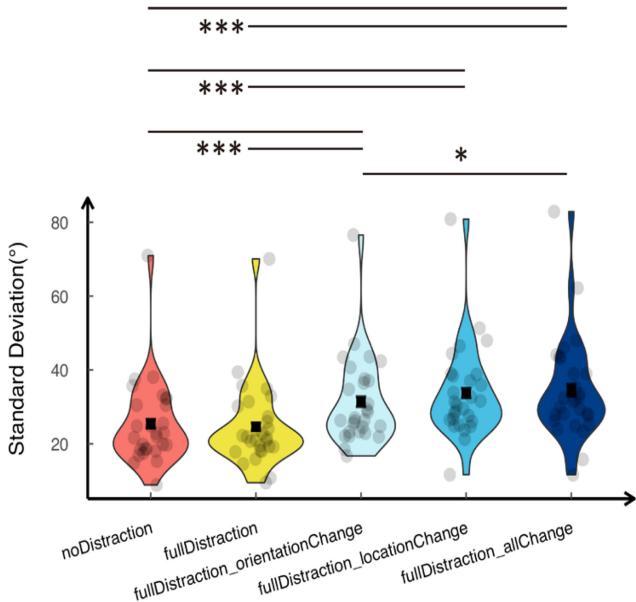
## Results and Discussion

We tested whether the manipulations of orientations or locations on distractors during the delay would cause interference for full-distraction conditions. The one-way repeated-measures ANOVA revealed a significant main effect of conditions on  $SD$ ,  $F(4, 108) = 52.83$ ,  $p < .001$ ,  $\eta_p^2 = 0.66$ . Post hoc  $t$  tests with Bonferroni correction found that  $SD$  had no difference between no-distraction and full-distraction conditions,  $t_{(\text{no-full})} (23) = 1.54$ ,  $p = 1.00$ , Cohen's  $d = 0.24$ , replicating the discoveries in Experiment 3. More importantly, we found a significant distraction effect for all the added three full-distraction conditions,  $t_{(\text{no-orientation change})} (27) = -7.11$ ,  $p < .001$ , Cohen's  $d = -1.77$ ;  $t_{(\text{no-location change})} (27) = -9.47$ ,  $p < .001$ , Cohen's  $d = -2.46$ ;  $t_{(\text{no-both change})} (27) = -8.07$ ,  $p < .001$ , Cohen's  $d = -2.69$ . Besides, the added three conditions all had a worse memory performance than the original full-distraction condition,  $t_{(\text{full-orientation change})} (27) = -8.52$ ,  $p < .001$ , Cohen's  $d = -1.42$ ;  $t_{(\text{full-location change})} (27) = -11.79$ ,  $p < .001$ , Cohen's  $d = -1.91$ ;  $t_{(\text{full-both change})} (27) = -9.51$ ,  $p < .001$ , Cohen's  $d = -2.07$ . Moreover, we also found a significant worse memory performance for the full-distraction (both change) condition than the full-distraction (orientation change) condition but not the other comparisons,  $t_{(\text{orientation change-both change})} (27) = -3.08$ ,  $p = .05$ , Cohen's  $d = 0.93$ ;  $t_{(\text{orientation change - location change})} (27) = -2.28$ ,  $p = .31$ , Cohen's  $d = 0.69$ ;  $t_{(\text{location change-both change})} (27) = 0.93$ ,  $p = 1.00$ , Cohen's  $d = 0.24$  (see Figure 9).

These results indicated that the success of distractor resistance in full-distraction condition was due to the suppression of both distractors' orientations and locations at encoding, where locations might play a more important role. It suggested that pre-encoding distractors at encoding allowed people allocate less resources on distractors, which led to successful distractor resistance.

**Figure 9**

*Results for Experiment 5*



Note. Standard deviation ( $SD$ ) for no-distraction (red), full-distraction (yellow), full-distraction (orientation change) (light blue), full-distraction (location change) (middle blue), full-distraction (both change) (dark blue) conditions. Grey dots represent single participant's data. Black dots indicate condition mean and error bars denote 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

\* $p < .05$ . \*\*\* $p < .001$ .

## Experiment 6

According to the preceding experiments, we have found a robust distraction effect when the distractors appeared during the WM delay period. Since the distractors always appeared at a different location than the targets, we cannot tell whether the distraction-related interference should occupy extra spatial resource. To answer the question, we added a new delay-distraction condition, where the distractors shared the same locations with the targets. There are two possible predictions. First, if the distraction-related effect is related to extra space occupied by distractors during the delay, the delay (same location) condition won't induce interference. On the contrary, if distractor itself is the determination of the interference, then the delay (same location) condition should cause similar interference as the delay (different location) condition. To generalize our discoveries and exclude an alternative explanation that the distraction-induced interference might be caused by the abrupt appearance of distractors, we used two different ways of presenting distractors during the delay. One is the same as Experiment 1 (Experiment 6a), while the other presented distractors gradually during the delay without a sudden onset (Experiment 6b).

## Method

### Participants

Twenty-four different volunteers (ages 18–25) participated in Experiment 6a (6 male, 18 female) and Experiment 6b (8 male, 16

female). All participants had a normal or corrected-to-normal vision as well as normal color vision, gave informed written consent, and were remunerated ¥60–70 (approximately US\$8–10) for their time.

### Procedure

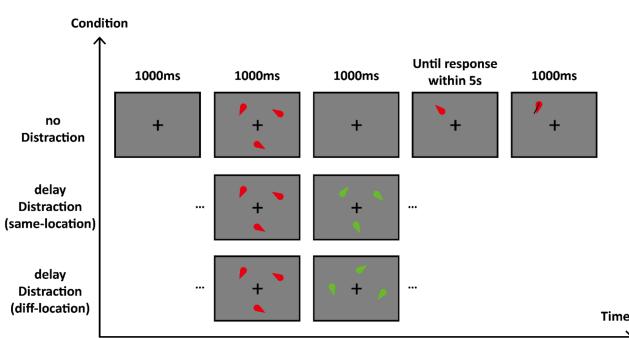
The no-distraction condition and the delay-distraction (different location) condition were the same as those in Experiment 1. The newly added delay-distraction (same location) condition was also similar to the different-location condition, except for the shared locations with targets during the delay (see Figure 10). Notably, in Experiment 6b, we presented the distractors from completely transparent to completely opaque within 500 ms to implement the gradual effect.

### Results and Discussion

We tested whether the distraction effect was based on object or location by comparing the performances between conditions sharing the same locations between the targets and the distractors or not. For Experiment 6a, the one-way repeated-measures ANOVA revealed a significant main effect of conditions on  $SD$ ,  $F(1.97, 45.38) = 60.81$ ,  $p < .001$ ,  $\eta_p^2 = 0.73$ . Post hoc  $t$  tests with Bonferroni correction found that  $SD$  were significantly lower in no-distraction condition compared with the other two conditions,  $t_{(\text{no-same location})} (23) = -9.69$ ,  $p < .001$ , Cohen's  $d = -2.95$ ;  $t_{(\text{no-different location})} (23) = -9.08$ ,  $p < .001$ , Cohen's  $d = -2.51$ . However, no difference was found between delay-distraction (same location) condition and delay-distraction (different location) condition,  $t_{(\text{same-different location})} (23) = 1.54$ ,  $p = .41$ , Cohen's  $d = 0.44$  (see Figure 11).

For Experiment 6b, the one-way repeated-measures ANOVA revealed a significant main effect of conditions on  $SD$ ,  $F(2, 46) = 50.46$ ,  $p < .001$ ,  $\eta_p^2 = 0.69$ . Post hoc  $t$  tests with Bonferroni correction found that  $SD$  were significantly lower in no-distraction condition compared with the other two conditions,  $t_{(\text{no-same location})} (23) = -8.78$ ,  $p < .001$ , Cohen's  $d = -2.76$ ;  $t_{(\text{no-different location})} (23) = -8.84$ ,  $p < .001$ , Cohen's  $d = -2.15$ . However, no difference was found between delay-distraction (same location) condition and delay-distraction (different location) condition,  $t_{(\text{same-different location})} (23) = 2.00$ ,  $p = .17$ , Cohen's  $d = 0.61$  (see Figure 12).

**Figure 10**  
Schematic Illustration of the No-Distraction, Delay-Distraction (Same Location), and Delay-Distraction (Different Location) Conditions in Experiments 6a and 6b



*Note.* The procedure was similar to Experiment 1 except for the delay-distraction (same location) condition, where distractors were presented at the same location as the targets during the delay. See the online article for the color version of this figure.

(23) =  $-8.84$ ,  $p < .001$ , Cohen's  $d = -2.15$ . However, no difference was found between delay-distraction (same location) condition and delay-distraction (different location) condition,  $t_{(\text{same-different location})} (23) = 2.00$ ,  $p = .17$ , Cohen's  $d = 0.61$  (see Figure 12).

These findings demonstrated that the distraction effect during the maintenance did not rely on extra occupation of spatial resources, as both the different-location and same-location condition caused the same cost in WM fidelity. Moreover, different ways of presenting distractors could exclude another alternative explanation that the distraction-induced interference in the delay-distraction condition was caused by the abrupt appearance of the distractors.

### Understand the Dissociated Distraction Effect

The primary discovery was that distractors' impact on memory performance depended on the stage of memory process (encode or delay) when they were introduced. We also found some marginal conditions modulating such effect, such as difficulty to distinguish targets from distractors and feature changes during the delay. These results were in accordance with our speculations: task relevance and visual uncertainty drove the limited resource allocation in visual WM tasks. Specifically, the task relevance was a gating mechanism. When targets and distractors were presented simultaneously with equal uncertainty, input from distractors was effectively attenuated to avoid distraction cost. If the uncertainty of distractors were strengthened, such as first appearing during the delay or changing features during the delay, more resources would go to them, lowering memory performance for targets. In general, the competing power of a stimulus could be a function of its task relevance and uncertainty. However, details about how such function worked was not clear and the potential differences between individuals were unlikely to be captured. In the next section, we proposed a computational Bayesian model, made quantitative predictions about the roles of task relevance and visual uncertainty, and fit out a specific subject's characteristics exhibited in the task.

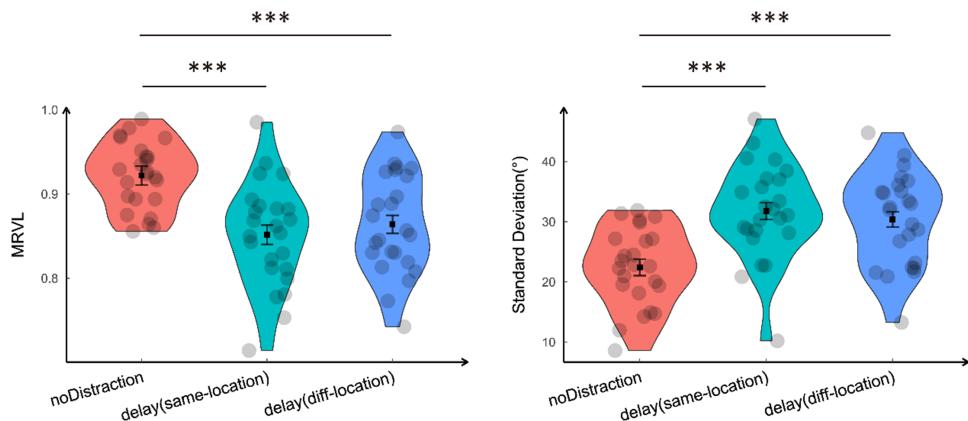
### The Computational Model

The current model aimed to explain how WM allocated its limited resources when being confronted with perceptual distractors at different stages. First, we reformulated the continuous recall task as probabilistic inference in a generative model. Then, we implemented the two main factors, task relevance and visual uncertainty, which were supposed to modulate the resource allocation between targets and distractors across the encoding and delay periods. Finally, we made explicit predictions about the interference effect of different conditions by running the model. Predictions were then compared with human performance and predictions from alternative models to evaluate model performance.

