

# Prioritization Sharpens Working Memories but Does Not Protect Them From Distraction

Ziyao Zhang and Jarrod A. Lewis-Peacock  
Department of Psychology, University of Texas at Austin

Perceptual distraction distorts visual working memory representations. Previous research has shown that memory responses are systematically biased toward visual distractors that are similar to the memoranda. However, it remains unclear whether the prioritization of one working memory representation over another reduces the impact of perceptual distractors. In five behavioral experiments, we used different forms of retrospective cues (indicating the likelihood of testing each item and/or the reward for responding correctly to each item) to manipulate the prioritization of items in working memory before visual distraction. We examined the effects of distraction with nonparametric analyses and a novel distractor intrusion model. We found that memory responses were more precise (lower absolute response errors and stronger memory signals) for items that were prioritized. However, these prioritized items were not immune to distraction, and their memory responses were biased toward the visual distractors to the same degree as were unprioritized items. Our findings demonstrate that the benefits associated with prioritization in working memory do not include protection from distraction biases.

*Keywords:* working memory, mnemonic bias, perceptual interference, prioritization, distractor intrusion model

Visual working memory closely interacts with visual perception to influence behaviors. The sensory recruitment account posits these two processes are interdependent, such that information in working memory is represented in early visual cortices that are also actively involved in visual perception (D'Esposito, 2007; D'Esposito & Postle, 2015; Harrison & Tong, 2009; Pasternak & Greenlee, 2005; Postle, 2006, 2015; Serences et al., 2009). Working memory contents can alter visual perception and bias visual attention (Carlisle et al., 2011; Kang et al., 2011; Kiyonaga &

Egner, 2013; Olivers et al., 2006; Soto et al., 2008; Teng & Kravitz, 2019). Reciprocally, visual perception of distracting stimuli during maintenance detrimentally distorts working memory representations (Lorenc et al., 2018; Mallett et al., 2020; Rademaker et al., 2015; Teng & Kravitz, 2019; Van der Stigchel et al., 2007) and disrupts working memory task performance (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014; Marini et al., 2017; Nemes et al., 2012; Wildegger et al., 2015). When perceptual distractors are from the same feature space as memoranda, memory errors can show an attraction bias toward perceptual distractors, which might be dependent on the similarity between maintained information and distracting information (Lorenc et al., 2018; Mallett et al., 2020; Rademaker et al., 2015; Teng & Kravitz, 2019; Van der Stigchel et al., 2007). This finding has been replicated with simplistic features, such as colors, locations, and orientations, as well as complex stimuli, such as faces.

How information is maintained in working memory modulates its representational fidelity and its interaction with perceptual processes (Muhle-Karbe et al., 2021; Olivers et al., 2011; Stokes et al., 2020). Research over the past decade has suggested that information can be coded into working memory in different representational formats depending on its prioritization status. Studies using retro-cues to indicate the relevance (i.e., the likelihood of being tested) of working memory representations have shown that prioritized (i.e., task-relevant) representations are maintained via active coding with an observable and sustained neural signature. In contrast, unprioritized representations that are less relevant to the current task can be supported through distinct “activity-silent” mechanisms that do not require persistent neural activity (LaRocque et al., 2014, 2013; Lewis-Peacock et al., 2012; Myers et al., 2017; Rose et al., 2016; Stokes, 2015; Wolff et al., 2017). Such a distinction in coding formats between prioritized and unprioritized

This article was published Online First November 17, 2022.

Ziyao Zhang  <https://orcid.org/0000-0003-3219-9221>

Jarrod A. Lewis-Peacock  <https://orcid.org/0000-0002-9918-465X>

This work was completed with support from the National Institute of Health Grant R01EY028746 awarded to Jarrod A. Lewis-Peacock. This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank Elizabeth Lorenc for providing the face stimulus set and for helpful discussions on this project. We also thank Remington Mallett and Xuexin Wei for helpful comments on the project. Data has been shared at <https://osf.io/9bsqw/>. This work has not been presented publicly before the submission.

Jarrod A. Lewis-Peacock and Ziyao Zhang conceived of and designed the experiments. Ziyao Zhang implemented the experiment and collected the data. Ziyao Zhang analyzed the data and visualized results. Jarrod A. Lewis-Peacock and Ziyao Zhang wrote the original draft of the manuscript. Jarrod A. Lewis-Peacock and Ziyao Zhang reviewed and edited the final draft of the manuscript.

Correspondence concerning this article should be addressed to Jarrod A. Lewis-Peacock, Department of Psychology, University of Texas at Austin, 108 East Dean Keeton Street, Stop A8000, Austin, TX 78712, United States. Email: [jalewpea@utexas.edu](mailto:jalewpea@utexas.edu)

items might lead to differences in representation fidelity and how they interact with ongoing visual processes. It has been well established that the retention of prioritized representations is improved compared with unprioritized representations or neutral trials without retro-cues (retro-cue benefits; Griffin & Nobre, 2003; Lepsien et al., 2005; Jerke & Oberauer, 2013; van Moorselaar et al., 2015; for a review, see Souza & Oberauer, 2016). Prioritized (putatively active) representations can effectively guide perceptual decisions and cause shifts in external attention whereas unprioritized (putatively silent) representations have weaker and sometimes null effects on perceptual processes (Fritsche & de Lange, 2019; Mallett & Lewis-Peacock, 2018; Muhle-Karbe et al., 2021; Olivers et al., 2011; Stokes et al., 2020; van Loon et al., 2017; van Moorselaar et al., 2014; Wolff et al., 2017). The differential effects of prioritized and unprioritized working memory representations on visual processes suggest that a prioritization manipulation in working memory should moderate the interaction between working memory and visual perception.

However, it is unclear whether the prioritization of working memories provides protection against disruption from the perception of task-irrelevant information (perceptual interference effect). One group of studies examined the perceptual interference effect of distractors on visual working memory by inserting either a task-irrelevant mask or a secondary task during the maintenance phase of the primary memory task, but found mixed results. Some evidence has shown that prioritized memory representations are more resistant to interference from intervening visual masks (Makovski et al., 2008; Schneider et al., 2017; van Moorselaar et al., 2015), or an intervening secondary task (Makovski & Pertzov, 2015) compared with conditions where no item was prioritized in working memory, suggesting that prioritization could protect working memory representations from external distractions (Makovski & Pertzov, 2015; van Moorselaar et al., 2015). Others have shown that intervening tasks disrupted memory performance for prioritized and unprioritized representations to a similar degree (Hollingworth & Maxcey-Richard, 2013; Jerke et al., 2014).

Another group of studies examined the perceptual interference effect of visual distractor stimuli that are from the same feature space as the memoranda. Following the serial presentation of multiple memory items, the presentation of a visual distractor stimulus disrupted participants' memory performance. This disruption was greatest for the final (most recent) memory item, which was presumed to be prioritized over the earlier items in the list, suggesting that, contrary to prior accounts (Makovski & Pertzov, 2015