### Continuous Recall Task as Bayesian Inference

First of all, we illustrated how observers performed the WM task without distraction interference under the Bayesian framework (Brady & Tenenbaum, 2013). Our modeling proceeds in two stages, mirroring the two stages of the continuous recall task: how observers encode the study display and how they recall the target's information once they have access to the test display. For simplicity, assuming

**Figure 11**  
Results for Experiment 6a



*Note.* Standard deviation ( $SD$ ) for no-distraction (red), delay-distraction (same location) (cyan), and delay-distraction (blue) conditions. Grey dots represent single participant's data. Black dots indicate condition mean and error bars denote 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

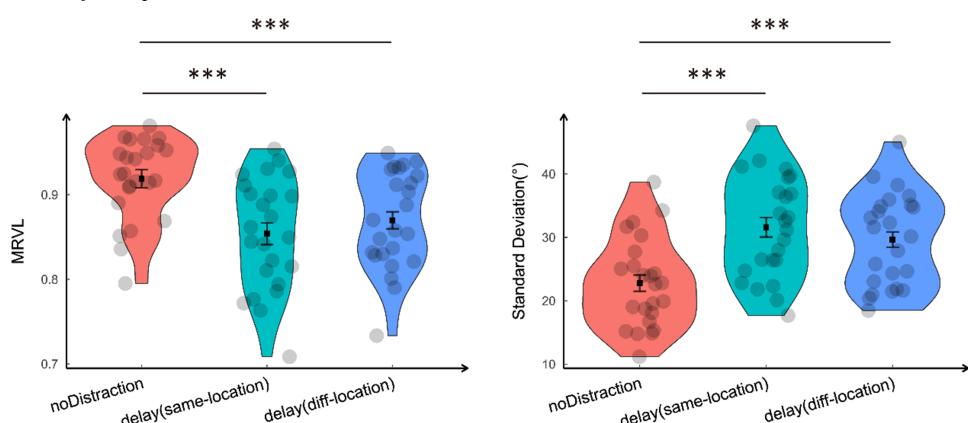
\*\*\* $p < .001$ .

there is only one item on the display, whose orientation should be remembered according to the task demand (see Figure 13). Before encountering the display, observers had no information about the item's orientation, corresponding to a uniform distribution. By encoding the display, observers update their belief of the stimulus' orientation, forming up a von-mises distribution to represent perceived information. The peak point of this distribution will be around the actual orientation of the stimulus, meaning here is the most "likely" spot. This "sweet spot" may slightly drift away from trial to trial, but the grand expectation surely fits the very center (shown in the figure). The declining shoulders represent the declining probability that you sample a value remotely away from the center point out of the distribution. According to the Bayes' rule, what

people do in the test phase is to infer the displayed orientation from this memory distribution and report it to finish recall tasks.

To implement such Bayesian inference and probabilistic calculation, the computation is hypothesized as a sampling process (Sanborn & Chater, 2016, 2017). Sampling is a statistical concept, and it has similar meaning in cognitive modeling: to establish a representative estimation of a population by retrieving a sample out of it. Seeing WM as sampling processes is based on the idea that information is never perfectly stored in or retrieved from memory, but with certain accuracy loss, which complies with probabilistic rules. Just like what happens in statistical practice, larger samples always indicate higher accuracy. Under sampling perspective, limited WM resource is equivalent to limited sampling capacity, and more

**Figure 12**  
Results for Experiment 6b

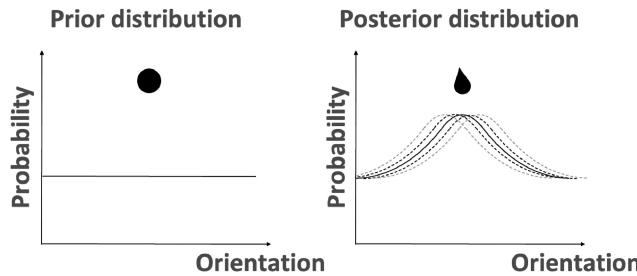


*Note.* Standard deviation ( $SD$ ) for no-distraction (red), delay-distraction (same location) (cyan), and delay-distraction (blue) conditions. Grey dots represent single participant's data. Black dots indicate condition mean and error bars denote 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

\*\*\* $p < .001$ .

**Figure 13**

Updated Probabilistic Distributions According to the Bayes' Rule



*Note.* Assuming there is only one item on the display, observers update their belief from uniform prior distribution into normal posterior distribution centered on the true value (trial-to-trial deviations shown in dashed lines).

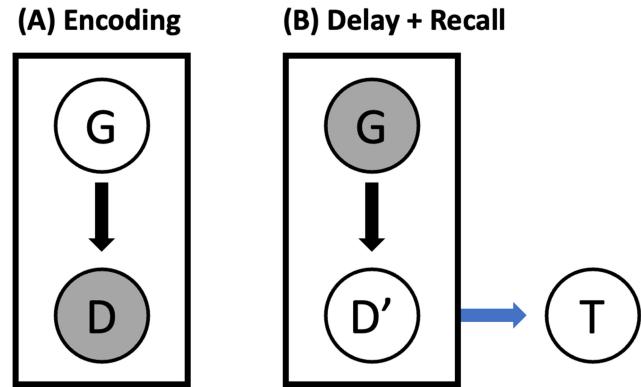
allocated resources correspond to more sampling numbers, or a larger sample in statistical terms.

In general, there are two sampling processes across two periods of the WM task. The first sampling happens during the encoding period, where observers sample external sensory input and update their memory presentations from the prior uniform distribution to the posterior normal distribution. The allocated sampling number influences the posterior probabilistic distribution of targets. With more sampling numbers, the posterior distribution of targets will have higher kurtosis (i.e., more centralized). The second sampling happens during the delay and recall periods, where observers sample from distribution established in encoding stage and finally report it. The allocated sampling number of the second phase and the kurtosis of the distribution relying on the sampling number of the first period jointly determine the final averaged sampling error. Our model uses this averaged sampling error to predict participants' memory precision in the behavioral experiments. Notably, since we did not distinguish different targets or distractors in our experiments, our model only considered the resource allocation between targets and distractors, while assuming the same within these three targets and distractors. In the next section, we would describe in detail factors that might influence the resource allocation during the sampling process. We now focus on how to formalize the two-stage WM model in mathematical expressions.

The graphical model representation of the two-stage WM model (shown in Figure 14) specifies how the stimuli were initially encoded into memory and finally retrieved to guide recalling tasks. At the first stage, participants observed the study display ( $D$ ), and they used this display to inversely infer the probabilistic distributions that generate the observed display items ( $G$ ), corresponding to their internal representations. Thus, our model infers the posterior distributions  $p(G_k|D_k)$  of the variable  $G_k$  ( $k = 1, 2, \dots, n$ ,  $n$  denotes the number of items on the screen,  $G_k$  denotes the representative value of the  $k$ th item in the WM, and  $D_k$  denotes the observed value of the  $k$ th item on the display). Our model assumes that the prior distribution follows a uniform distribution on the range of items' possible value, and the likelihood distribution follows a normal distribution centered on  $G_k$  with a variance (captured by the likelihood noise,  $I_{sd}$ ). See formula (1). Observers update their beliefs by sampling according to this Bayes' rule. Therefore, more

**Figure 14**

Graphical Model Notation for the Two-Stage WM Model at Encoding (A) and During the Delay and Recall Periods (B)



*Note.* Shaded nodes are observed. The solid black arrows correspond to our model of how a display is generated in our mind. The blue arrow corresponds to the process observers sampled the probed item from the generated display from their updated memory representations. At encoding, we observe the study display ( $D$ ), and we use this to infer the probabilistic distributions that may have generated the observed display items ( $G$ ). During the delay and recall periods, we have access to the generative display we encoded in memory ( $G$ ). We thus infer what the study display looked like ( $D'$ ) and generate samples ( $t$ ) to complete the continuous recall task. See the online article for the color version of this figure.

sampling number results in a more centralized normal posterior distribution.

$$\begin{aligned} p(G_k|D_k) &\propto p(G_k)p(D_k|G_k) \\ D_k|G_k &\sim N(G_k, \in) \end{aligned} \quad (1)$$

$$t_k^{(1)}, t_k^{(2)}, \dots, t_k^{(nk)} \sim p(D'_k). \quad (2)$$

At the second stage, observers maintain the updated posterior distributions  $p(G_k|D_k)$  from the first stage in their memory and infer what the study display ( $D_k^{(n)}$ ) looks like. When probe shows up, observers complete the recalling task by generating the target's samples  $t_k^{(i)}$  with the inferred distribution  $p(D_k^{(n)})$ . See formula (2). Finally, the circular mean ( $M$ ) of these samples denotes the report value, and the  $SD$  denotes the memory precision, which is the square root of the variance. Therefore, more sampling number is associated with less  $SD$ , corresponding to a better task performance in the continuous recall task.

### Resource Allocation Based on Two Factors

When distractors were introduced during the encoding or delay periods, the sampling process for targets would be influenced. Here, we described two main factors that might influence the resource allocation between targets and distractors across both two periods.

The first factor is *task relevance*, which represents the priority of an item required to perform well in the current task. Task relevance is determined by both top-down task goals and bottom-up stimulus saliency. This idea is analogical to the biased-competition framework (Bundesen et al., 2011; Desimone & Duncan, 1995) and the priority map theory (Fecteau & Munoz, 2006; Serences & Yantis, 2006;

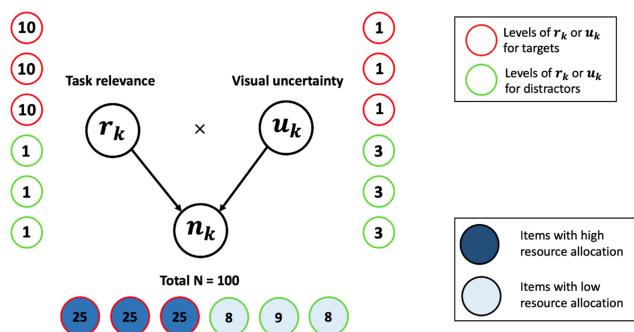
Sprague et al., 2018), which assumes that top-down modulations bias bottom-up saliency signals, resulting in a priority map coding for behavioral relevance. The second factor is *visual uncertainty*, defined as the degree of unfamiliarity with each item. Decreasing uncertainty about future states drives our attention and behaviors. For example, previous research found that uncertainty in WM can promote resampling of the environment to reduce the uncertainty (Sullivan et al., 2012). In neuroscience research, sensory novelty is known to enhance neural responses throughout visual, frontal, and temporal areas (Gottlieb et al., 2013; Ranganath & Rainer, 2003) and activate reward-responsive dopaminergic areas (Berlyne, 1950; Düzel et al., 2010; Lehman & Stanley, 2011).

In our model, we let  $r_k$  denote the *task relevance* and  $u_k$  denote the *visual uncertainty* of the  $k$ th displayed item. These two factors determine the importance of the stimuli, which can be computed by the product of  $r_k$  and  $u_k$ . Our model defines a constant  $N$  as the total number of samples at each stage, which represents the total limited cognitive resources in the WM. Therefore, the sampling number  $n_k$  for the  $k$ th item can be computed as follows (see Figure 15 for an example):

$$n_k = \frac{r_k u_k}{\sum_i r_i u_i} N. \quad (3)$$

Task relevance  $r_k$  is mainly determined by the current task requirement, like whether the item is task relevant or irrelevant. For example, in our current study, participants were asked to memorize the orientations of the red teardrops while ignoring the green ones. Intuitively, the  $r_k$  of red teardrops (targets) should be much higher than green teardrops (distractors). Besides, it is also influenced by stimulus saliency. For instance, in Experiment 4, it is harder to separate targets from distractors by using pre-cue than prominent color difference. Therefore, the relative task relevant ratio between targets and distractors would be reduced in the same-color conditions compared to the original different-color conditions.

**Figure 15**  
Two Factors ( $r_k$  and  $u_k$ ) That Influence the Limited Resource Allocation ( $n_k$ )



**Note.** First, we assumed a limited cognitive resource allocated among stimuli ( $N = 100$  in total). Task relevance ( $r_k$ ) and visual uncertainty ( $u_k$ ) are multiplied for each item to determine its allocated resources ( $n_k$ ). The total resources ( $N$ ) are allocated among items based on their allocated resources ( $n_k$ ). In this example, targets (red circles) have higher task relevance than distractors (green circles) (left column), while distractors have higher visual uncertainty than targets (right column). Combining these two factors, the limited resources are allocated between targets and distractors. See the online article for the color version of this figure.

Visual uncertainty  $u_k$  comes from two sources. The first source  $u_o$  comes from the item's orientation, which can be calculated by the information entropy or variance of its probabilistic distribution ( $G_k$ ). See formula (4). Generally, if the distribution of a certain item is more centralized, its  $u_o$  would be smaller. The second source  $u_l$  comes from the item's location. If item's location changes during the delay period (Experiment 5), it will increase the amount of information. To simplify the calculation, we assume the process of choosing new locations for distractors follows a uniform distribution. According to our experiment settings, the location-induced entropy is approximately calculated by randomly selecting three locations for distractors within a fixed size 2D space, given six locations have been occupied during the encoding. The fixed size is a rectangle ( $4.0^\circ \times 6.0^\circ$ ) and each possible location is a circle with a diameter of  $1.6^\circ$ . Therefore, the total visual uncertainty  $u_k$  can be calculated by combining  $u_o$  and  $u_l$  together. See formula (5). When items' locations change during the delay,  $\lambda$  equals to 1. When items' locations remain the same during the whole WM task,  $\lambda$  equals to 0.

$$u_o = - \int \log(p(G_k)) p(G_k) dG_k \quad (4)$$

$$u_k = u_o + u_l. \quad (5)$$

## Model Predictions

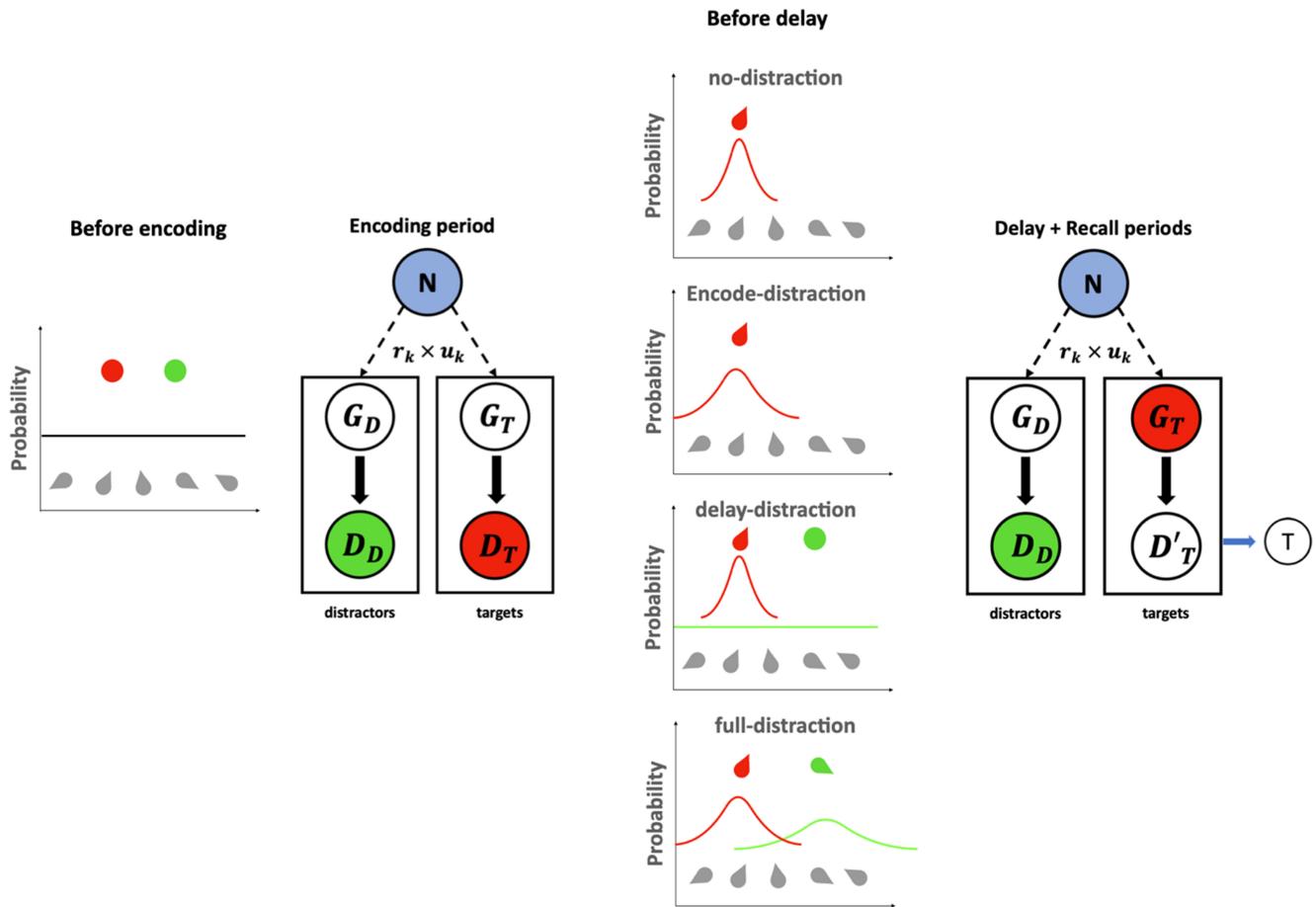
Based on our two-stage WM model and the resource allocation principles, we make specific predictions for all conditions in Experiments 1–5—no-distraction, encode-distraction, delay-distraction, full-distraction, full-distraction (same color), full-distraction (orientation change), full-distraction (location change), and full-distraction (both change) conditions. As conditions in Experiments 6a and 6b had the same resource allocation processes as Experiments 1–3, we did not include them in this part. Notably, our model assumed both task relevance ( $r_k$ ) and visual uncertainty ( $u_k$ ) would influence resource allocation across both two stages. The second-stage samples from the sampling results at the first stage, whose statistics predict the final performance. Therefore, both sampling processes at two stages had impact on the final recall performance (see Figure 16). Here, we illustrate our model's predictions for all conditions as follows.

For the encoding-distraction condition, the limited sampling number should be divided between targets and distractors only at encoding. Since they are both encountered for the first time,  $u_{target}$  and  $u_{distractor}$  are both calculated based on the uniform prior distribution. Therefore,  $r_k$  plays the main role to allocate most resources to targets, leading to an effective resistance to distractors.

According to the calculation formula of visual uncertainty ( $u_k$ ), it can be computed based on the current probabilistic distribution at different stages. If observers encounter the stimuli for the first time,  $u_k$  should be computed based on the uniform prior distribution. After observers process the item, its distribution would be updated to the normal posterior distribution, leading to a reduction of  $u_k$ . Therefore, for the delay-distraction condition, even though  $r_k$  is still higher for targets than distractors, a much higher  $u_k$  for distractors than targets result in most resources being allocated to distractors during the delay. Our model assumes this is the main reason driving

**Figure 16**

Illustration of the Resource Allocation and Probabilistic Distribution at Two Stages



**Note.**  $N$  represents the total sampling number for all items at each stage, being divided between targets (with  $t$  subscript) and distractors (with  $d$  subscript). Shaded notes are observed, where red represents targets while green represents distractors. The resource allocation is determined by task relevance ( $r_k$ ) and visual uncertainty ( $u_k$ ). Before the encoding period, probabilistic distributions for both targets and distractors are uniform. After processing them, probabilistic distributions change into normal distributions centered on the true value with different variances. More allocated resources in the first stage would result in a more centered distribution with less variance before the delay period. For both the no-distraction and encode-distraction conditions, no distractors' distributions are needed, and the distributions of targets are more variant for the encode-distraction condition. For the delay distraction condition, the distributions of distractors are uniform. For the full distraction condition, the normal distributions for distractors have more variances than targets mainly because of the different task relevance ( $r_k$ ) on the first stage. The final sampling results are based on the distributions before the delay, whose statistics determine the final memory performance. See the online article for the color version of this figure.

the significant distraction interference in the delay-distraction condition.

For the full-distraction condition, both the encoding and delay stages involve the resource allocation between targets and distractors. At encoding, the main determining factor is  $r_k$ , similar to the encoding-distraction condition. Thus, most sampling numbers will be allocated to targets than distractors to protect targets from interference. Even though the posterior distributions of distractors are nosier than targets, they won't occupy many resources during the delay, since their features remain the same. Therefore, our model predicts the full-distraction condition can still resist distraction successfully.

However, when distractors' features are changed during the delay as in Experiment 5, the posterior distribution of distractors would be

much nosier than targets, leading to more resource's consumption. Therefore, our model predicts significant distraction interference for those variant full-distraction conditions. Specifically, for the full-distraction (orientation change) condition, we assume the distribution of distractors becomes uniform again instead of the posterior normal distribution, so that  $u_o$  is added. For the full-distraction (location change) condition,  $u_{distractor}$  increases because of the added  $u_l$ . For the full-distraction (both change) condition, both  $u_o$  and  $u_l$  lead to the most severe distraction-induced interference.

Finally, we included another free parameter  $r_k$  in our model to understand the effects of same-color conditions in Experiment 4. According to the definition of  $r_k$ , using pre-cues to decrease the distinguishability between targets and distractors will decrease our subjective judgement of the task relevance. Our model assumes

this influence persists across both periods, leading to significant distraction interference for all same-color conditions.

### Model Fitting

For model fitting, we set the total sampling number ( $N$ ) as 100 to simulate the limited capacity of WM at each stage. In general, we fit three free parameters for each participant. The task relevance ratio ( $r$ ) represents the relative task relevance between targets and distractors. A greater value of  $r$  means observers weight targets higher than distractors. Notably, the task relevance ( $r$ ) remains the same during both two stages. The likelihood noise ( $l_{sd}$ ) and sampling noise ( $s_{sd}$ ) represent the noise level in the Bayesian inference process and sampling from the distribution process respectively. A higher level of noise means higher variability during these two processes. Besides, the visual uncertainty ratio ( $u$ ) represents the relative uncertainty between targets and distractors, which is not a free parameter needed for estimating. Both visual uncertainties for targets and distractors can be computed by the information entropy of its probabilistic distribution, see formula (4). Therefore, combining parameter  $r$  and  $u$ , different sampling numbers were assigned to targets at two stages for different conditions.

According to our generative model, we sampled targets at both stages whose sampling number was determined by the rules we have proposed. Then, we calculated the sampling means of the sample results at the final stage. We repeated this sampling process 100 times and calculated the  $SD$  of all these sampling means. This simulated  $SD$  will be used to represent the model estimated memory precision and being compared with the  $SD$  in our behavioral experiments. To find the best-fitting parameters for each participant, we set a range of possible parameter spaces and simulated  $SD$  with each parameter combination. Then, we calculated the mean-squared error (MSE) between the simulated  $SD$  and the real  $SD$  and found the parameters, which minimized MSE for each participant. We performed several pilot simulations to find a reasonable parameter space. For Experiments 1–3, the parameter range for  $r$  was from 1 to 20 in step of 1.  $l_{sd}$  and  $s_{sd}$  were set from  $\pi$  divided by 1–20 in step of 2. Therefore, there were 2,000 parameter combinations for Experiments 1–3 and Experiment 5. For Experiment 4, another task relevance ration ( $r'$ ) was needed to be estimated for those same-color conditions. We also conducted parameter recovery analysis to verify the reliability of such setting (see the online [supplemental materials](#) for details). According to the model assumption, decreasing the task relevance can explain the revive of distraction effect in these same-color conditions. Therefore, we set the parameter range for  $r'$  from 1 to 10 in step of 1. According to the parameter selecting results from the previous results, we narrowed down the parameter range for  $l_{sd}$  and  $s_{sd}$  into  $\pi$  divided by 1–10 in step of 2. Therefore, the total number of parameter combinations for Experiment 4 was 5,000.

Finally, we resimulated the sampling results by using the best-fitting parameters for each participant and conducted the same data analysis (repeated-measurement ANOVA) as the behavioral experiments. If the simulated results can repeat the critical discoveries, we can further confirm the reliability of our model.

### Model Comparison

For model comparisons, we had four alternative models. First, to prove the necessity of the two proposed factors (task relevance and

visual uncertainty) in our model, we had Model 2 regarding the task relevance and Model 3 regarding the visual uncertainty as the only determining factor on cognitive resource allocation. Second, to prove our discoveries cannot be explained by solely task-related orientations or spatial locations, we had Model 4 assuming only orientation change and model 5 assuming only location change can explain the distraction effect for the additional full-distraction conditions. We fitted these alternative models with the same parameter space as the two-stage WM model for each participant. Then, we used the Akaike information criterion (AIC) (Hurvich & Tsai, 1989) and the Bayesian information criterion (BIC) (Schwarz, 1978) to compare the performance of these models. The AIC and BIC take into account the likelihood of the fitted models and the flexibility of each model in terms of the number of fitted parameters. A lower AIC and BIC indicates a better fit. To increase the robustness of the information criterion, we calculated AIC and BIC across all participants except one and repeated this process until all participants had been left out. For both AIC and BIC, we found the best-fitting model for each iteration using the mean value of the information criterion. We also counted the total times of choosing each model as the best-fitting model. Finally, to determine whether the difference in fit between the best fitting model and the other models was meaningful, we calculated, for each iteration, the difference in information criterion between the overall best-fitting model and each of the other models. By bootstrapping these differences 10,000 times, we computed 95% confidence intervals around the mean difference between the best fitting model and each of the other models. If the confidence interval on the mean difference does not include zero for all competitor models, then we can conclude that the best fitting model fits better than all other models. For Experiments 1–4, we only compared our model with the alternative Models 2 and 3, since there was no difference between our model and the alternative Models 4 and 5 when distractors did not change during the full-distraction condition. For Experiment 5, we compared our model with all the other four alternative models.

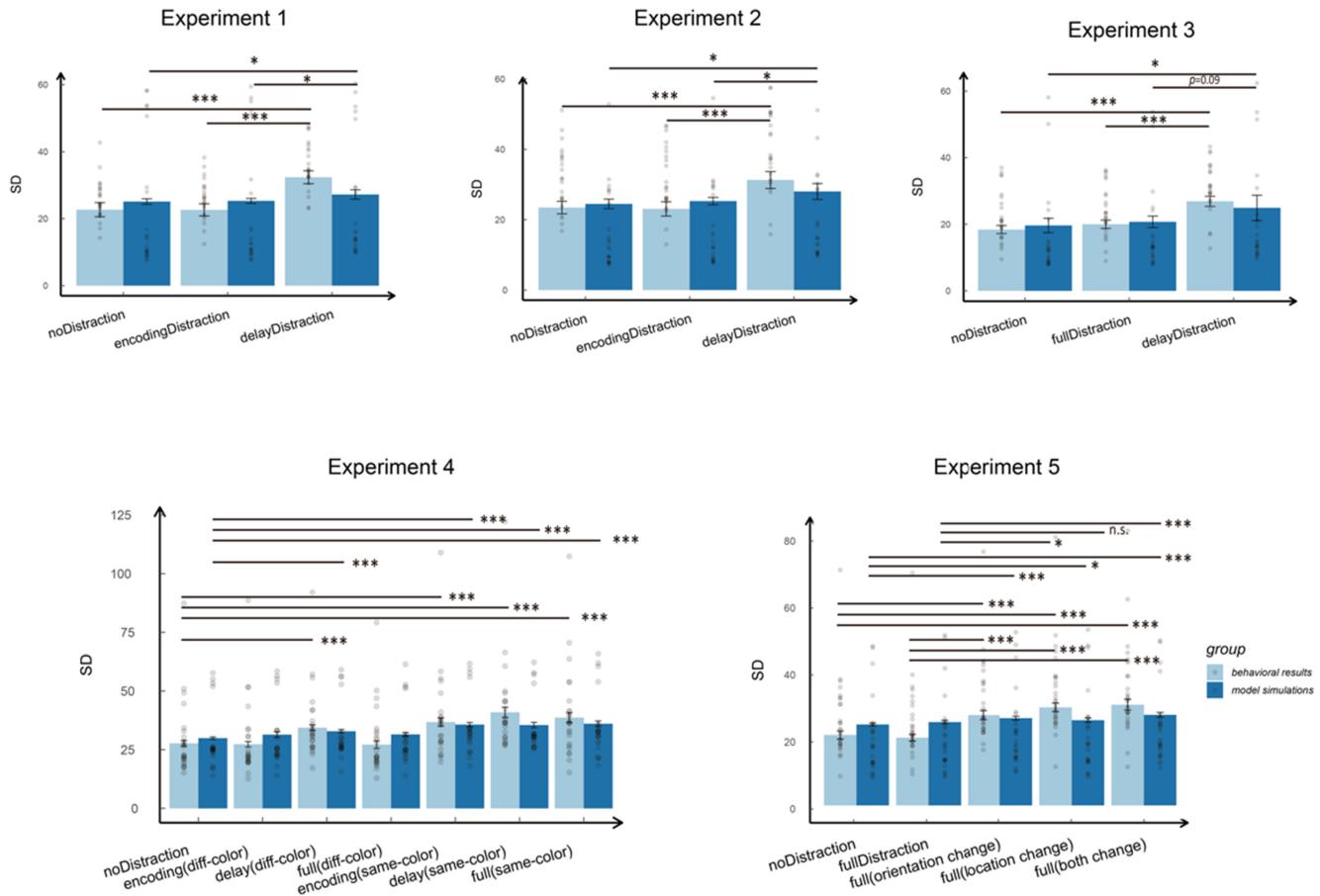
### Model Results and Discussions

All model code and simulated results have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/j8b92/>.

### Model Fitting Results

By searching for a wide parameter space, we found the parameters that minimized the MSE between the simulated  $SD$  and the real  $SD$  for each participant in Experiments 1–5 (for each participant's best-fitting parameters, see the online [supplemental material](#)). We also summarized the minimized MSE for all participants (Experiment 1: mean =  $20.30^\circ$ ,  $SD = 22.52^\circ$ ; Experiment 2: mean =  $42.22^\circ$ ,  $SD = 49.41^\circ$ ; Experiment 3: mean =  $6.24^\circ$ ,  $SD = 8.22^\circ$ ; Experiment 4: mean =  $31.87^\circ$ ,  $SD = 12.06^\circ$ ; Experiment 5: mean =  $22.78^\circ$ ,  $SD = 34.71^\circ$ ) (for each participant's minimized MSE, see the online [supplemental material](#)).

To further explore whether our model could capture the dissociated distraction effect, we performed the same statistical analysis (one-way repeated-measures ANOVA) on the re-simulated  $SD$  with the best-fitting parameters for each participant. The re-simulated  $SD$

**Figure 17***Results of Final Model Simulations and Behavioral Experiments*

**Note.** Our model captured all the behavioral discoveries. In general, *SD* was significantly higher in delay-distraction condition (Experiments 1–3), full-distraction (same-color) condition (Experiment 4), and full-distraction (orientation change, location change, both change) conditions (Experiment 5) compared with no-distraction condition. However, no difference was found between encoding-distraction condition, full-distraction condition, and no-distraction condition. Light blue represents behavioral results, while dark blue represents model results. Notably, the model results are simulated with the parameter setting as the median of all participants' best-fitting parameters. Grey dots represent each participant (for behavioral experiments) or each simulation (for model simulations) results. Error bars represent 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

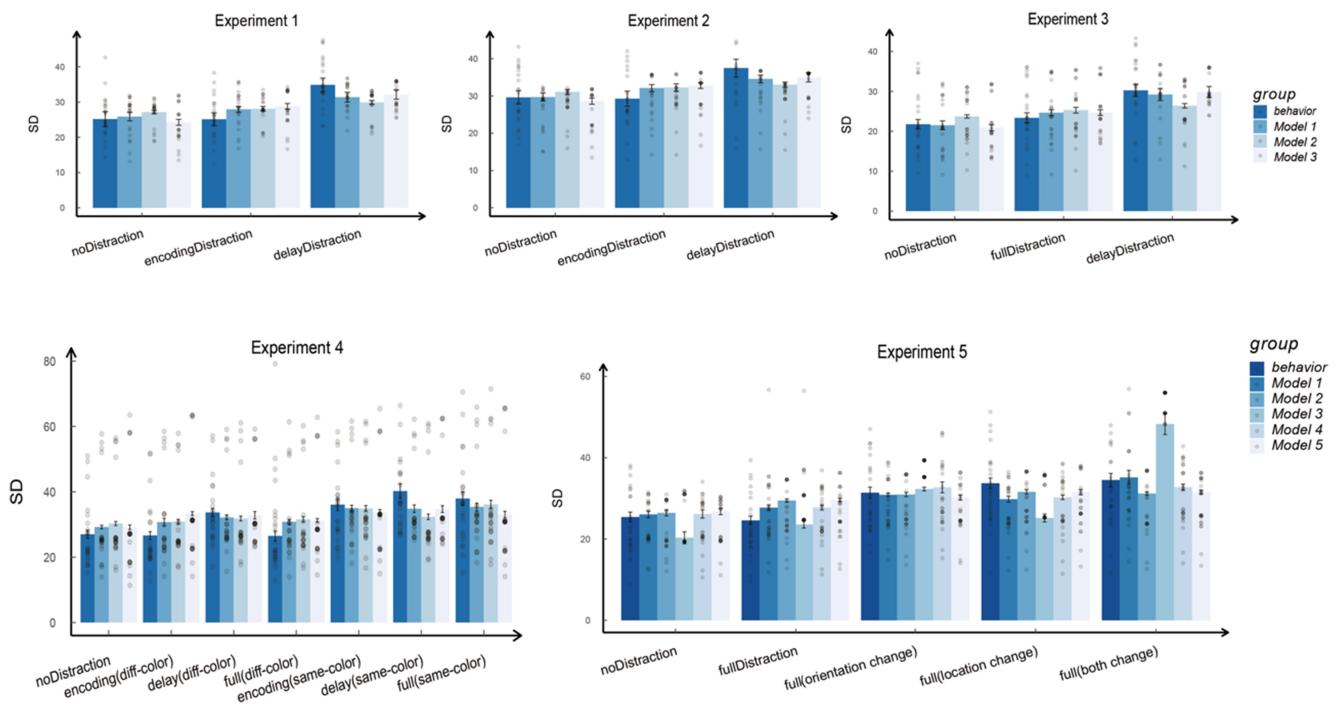
\* $p < .05$ . \*\*\* $p < .001$ .

performed very similar to our behavioral experiments (see Figure 17). For Experiments 1 and 2, *SD* was significantly higher in delay-distraction condition compared with no-distraction and encoding-distraction condition, but no difference was found between no-distraction condition and encoding-distraction condition—for Experiment 1:  $t_{(\text{no-encoding})} (22) = -0.28, p = 1.00$ , Cohen's  $d = -0.01$ ;  $t_{(\text{no-delay})} (22) = -3.00, p = .01$ , Cohen's  $d = -0.12$ ;  $t_{(\text{encoding-delay})} (22) = -2.71, p = .03$ , Cohen's  $d = -0.10$ ; for Experiment 2:  $t_{(\text{no-encoding})} (23) = -0.71, p = 1.00$ , Cohen's  $d = -0.07$ ;  $t_{(\text{no-delay})} (23) = -0.71, p = .01$ , Cohen's  $d = -0.07$ ;  $t_{(\text{encoding-delay})} (23) = -2.46, p = .05$ , Cohen's  $d = -0.24$ . These discoveries repeated the critical dissociation effect when distractors occurred at encoding versus during the delay. For Experiment 3, *SD* was significantly higher in delay-distraction condition compared with no-distraction, while no difference was found between full-distraction condition and no-distraction condition, as well as

between delay-distraction condition and full-distraction condition,  $t_{(\text{no-delay})} (23) = -2.83, p = .02$ , Cohen's  $d = -0.08$ ;  $t_{(\text{no-full})} (23) = -0.59, p = 1.00$ , Cohen's  $d = -0.08$ ;  $t_{(\text{full-delay})} (23) = -2.25, p = .09$ , Cohen's  $d = -0.32$ . For Experiment 4, *SD* was significantly higher in delay-distraction (different color) condition compared to no-distraction,  $t_{(\text{no-delay\_diff})} (23) = -6.27, p < .001$ , Cohen's  $d = -0.92$ , but not for encode-distraction (different color) and full-distraction (different color) conditions,  $t_{(\text{no-encode\_diff})} (23) = -2.35, p = 0.58$ , Cohen's  $d = -0.47$ ;  $t_{(\text{no-full\_diff})} (23) = -3.07, p = 0.13$ , Cohen's  $d = -0.45$ . However, when we used pre-cues to distinguish between targets and distractors, all same-color conditions had significant higher *SD* than the no-distraction condition,  $t_{(\text{no-encode\_same})} (23) = -12.14, p < 0.001$ , Cohen's  $d = -1.81$ ;  $t_{(\text{no-delay\_same})} (23) = -9.80, p < .001$ , Cohen's  $d = -1.75$ ;  $t_{(\text{no-full\_same})} (23) = -12.78, p < .001$ , Cohen's  $d = -1.92$ . For Experiment 5, no difference was found between full-distraction

condition and no-distraction condition,  $t_{(\text{no-full})} (27) = -1.84$ ,  $p = .69$ , Cohen's  $d = -0.07$ , while the other three full-distraction conditions (orientation change, location change, and both change) all had a significantly higher  $SD$  than no-distraction condition,  $t_{(\text{no-orientation change})} (27) = -4.84$ ,  $p < .001$ , Cohen's  $d = -0.17$ ;  $t_{(\text{no-location change})} (27) = -0.12$ ,  $p = .02$ , Cohen's  $d = -0.12$ ;  $t_{(\text{no-both change})} (27) = -7.32$ ,  $p < .001$ , Cohen's  $d = -0.26$ . Besides, we also found the  $SD$  for full-distraction condition was significantly lower than the full-distraction (orientation and both change) but not than the full-distraction (location change),  $t_{(\text{full-orientation change})} (27) = -3.00$ ,  $p = .03$ , Cohen's  $d = -0.11$ ;  $t_{(\text{full-location change})} (27) = -1.43$ ,  $p = 1.00$ , Cohen's  $d = -0.05$ ;  $t_{(\text{full-both change})} (27) = -5.49$ ,  $p < .001$ , Cohen's  $d = -0.20$ . Moreover, we found significant higher  $SD$  for full-distraction (both change) condition than full-distraction (location change) condition,  $t_{(\text{orientation-location change})} (27) = 1.57$ ,  $p = 1.00$ , Cohen's  $d = 0.06$ ;  $t_{(\text{orientation-both change})} (27) = -2.49$ ,  $p = 0.15$ , Cohen's  $d = -0.09$ ;  $t_{(\text{location-both change})} (27) = -4.06$ ,  $p < 0.001$ , Cohen's  $d = -0.15$ . To sum up, the re-simulation results based on our model repeated the distractor-induced interference in the delay-, full (same color)-, full (orientation change)-, full (location change)-, and full (both change)-distraction conditions, while no cost on memory precision in the encoding- and full-distraction conditions. These results suggested that our two-stage WM model could capture human behaviors very well.

**Figure 18**  
Results of Behavioral Experiments and All Model Fitting Results



*Note.* Model 1 represents our two-stage WM model with two factors. Model 2 represents the alternative model with only the task relevance factor. Model 3 represents the alternative model with only the visual uncertainty factor. Models 4 and 5 represent the alternative model regarding orientation change and location change separately as the only reason for distraction effect in the additional full-distraction conditions in Experiment 5. Grey dots represent each participant (for behavioral experiments) or each simulation (for model simulations) results. Error bars represent 95% CI within subjects (Cousineau, 2005). See the online article for the color version of this figure.

## Model Comparison Results

Having seen that our two-stage WM model captured the dissociated distraction effect well, we wanted to explore whether our model was superior to the other alternative models when explaining our behavioral discoveries. We compared our model with four alternative models, which regarded only one parameter or regarded only one feature as the determination for resource allocation in WM. We fitted these four alternative models with the same parameter space as our model for each participant. Since there was no difference between our model and the alternative Models 4 and 5 when distractors did not change during the full-distraction condition, we only compared the results between our model and the alternative Models 2 and 3 for Experiments 1–4. However, for experiment 5, we compared our model with the other four alternative models. Intuitively, we could see a better fit for our model than the other four alternative models (see Figure 18). Moreover, we used both AIC and BIC as the criteria for model fitting and plotted the number of participants best fit by each model and the mean difference relative to the best model based on both AIC and BIC (see Figure in the online supplemental material). As a result, our two-factor model performed best for most participants' data. Confidence intervals on the difference in fit between the best model and the other models did not include zero, suggesting our two-factor WM model fitted reliably better.

## Model Discussion

In general, our two-stage WM model could capture the behavioral data very well and strongly beat the other four alternative models. These discoveries supported the validity of our resource allocation principle and using it to understand the dissociated distraction effect. Here, we discussed why and how the model made it reasonable in our current WM task.

Importantly, we need to combine the influences of task relevance and visual uncertainty on both the encoding and delay periods to understand the dissociated distraction effect. For the encoding-distraction condition, the limited sampling number was divided between targets and distractors only at encoding. Since both targets and distractors were firstly encountered at encoding, visual uncertainty was the same for targets and distractors. Therefore, task relevance played the main role to allocate most resources to targets, leading to an effective resistance to distractors. For the delay-distraction condition, the limited sampling number was divided between targets and distractors only during the delay. Even though task relevance was similar to the encoding-distraction condition, a lot of resources were captured by distractors because of their much higher visual uncertainty than targets' during the delay. Therefore, lacking resources during the delay resulted in a cost on targets' memory performance. For the full-distraction condition, the resource allocation process was the same as the encoding-distraction condition. During the delay, even distractors had a higher visual uncertainty than targets because distractors had a noisier posterior distribution than targets after the encoding period, it was much better than a total uniform distribution as the delay-distraction condition. Therefore, when being combined with task relevance, most resources were still allocated to targets during the delay period, leading to a successful resistance to distractors in the full-distraction condition.

As for those same color conditions, since targets were less distinguishable from distractors, the task relevance difference between targets and distractors decreased, leading to less resource allocation for targets during both periods. Therefore, the distraction-related cost increased for all three same color. Finally, for the additional three full-distraction conditions (with orientation change, location change, or both change), the distraction-related cost also revived since any change made to distractors would cause a much higher visual uncertainty. The increased visual uncertainty occupied more WM resources during the delay period, leading to an interference on targets' memory performance in the end. One thing to note is that our model's simulation results were not as strong as the behavioral results for the full-distraction (location change) condition (see Figure 17), which could also be reflected in the sampling results (see Figure in the online [supplementary material](#)). There might be two main reasons. First, our model calculated the visual uncertainty caused by location change based on an estimation of area selection, which could lead to some deviations. Second, in the full-distraction (location change) condition, the location change might lead to orientation change since features and locations were bound together after encoding. Future studies are needed to investigate whether and how location change can be separated from feature change.

## General Discussion

### Summary

The human cognitive system is characterized by limited capacity, which embeds a potential resource allocation mechanism. In the

current study, we aimed to address the resource allocation problem in WM in terms of the distraction effect (Lorenc et al., 2021). To do so, we presented the distractors during the encoding or delay or throughout both periods, respectively. In general, we found a dissociated distraction effect between the encoding and delay conditions. Specifically, a reliable distraction-related cost of WM fidelity was found when distractors were only presented during the delay, but not as long as distractors were introduced at encoding (Experiments 1–3). Moreover, we revealed that the success of distractor resistance was due to the efficient suppression of distractors at encoding. We found a revival of interference when decreasing the distinguishability between distractors and targets (Experiment 4) as well as when we changed the orientations or locations of distractors during the delay (Experiment 5). Finally, we revealed that the robust distraction-related cost during the delay did not require extra spatial resources, as the distraction effect existed no matter whether distractors shared the same locations with targets or not (Experiments 6a and 6b).

The main findings of the dissociated distraction effect across all experiments could be explained within a unified framework of task relevance and visual uncertainty. When uncertainty levels were equal for targets and distractors, the task relevance advantage helped participants to resist distraction and WM resources would be primarily allocated toward targets. However, when distractors came with stronger uncertainty, they gained more resources, and the resistance was weakened if not broken.

To reveal the details of the mechanism that underlies the role of task relevance and uncertainty, we proposed a two-stage Bayesian model. Importantly, our model replicated different distraction effects across different conditions, and the final model re-simulation revealed the robustness of our model. The model comparison also supported the superiority of our model compared with other alternative models. In general, our findings revealed a dissociated interference effect when distractors were introduced at different times. Moreover, our model uncovered the underlying mechanisms by regarding the WM task as a sampling process and advanced our understanding of the limited resource allocation.

### Dissociated Distraction Effect of Mnemonic Fidelity

From six behavioral experiments, we found a reliable interference of WM mnemonic fidelity when distractors were presented during the delay. These results were consistent with previous studies, finding that irrelevant stimuli occurring during the delay degraded performance on a delayed discrimination task, including spatial frequency (Bennett & Cortese, 1996; Magnussen et al., 1991; Nemes et al., 2011), direction of motion (Magnussen & Greenlee, 1992; McKeefry et al., 2007; Pasternak & Zaksas, 2003), color (Hakim et al., 2019, 2020; Nemes et al., 2012; Nilsson & Nelson, 1981), or faces (Clapp et al., 2010; Mallett et al., 2020). For example, a recent study discovered that irrelevant information could be filtered out well at encoding but had a detrimental effect during the delay, and these effects remained the same across the lifespan (Tabi et al., 2021). Moreover, another study investigated the impact of interference on visual orientation during the maintenance period (Rademaker et al., 2015). In their study, they presented a second orientation during the delay while participants maintained a single orientation in memory. They parametrically varied the differences between the two orientations and found that memory for the target's

orientation was noisier when the target and distractor differed. Our current study verified this finding and extended the distraction effect to multiple items.

Interestingly, participants seemed capable of rejecting interference from distractors as long as they were processed along with targets at encoding (Experiments 1–3). These results were consistent with previous findings on the active suppression process. For example, Vogel et al. (2005) revealed that people with high WM capacity could resist distractors' interference by not representing distractors during the delay. A recent study also confirmed the active inhibition process during the encoding period with both spontaneous (PD, distractor positivity) and sustained (CDAp, contralateral delay activity [CDA] of positive polarity) event-related potential (ERP) components (Feldmann-Wüstefeld et al., 2018; Feldmann-Wüstefeld & Vogel, 2019). However, previous studies also revealed individual differences in performing the distractor-filtering process (Vogel et al., 2005; Vogel & Machizawa, 2004), showing that people with low WM capacity and old ages (Gazzaley et al., 2005, 2008; McNab et al., 2015) were suffered from suppression deficient. Although we did not consider the individual differences in our experiments, we explicitly manipulated the efficiency of distinguishing targets and distractors in Experiment 4. By using a pre-cue instead of an obvious color difference to distinguish targets and distractors, we found a revival of the distraction-related cost on WM precision. We claimed that these operations might correspond to the circumstances that people with low WM capacity and old ages had difficulty separating targets from distractors. Moreover, we fitted our model with participants' data in Experiment 4 by adding an extra free parameter, which denoted the relative task relevance ratio between targets and distractors in these same-color conditions. The model-fitting results revealed a lower relative task relevance ratio in same-color conditions compared with different-color conditions (see the online supplemental material). Therefore, we argued that the individual differences and other manipulations that change the distinguishability between targets and distractors could be explained by our task relevance factor.

In general, our current study provided the first empirical evidence revealing a dissociated distraction effect on memory precision when distractors were introduced at different WM stages. That is, the distraction-related costs only exist when novel distractors were introduced during the delay but not at encoding. The effectiveness of suppression distractors at encoding also influenced the final memory performance. These results advanced our understanding of distraction-related interference on WM fidelity.

### **Distraction Effect and Visuo-Spatial Resources**

Whether the successful distractor resistance was due to the inhibitory of the distractor locations or the processing of distractors was another core question. According to the model of WM proposed by Baddeley (Baddeley et al., 1986; Baddeley & Hitch, 1974), the visuo-spatial sketchpad maintains both visual and spatial information. Initial studies focused on whether the visuo-spatial sketchpad was essentially visual or spatial. For example, Baddeley (2012) claimed that the visuo-spatial sketchpad was predominantly spatial because a concurrent spatial task impaired performance on a visuo-spatial imagery task while a concurrent visual task did not. However, other studies failed to replicate these results and found a dissociation

between visual and spatial WM processes (e.g., Beech, 1984; Brooks, 1967; Della Sala et al., 1999; Logie & Marchetti, 1991; Quinn, 1988; Tresch et al., 1993). Moreover, another alternative interpretation proposed that the observed dissociations might reflect the operation of a visuo-spatial WM system, containing a combination of spatial and visual processes (Baddeley & Logie, 1999; Logie, 1995).

Accordingly, there were three hypotheses regarding how the efficient suppression of distractors at encoding guaranteed the successful distractor resistance in the full-distraction condition. The successful resistance might be due to the inhibition of distractors' locations, or task-directed features, or objects binding locations and features together. In Experiment 5, we manipulated the orientations and locations of distractors during the delay and found that both changes caused a significant distraction effect. These results indicated that both visual and spatial resources were important for distraction resistance, so that either change would destroy the efficient distractor resistance. Therefore, our discoveries advanced our understanding of the collaboration between visual and spatial WM processes. Besides, the combined view of visuo-spatial WM also parallels Baddeley's account of articulatory rehearsals processes and phonological storage in verbal WM, indicating similar functional architectures within the WM system.

Recent studies used electroencephalography to investigate spatial versus item-based distraction effects (Hakim et al., 2020, 2021). They used lateralized alpha power and CDA to track spatial capture and item-based capture separately. They found that task-irrelevant distractors captured spatial attention but not item-based attention and claimed that task-irrelevant distractors were not encoded into WM via item-based capture. Even though spatial capture was involuntary for salient distractors, we still did not know whether spatial capture was necessary for distraction-related interference. In Experiments 6a and 6b, we presented distractors at the same locations as targets and found similar distraction-induced interference as when distractors were presented at different locations. Moreover, the Experiment 5 also revealed that both changing orientations and locations could lead to a similar distraction effect. These results indicated that spatial resources could be occupied by distractors. However, it was not necessary for the distraction effect. In the view of Bayesian computation, both the spatial and non-spatial resources could be combined into a common resource divided between targets and distractors. Our model also regarded both spatial and non-spatial factors when determining the visual uncertainty factor. Alternative models failed to capture this phenomenon, which assumed only the location or orientation change mattered. Therefore, our current studies advance our understanding of the relationship between spatial and visual resources within WM.

In general, our results suggested that spatial locations were not necessary for distraction-related interference and not sufficient for successful distraction resistance. However, we cannot deny the importance of spatial locations in WM. Instead, we should combine it with the visual process and focus on how these two complement each other to help promote WM capacity.

### **Using Bayesian Model to Understand the Dissociated Effect**

Another important contribution of our current study is having proposed a unified computational mechanism to reconcile the

dissociated distraction effect. Our two-stage WM model not only replicated the dissociated distraction effect under different conditions, but also outperformed the other four alternative models. These results revealed the superiority of our model by combining the Bayesian inference process and the resource allocation principles. In general, our model assumed a limited cognitive capacity during both periods and regarded the WM task as a Bayesian sampling process (Brady & Tenenbaum, 2013). Despite some differences in assumptions, previous prominent cognitive models of visual WM also emphasized the limited WM capacity, including the slot model (Luck & Vogel, 1997), slot-averaging model (Zhang & Luck, 2008), resource model (Bays & Husain, 2008; van den Berg et al., 2012), and sample-size models (Sewell, 2014; Smith et al., 2016). For example, the sample-size model regarded visual WM as a resource comprised of a finite number of noisy stimulus samples. The sample-size model predicted the invariance of the sum of squared sensitivities across items for different sample sizes. Therefore, if there is no preferential weighting of items, then each stimulus will be represented by the same odds of the total samples. Our two-stage WM model also assumed a limited cognitive resource to be divided between targets and distractors, and we assumed the resource allocation would be the same within targets and distractors once determined by the resource allocation principles.

Importantly, we proposed one resource allocation principle at both stages by considering two main factors. The first factor is task relevance, which is determined by the role of items. Intuitively, the task relevance of targets is much higher than distractors, no matter during which period. Task relevance was always considered as the top-down determination of our attentional resource allocation (Fecteau & Munoz, 2006). Studies from neuropsychology demonstrated that both local brain activity (prefrontal cortex) (Baluch & Itti, 2011; Knudsen, 2007; Opris et al., 2005; Rossi et al., 2007; Tomita et al., 1999) and neural oscillations (de Vries et al., 2020; Fries, 2009; Fries et al., 2002) supported the guidance of task relevant attention. Recently, different cognitive and neuroscientific mechanisms have been proposed to illustrate the distractor filtering during the encoding (Liesefeld et al., 2020) and delay (Lorenc et al., 2021) periods. Liesefeld et al. (2020) proposed that the successful suppression of distraction should maximize the difference between targets and distractors gain, and Lorenc et al. (2021) claimed that the resistance of distractors depended on the relative activation level (priority) of targets. Both frameworks took the limited WM capacity as a prerequisite and regarded the relative importance between targets and distractors as the determining factor. Our model agreed with this claim by adding the relative task relevance ratio between targets and distractors as one of the determining factors of cognitive resource allocation. Beyond that, we also manipulated the relative task relevance ratio to fit our results in Experiment 4. By decreasing the relative task relevance ratio, our model captured the distraction-related interference in these same-color conditions.

Another crucial factor in our model is the visual uncertainty of items, which explains most of the dissociated distraction effect. Visual uncertainty is always considered as an intrinsic reward that drives people's attention so that more cognitive resources would be automatically distributed to uncertain staff. For example, the memory benefit of information appearing at the event boundary could be a result of unpredictability-driven attentional priority

(Heusser et al., 2018; Newson & Engquist, 1976; Polyn et al., 2009; Schwan & Garsoffky, 2004; Swallow et al., 2009). Moreover, in reinforcement learning, uncertainty always serves as an intrinsic motivation to explore an unknown environment (Oudeyer et al., 2007; Oudeyer & Kaplan, 2006; Pathak et al., 2017; Sun et al., 2011). Since people have a strong aversion to uncertainty (Hogg, 2000), humans tend to resolve uncertainty as quickly as possible (Bennett et al., 2016). Consequentially, uncertain items would involuntarily obtain more cognitive resources during the encoding and delay periods. Previous research also revealed that compared with constant items, people involuntarily allocated more WM workspace to the ever-changing ones. For example, Berggren et al. (2020) found that response time was slower, and the ERP components emerged later and smaller for the sustained target as compared to transient target objects. Similarly, Schurgin et al. (2018) discovered that people freed their online WM workspace of repeated objects with the help of long-term memory by showing a decreased amplitude of CDA. These discoveries indicated that people voluntarily allocate fewer resources to more familiar items over the course of the experiment. However, our current study did not manipulate the familiarity across the experiment since all items' orientations and locations were randomly selected. Instead, our model focused on the change of visual uncertainty within a single trial across the encoding and delay periods. Future work is needed to investigate the dynamic change of visual uncertainty with the learning effect across the whole experiment.

Accordingly, our model included visual uncertainty as a determining factor in cognitive resource allocation and assumed that items with higher uncertainty (quantified by entropy) would capture more resources. Similar to the task relevance, the relative visual uncertainty between targets and distractors determined the resource allocation at both two stages. When both targets and distractors were presented at encoding, their visual uncertainty was comparable. However, if the distractors had not been processed at the beginning, their higher uncertainty would occupy most cognitive resources during the delay period, leading to a distraction-related cost. Moreover, our model predicted that changes that happened to distractors during the delay would increase the visual uncertainty, leading to a revive in distraction-related interference. Importantly, both the behavioral results in Experiment 5 and our model's predictions revealed that changes to both task relevant orientations and locations could induce distraction. Therefore, the successful resistance to distractors in the full-distraction condition cannot be simply explained by the inhibitory of distractors' locations. On the contrary, a better way to explain this effect is to include both spatial and non-spatial features in a unified way. Our model quantified the effect of both orientations and locations by adding their influence on the visual uncertainty parameter, which determined the resource allocation and influenced the final memory performance.

Combining these two factors, our model predicted how much resources would be allocated to targets and distractors during both periods. Although we admitted that we could not exclude other factors that may influence the resource allocation process, the good performance of our current model, with the full replication of all behavioral experiments as well as better performance than alternative models, supported the superiority of our two-stage Bayesian model. Moreover, our model also suggested that the same resource allocation principle can be applied at both two stages, providing a parsimonious perspective to understand the dissociated distraction effect.

## Model Generalization and Limitations

Our model can not only predict performance in the current study but also provide new perspectives to related phenomena in a broader WM field. First, we compared our current study to a similar effect in verbal WM named orienting response (OR). OR is a classic WM interference effect in a novel, unexpected, a salient stimulus elicits an automatic attentional reaction (Sokolov, 1963). Irrelevant information presented during the encoding or delay periods had a detriment impact on WM performance for visually or auditorily items (e.g., Andrade et al., 2002; Beaman et al., 2007; Chein & Fiez, 2010; Colle & Welsh, 1976; Hanley & Bakopoulou, 2003; Röer et al., 2011). Habituation of the OR (Sokolov, 1963) refers to a reduction in the detrimental effect of an irrelevant stimulus on relevant task performance following practice with, or exposure to, features of the irrelevant stimulus (Elliott & Cowan, 2001; Sörqvist et al., 2012; Waters et al., 1977). A changing-state effect refers to a single repetitive token that generates little or no disruption when compared with a sequence made up of two or more different tokens (Bell et al., 2010; Bonetti & Turatto, 2019; Jones et al., 1992; Turatto et al., 2018). This effect is explained by the fact that novel or changing stimuli elicit an OR while the OR habituates with repeated stimulus representation (Thompson & Spencer, 1966).

According to Cowan's embedded-processes model (1988, 1995, 1999), habituation of the OR serves as an attentional filter. It is assumed that a neural model of the repeated stimulus is formed to which each incoming stimulus is compared. If an important discrepancy from the neural model is detected or the neural model is not yet established, an OR occurs (Gati & Ben-Shakhar, 1990; Sokolov, 1963). This theoretical idea is allied to the role of visual uncertainty in our model. In our experiments, when distractors remained the same during the delay, the visual uncertainty for distractors decreased, resulting in no cost to WM performance in the full-distractor condition. However, when distractors' features changed, the OR was elicited, and visual uncertainty increased, resulting in a revival of the distraction-related interference in the full-distraction (orientation change, location change, and both change) conditions. Therefore, our experiments expanded the habituation of the OR from a trial-wised setting into within-trial habituation. Moreover, our model could provide a unified framework to quantify both effects. When task-irrelevant distractors were repeated during the experiment, the decrease of distractors' uncertainty occupied less cognitive resources, resulting in a less detriment to the main task. Therefore, our model could provide a computational view of the habituation effect.

Second, researchers focusing on WM distractor-resistance mechanisms proposed that distractors filtering can be implemented in a biased-competition framework (Desimone & Duncan, 1995). According to the framework, various stimuli compete for access to visual WM. Bottom-up stimulus saliency is modulated by top-down task goals, resulting in a priority map that controls the allocation of processing resources. Within this framework, distractors can be handled by both distractor suppression and target enhancement. However, researchers claimed that distractor suppression and target enhancement might be difficult to disentangle (Liesefeld et al., 2020). For example, the differences for probes at target versus distractor locations can be interpreted by both the ability to inhibit distractors and enhance targets (Feldmann-Wüstefeld & Vogel, 2019; Fukuda & Vogel, 2009). Our model could solve this problem by

modeling the resource allocation process at each stage and comparing the allocated sampling numbers between targets and distractors. By simulating results with different parameters, we could find the optimal allocation patterns that fit human behaviors. One possibility could be that different people have different resource allocation strategies, which might be ignored by previous averaging methods.

However, we have to admit that our current model is still a simplification of the WM process. To make our model more computationally realizable, we explicitly made some assumptions. For example, our model assumed that the limited resources remained the same during both periods and observers used up all resources during both periods. Since we asked participants to remember three targets' orientations at the same time, using up all WM resources was acceptable. However, for studies only have one target might need to make some adjustments. Besides, our model did not cover sequential presented stimuli as well as considering the time decaying factors. Future studies are needed for further exploration. In some subtle cases, our model might overestimate or underestimate the distraction effect. However, the superiority of our model is to use only three (or four) free parameters to capture the main dissociated distraction effect and consistently predicted human behaviors across all experiments. This can provide us a new perspective to understand the underlying mechanism of WM distraction resistance.

## Conclusion

Using six experiments, we revealed a dissociated distraction effect of mnemonic fidelity in human WM. A robust interference effect existed when novel distractors were presented during the delay. However, people can resist the interference if distractors were processed with targets at encoding. It should be noted that successful distraction resistance at encoding depended on the efficiency and continuity of distinguishing targets from distractors. Spatial changes could induce but were not necessary for the distraction effect. These results could be explained by taking stimuli's task relevance and visual uncertainty into account. We also established a computational model to reveal the underlying mechanism. We regarded the WM task as a two-stage Bayesian inference process and parametrized the two proposed factors in our model. By fitting each participant's data throughout different experiments, our model replicated the dissociated distraction effect respectively and outperformed other alternative models. Therefore, we claimed that both our behavioral experiments and computational model advanced our understanding of the WM resource allocation mechanism.

## Context of the Research

For the past few years, our research focused on the dynamic resource allocation of WM. WM is well-known for its limited capacity, and traditional research focused on the constrain. However, our lab regards the limitation as a trigger for human intelligence. How to use limited resources to achieve behavioral goals is the core question we want to address.

In this study, we aimed to understand how people allocate limited WM resources when facing external distractors. We systematically investigated how distractors influence WM performance during different periods and found a dissociated distraction effect. Besides, we proposed a Bayesian model to understand how resource allocation contributes to the dissociated distraction effect. The current study

sheds light on research in the WM field and deepens our understanding of the resource allocation principles. Following this study, we would like to conduct further research to generalize our model to other areas. We hope our research on WM could help us understand more about our cognitive system with limited resources.

## References

- Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, 30(3), 221–256. <https://doi.org/10.1006/cogp.1996.0007>
- Andrade, J., Kemps, E., Werniers, Y., May, J., & Szmałec, A. (2002). Insensitivity of visual short-term memory to irrelevant visual information. *The Quarterly Journal of Experimental Psychology Section A*, 55(3), 753–774. <https://doi.org/10.1080/02724980143000541>
- Arnell, K. M., & Stubitz, S. M. (2010). Attentional blink magnitude is predicted by the ability to keep irrelevant material out of working memory. *Psychological Research*, 74(5), 457–467. <https://doi.org/10.1007/s00426-009-0265-8>
- Awh, E., & Vogel, E. (2008). The bouncer in the brain. *Nature Neuroscience*, 11(1), 5–6. <https://doi.org/10.1038/nn0108-5>
- Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63(1), 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Baddeley, A. D., Hitch, G., & Bower, G. A. (Ed.). (1974). *Psychology of learning and motivation*. Academic Press.
- Baddeley, A. D., Logie, R., Bressi, S., Sala, S. D., & Spinnler, H. (1986). Dementia and working memory. *The Quarterly Journal of Experimental Psychology Section A*, 38(4), 603–618. <https://doi.org/10.1080/14640748608401616>
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge University Press.
- Baluch, F., & Itti, L. (2011). Mechanisms of top-down attention. *Trends in Neurosciences*, 34(4), 210–224. <https://doi.org/10.1016/j.tins.2011.02.003>
- Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9(10), Article 7. <https://doi.org/10.1167/9.10.7>
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321(5890), 851–854. <https://doi.org/10.1126/science.1158023>
- Beaman, C. P., Bridges, A. M., & Scott, S. K. (2007). From dichotic listening to the irrelevant sound effect: A behavioural and neuroimaging analysis of the processing of unattended speech. *Cortex*, 43(1), 124–134. [https://doi.org/10.1016/S0010-9452\(08\)70450-7](https://doi.org/10.1016/S0010-9452(08)70450-7)
- Beech, J. R. (1984). The effects of visual and spatial interference on spatial working memory. *The Journal of General Psychology*, 110(2), 141–149. <https://doi.org/10.1080/00221309.1984.9709959>
- Bell, R., Dentale, S., Buchner, A., & Mayr, S. (2010). ERP correlates of the irrelevant sound effect. *Psychophysiology*, 47(6), 1182–1191. <https://doi.org/10.1111/j.1469-8986.2010.01029.x>
- Bennett, D., Bode, S., Brydavall, M., Warren, H., & Murawski, C. (2016). Intrinsic valuation of information in decision making under uncertainty. *PLoS Computational Biology*, 12(7), Article e1005020. <https://doi.org/10.1371/journal.pcbi.1005020>
- Bennett, P. J., & Cortese, F. (1996). Masking of spatial frequency in visual memory depends on distal, not retinal, frequency. *Vision Research*, 36(2), 233–238. [https://doi.org/10.1016/0042-6989\(95\)00085-E](https://doi.org/10.1016/0042-6989(95)00085-E)
- Berggren, N., Nako, R., & Eimer, M. (2020). Out with the old: New target templates impair the guidance of visual search by preexisting task goals. *Journal of Experimental Psychology: General*, 149(6), 1156–1168. <https://doi.org/10.1037/xge0000697>
- Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *British Journal of Psychology*, 41(1–2), 68–80. <https://doi.org/10.1111/j.2044-8295.1950.tb00262.x>
- Bonetti, F., & Turatto, M. (2019). Habituation of oculomotor capture by sudden onsets: Stimulus specificity, spontaneous recovery and dishabituation. *Journal of Experimental Psychology: Human Perception and Performance*, 45(2), 264–284. <https://doi.org/10.1037/xhp0000605>
- Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review*, 120(1), 85–109. <https://doi.org/10.1037/a0030779>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Brooks, L. R. (1967). The suppression of visualization by reading. *The Quarterly Journal of Experimental Psychology*, 19(4), 289–299. <https://doi.org/10.1080/14640746708400105>
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia*, 49(6), 1446–1457. <https://doi.org/10.1016/j.neuropsychologia.2010.12.006>
- Case, R. (1972). Validation of a neo-Piagetian mental capacity construct. *Journal of Experimental Child Psychology*, 14(2), 287–302. [https://doi.org/10.1016/0022-0965\(72\)90051-3](https://doi.org/10.1016/0022-0965(72)90051-3)
- Chein, J. M., & Fiez, J. A. (2010). Evaluating models of working memory through the effects of concurrent irrelevant information. *Journal of Experimental Psychology: General*, 139(1), 117–137. <https://doi.org/10.1037/a0018200>
- Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2010). Mechanisms of working memory disruption by external interference. *Cerebral Cortex*, 20(4), 859–872. <https://doi.org/10.1093/cercor/bhp150>
- Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, 15(1), 17–31. [https://doi.org/10.1016/S0022-5371\(76\)90003-7](https://doi.org/10.1016/S0022-5371(76)90003-7)
- Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552. <https://doi.org/10.1016/j.tics.2003.10.005>
- Cools, R., Miyakawa, A., Sheridan, M., & D'Esposito, M. (2010). Enhanced frontal function in Parkinson's disease. *Brain*, 133(1), 225–233. <https://doi.org/10.1093/brain/awp301>
- Cousineau, 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. <https://doi.org/10.20982/tqmp.011.p042>
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2), 163–191. <https://doi.org/10.1037/0033-2950.104.2.163>
- Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford University Press.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press.
- Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, 24(4), 1158–1170. <https://doi.org/10.3758/s13423-016-1191-6>
- Cox, H. (1964). On the estimation of state variables and parameters for noisy dynamic systems. *IEEE Transactions on Automatic Control*, 9(1), 5–12. <https://doi.org/10.1109/TAC.1964.1105635>
- Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia*, 37(10), 1189–1199. [https://doi.org/10.1016/S0028-3932\(98\)00159-6](https://doi.org/10.1016/S0028-3932(98)00159-6)

- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(1), 193–222. <https://doi.org/10.1146/annurev.ne.18.030195.001205>
- D'esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature*, 378(6554), 279–281. <https://doi.org/10.1038/378279a0>
- de Vries, I. E., Slagter, H. A., & Olivers, C. N. (2020). Oscillatory control over representational states in working memory. *Trends in Cognitive Sciences*, 24(2), 150–162. <https://doi.org/10.1016/j.tics.2019.11.006>
- Düzel, E., Bunzeck, N., Guitart-Masip, M., & Düzel, S. (2010). Novelty-related motivation of anticipation and exploration by dopamine (NOMAD): Implications for healthy aging. *Neuroscience & Biobehavioral Reviews*, 34(5), 660–669. <https://doi.org/10.1016/j.neubiorev.2009.08.006>
- Elliott, E. M., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal, color-word interference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(3), 654–667. <https://doi.org/10.1037/0278-7393.27.3.654>
- Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, 10(8), 382–390. <https://doi.org/10.1016/j.tics.2006.06.011>
- Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, Article 215. <https://doi.org/10.3389/fnhum.2010.00215>
- Feldmann-Wüstefeld, T., & Vogel, E. K. (2019). Neural evidence for the contribution of active suppression during working memory filtering. *Cerebral Cortex*, 29(2), 529–543. <https://doi.org/10.1093/cercor/bhw336>
- Feldmann-Wüstefeld, T., Vogel, E. K., & Awh, E. (2018). Contralateral delay activity indexes working memory storage, not the current focus of spatial attention. *Journal of Cognitive Neuroscience*, 30(8), 1185–1196. [https://doi.org/10.1162/jocn\\_a\\_01271](https://doi.org/10.1162/jocn_a_01271)
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual Review of Neuroscience*, 32(1), 209–224. <https://doi.org/10.1146/annurev.neuro.051508.135603>
- Fries, P., Schröder, J. H., Roelfsema, P. R., Singer, W., & Engel, A. K. (2002). Oscillatory neuronal synchronization in primary visual cortex as a correlate of stimulus selection. *Journal of Neuroscience*, 22(9), 3739–3754. <https://doi.org/10.1523/JNEUROSCI.22-09-03739.2002>
- Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301. <https://doi.org/10.1016/j.tics.2009.04.005>
- 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>
- Fukuda, K., & Vogel, E. K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science*, 22(3), 361–368. <https://doi.org/10.1177/0956797611398493>
- Gati, I., & Ben-Shakhar, G. (1990). Novelty and significance in orientation and habituation: A feature-matching approach. *Journal of Experimental Psychology: General*, 119(3), 251–263. <https://doi.org/10.1037/0096-3445.119.3.251>
- Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., & D'Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proceedings of the National Academy of Sciences*, 105(35), 13122–13126. <https://doi.org/10.1073/pnas.0806074105>
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, 8(10), 1298–1300. <https://doi.org/10.1038/nrn1543>
- Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, 17(11), 585–593. <https://doi.org/10.1016/j.tics.2013.09.001>
- Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, 15(8), 1176–1194. <https://doi.org/10.1162/08989290332598139>
- Hakim, N., Adam, K. C., 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. *Psychological Science*, 30(4), 526–540. <https://doi.org/10.1177/0956797619830384>
- Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2020). Perturbing neural representations of working memory with task-irrelevant interruption. *Journal of Cognitive Neuroscience*, 32(3), 558–569. [https://doi.org/10.1162/jocn\\_a\\_01481](https://doi.org/10.1162/jocn_a_01481)
- Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2021). Controlling the flow of distracting information in working memory. *Cerebral Cortex*, 31(7), 3323–3337. <https://doi.org/10.1093/cercor/bhab013>
- Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression, and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review*, 10(2), 435–444. <https://doi.org/10.3758/BF03196503>
- Heusser, A. C., Ezzyat, Y., Shiff, L., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(7), 1075–1090. <https://doi.org/10.1037/xlm0000503>
- Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European Review of Social Psychology*, 11(1), 223–255. <https://doi.org/10.1080/1479272043000040>
- Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2), 297–307. <https://doi.org/10.1093/biomet/76.2.297>
- Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.
- Jones, D., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *The Quarterly Journal of Experimental Psychology Section A*, 44(4), 645–669. <https://doi.org/10.1080/14640749208401304>
- Jost, K., & Mayr, U. (2016). Switching between filter settings reduces the efficient utilization of visual working memory. *Cognitive, Affective, & Behavioral Neuroscience*, 16(2), 207–218. <https://doi.org/10.3758/s13415-015-0380-5>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review Neuroscience*, 30(1), 57–78. <https://doi.org/10.1146/annurev.neuro.30.051606.094256>
- Lehman, J., & Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation*, 19(2), 189–223. [https://doi.org/10.1162/EVCO\\_a\\_00025](https://doi.org/10.1162/EVCO_a_00025)
- Liesefeld, A. M., Liesefeld, H. R., & Zimmer, H. D. (2014). Intercommunication between prefrontal and posterior brain regions for protecting visual working memory from distractor interference. *Psychological Science*, 25(2), 325–333. <https://doi.org/10.1177/0956797613501170>
- Liesefeld, H. R., Liesefeld, A. M., Sauseng, P., Jacob, S. N., & Müller, H. J. (2020). How visual working memory handles distraction: Cognitive mechanisms and electrophysiological correlates. *Visual Cognition*, 28(5–8), 372–387. <https://doi.org/10.1080/13506285.2020.1773594>
- Logie, R. H. (1995). *Visuo-spatial working memory*. Erlbaum.
- Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Advances in psychology*: Vol. 80. Mental Images in Human Cognition (pp. 105–115). North Holland. <https://doi.org/10.1016/S0166-4115>

- Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in visual working memory: Resistance is not futile. *Trends in Cognitive Sciences*, 25(3), 228–239. <https://doi.org/10.1016/j.tics.2020.12.004>
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281. <https://doi.org/10.1038/36846>
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17(3), 347–356. <https://doi.org/10.1038/nn.3655>
- Magnussen, S., & Greenlee, M. W. (1992). Retention and disruption of motion information in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(1), 151–156. <https://doi.org/10.1037/0278-7393.18.1.151>
- Magnussen, S., Greenlee, M. W., Asplund, R., & Dymes, S. (1991). Stimulus-specific mechanisms of visual short-term memory. *Vision Research*, 31(7–8), 1213–1219. [https://doi.org/10.1016/0042-6989\(91\)90046-8](https://doi.org/10.1016/0042-6989(91)90046-8)
- Mallett, R., Mummaneni, A., & Lewis-Peacock, J. A. (2020). Distraction biases working memory for faces. *Psychonomic Bulletin & Review*, 27(2), 350–356. <https://doi.org/10.3758/s13423-019-01707-5>
- McKeefry, D. J., Burton, M. P., & Vakrou, C. (2007). Speed selectivity in visual short term memory for motion. *Vision Research*, 47(18), 2418–2425. <https://doi.org/10.1016/j.visres.2007.05.011>
- McNab, F., & Dolan, R. J. (2014). Dissociating distractor-filtering at encoding and during maintenance. *Journal of Experimental Psychology: Human Perception and Performance*, 40(3), 960–967. <https://doi.org/10.1037/a0036013>
- McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, 11(1), 103–107. <https://doi.org/10.1038/nn2024>
- McNab, F., Zeidman, P., Rutledge, R. B., Smittenaar, P., Brown, H. R., Adams, R. A., & Dolan, R. J. (2015). Age-related changes in working memory and the ability to ignore distraction. *Proceedings of the National Academy of Sciences*, 112(20), 6515–6518. <https://doi.org/10.1073/pnas.1504162112>
- Mehta, M. A., Manes, F. F., Magnolfi, G., Schakian, B. J., & Robbins, T. W. (2004). Impaired set-shifting and dissociable effects on tests of spatial working memory following the dopamine D2 receptor antagonist sulpiride in human volunteers. *Psychopharmacology*, 176(3–4), 331–342. <https://doi.org/10.1007/s00213-004-1899-2>
- Nemes, V. A., Parry, N. R., Whitaker, D., & McKeefry, D. J. (2012). The retention and disruption of color information in human short-term visual memory. *Journal of Vision*, 12(1), Article 26. <https://doi.org/10.1167/12.1.26>
- Nemes, V. A., Whitaker, D., Heron, J., & McKeefry, D. J. (2011). Multiple spatial frequency channels in human visual perceptual memory. *Vision Research*, 51(23–24), 2331–2339. <https://doi.org/10.1016/j.visres.2011.09.003>
- Newton, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, 12(5), 436–450. [https://doi.org/10.1016/0022-1031\(76\)90076-7](https://doi.org/10.1016/0022-1031(76)90076-7)
- Nilsson, T. H., & Nelson, T. M. (1981). Delayed monochromatic hue matches indicate characteristics of visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, 7(1), 141–150. <https://doi.org/10.1037/0096-1523.7.1.141>
- Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, 142(7), 758–799. <https://doi.org/10.1037/bul0000046>
- Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, 144(9), 885–958. <https://doi.org/10.1037/bul0000153>
- Oberauer, K., & Lin, H. Y. (2017). An interference model of visual working memory. *Psychological Review*, 124(1), 21–59. <https://doi.org/10.1037/rev0000044>
- Opriš, I., Barborica, A., & Ferrera, V. P. (2005). Microstimulation of the dorsolateral prefrontal cortex biases saccade target selection. *Journal of Cognitive Neuroscience*, 17(6), 893–904. <https://doi.org/10.1162/0898929054021120>
- Oudeyer, P.-Y., & Kaplan, F. (2006). Discovering communication. *Connection Science*, 18(2), 189–206. <https://doi.org/10.1080/09540090600768567>
- Oudeyer, P.-Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(2), 265–286. <https://doi.org/10.1109/TEVC.2006.890271>
- Park, Y. E., Sy, J. L., Hong, S. W., & Tong, F. (2017). Reprioritization of features of multidimensional objects stored in visual working memory. *Psychological Science*, 28(12), 1773–1785. <https://doi.org/10.1177/0956797617719949>
- Pasternak, T., & Zaksas, D. (2003). Stimulus specificity and temporal dynamics of working memory for visual motion. *Journal of Neurophysiology*, 90(4), 2757–2762. <https://doi.org/10.1152/jn.00422.2003>
- Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. *Proceedings of the 34th International Conference on Machine Learning, in Proceedings of Machine Learning Research*, 70: 2778–2787. <https://doi.org/10.48550/arXiv.1705.05363>
- Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). Task context and organization in free recall. *Neuropsychologia*, 47(11), 2158–2163. <https://doi.org/10.1016/j.neuropsychologia.2009.02.013>
- Quinn, G. (1988). Interference effects in the visuo-spatial sketchpad. In M. Denis, J. Engelkamp, & J. T. E. Richardson (Eds.), *Cognitive and neuropsychological approaches to mental imagery* (Vol. 42 pp. 181–189). Springer.
- Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 41(6), 1650–1665. <https://doi.org/10.1037/xhp0000110>
- Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, 4(3), 193–202. <https://doi.org/10.1038/nrn1052>
- Röer, J. P., Bell, R., Dentale, S., & Buchner, A. (2011). The role of habituation and attentional orienting in the disruption of short-term memory performance. *Memory & Cognition*, 39(5), 839–850. <https://doi.org/10.3758/s13421-010-0070-z>
- Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354(6316), 1136–1139. <https://doi.org/10.1126/science.aaa7011>
- Rossi, A. F., Bichot, N. P., Desimone, R., & Ungerleider, L. G. (2007). Top-down attentional deficits in macaques with lesions of lateral prefrontal cortex. *Journal of Neuroscience*, 27(42), 11306–11314. <https://doi.org/10.1523/JNEUROSCI.2939-07.2007>
- Sanborn, A. N., & Chater, N. (2016). Bayesian Brains without probabilities. *Trends in Cognitive Sciences*, 20(12), 883–893. <https://doi.org/10.1016/j.tics.2016.10.003>
- Sanborn, A. N., & Chater, N. (2017). The sampling brain. *Trends in Cognitive Sciences*, 21(7), 492–493. <https://doi.org/10.1016/j.tics.2017.04.009>
- Savage, L. J. (1972). *The foundations of statistics*. Courier Corporation.
- Schurgin, M. W., Cunningham, C. A., Egeth, H. E., & Brady, T. F. (2018). Visual long-term memory can replace active maintenance in visual working memory. *BioRxiv*. <https://www.biorxiv.org/content/10.1101/381848v2.article-info>

- Schwan, S., & Garsoffsky, B. (2004). The cognitive representation of filmic event summaries. *Applied Cognitive Psychology*, 18(1), 37–55. <https://doi.org/10.1002/acp.940>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences*, 10(1), 38–45. <https://doi.org/10.1016/j.tics.2005.11.008>
- Sewell, G. L. (2014). *Quantum theory of collective phenomena*. Courier Corporation.
- Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, 283(5408), 1657–1661. <https://doi.org/10.1126/science.283.5408.1657>
- Smith, P. L., Lilburn, S. D., Corbett, E. A., Sewell, D. K., & Kyllingsbæk, S. (2016). The attention-weighted sample-size model of visual short-term memory: Attention capture predicts resource allocation and memory load. *Cognitive Psychology*, 89, 71–105. <https://doi.org/10.1016/j.cogpsych.2016.07.002>
- Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Pergamon Press.
- Sörqvist, P., Nösl, A., & Halin, N. (2012). Working memory capacity modulates habituation rate: Evidence from a cross-modal auditory distraction paradigm. *Psychonomic Bulletin & Review*, 19(2), 245–250. <https://doi.org/10.3758/s13423-011-0203-9>
- Sprague, T. C., Ithipuripat, S., Vo, V. A., & Serences, J. T. (2018). Dissociable signatures of visual salience and behavioral relevance across attentional priority maps in human cortex. *Journal of Neurophysiology*, 119(6), 2153–2165. <https://doi.org/10.1152/jn.00059.2018>
- Stokes, M. G. (2015). “Activity-silent” working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19(7), 394–405. <https://doi.org/10.1016/j.tics.2015.05.004>
- Suchow, J. W., Brady, T. F., Fougner, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, 13(10), Article 9. <https://doi.org/10.1167/13.10.9>
- Sullivan, B. T., Johnson, L., Rothkopf, C. A., Ballard, D., & Hayhoe, M. (2012). The role of uncertainty and reward on eye movements in a virtual driving task. *Journal of Vision*, 12(13), Article 19. <https://doi.org/10.1167/12.13.19>
- Sun, Y., Gomez, F., & Schmidhuber, J. (2011). Planning to be surprised: Optimal Bayesian exploration in dynamic environments. In J. Schmidhuber, K. R. Thórisson & M. Looks (Eds.), *Artificial general intelligence* (pp. 41–51). Springer. [https://doi.org/10.1007/978-3-642-22887-2\\_5](https://doi.org/10.1007/978-3-642-22887-2_5)
- Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, 138(2), 236–257. <https://doi.org/10.1037/a0015631>
- Tabi, Y. A., Maio, M. R., Fallon, S. J., Udale, R., Dickson, S., Idris, M. I., Nobis, L., Manohar, S. G., & Husain, M. (2021). Impact of processing demands at encoding, maintenance and retrieval in visual working memory. *Cognition*, 214, Article 104758. <https://doi.org/10.1016/j.cognition.2021.104758>
- Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, 73(1), 16–43. <https://doi.org/10.1037/h0022681>
- Tishby, N., & Polani, D. (2011). Information theory of decisions and actions. In V. Cuturidis, A. Hussain, & J. Taylor (Eds.), *Perception-action cycle* (pp. 601–636). Springer. [10.1007/978-1-4419-1452-1\\_19](https://doi.org/10.1007/978-1-4419-1452-1_19)
- Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature*, 401(6754), 699–703. <https://doi.org/10.1038/44372>
- Tresch, M. C., Sinnamom, H. M., & Seamon, J. G. (1993). Double dissociation of spatial and object visual memory: Evidence from selective interference in intact human subjects. *Neuropsychologia*, 31(3), 211–219. [https://doi.org/10.1016/0028-3932\(93\)90085-E](https://doi.org/10.1016/0028-3932(93)90085-E)
- Turatto, M., Bonetti, F., Pascucci, D., & Chelazzi, L. (2018). Desensitizing the attention system to distraction while idling: A new latent learning phenomenon in the visual attention domain. *Journal of Experimental Psychology: General*, 147(12), 1827–1850. <https://doi.org/10.1037/xge0000503>
- van den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, 109(22), 8780–8785. <https://doi.org/10.1073/pnas.1117465109>
- van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *eLife*, 7, Article e38677. <https://doi.org/10.7554/eLife.38677>
- Vogel, E. K., & Awh, E. (2008). How to exploit diversity for scientific gain: Using individual differences to constrain cognitive theory. *Current Directions in Psychological Science*, 17(2), 171–176. <https://doi.org/10.1111/j.1467-8721.2008.00569.x>
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984), 748–751. <https://doi.org/10.1038/nature02447>
- 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. <https://doi.org/10.1038/nature04171>
- Wan, Q., Cai, Y., Samaha, J., & Postle, B. R. (2020). Tracking stimulus representation across a 2-back visual working memory task. *Royal Society Open Science*, 7(8), Article 190228. <https://doi.org/10.1098/rsos.190228>
- Waters, W. F., McDonald, D. G., & Koresko, R. L. (1977). Habituation of the orienting response: A gating mechanism subserving selective attention. *Psychophysiology*, 14(3), 228–236. <https://doi.org/10.1111/j.1469-8986.1977.tb01166.x>
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, 4(12), Article 11. <https://doi.org/10.1167/4.12.11>
- Yu, Q., & Postle, B. R. (2018). Working memory: Separating the present and the future. *eLife*, 7, Article e43339. <https://doi.org/10.7554/eLife.43339>
- Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, 10(7), 301–308. <https://doi.org/10.1016/j.tics.2006.05.002>
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233–235. <https://doi.org/10.1038/nature06860>

Received June 3, 2021

Revision received November 25, 2022

Accepted December 9, 2022 ■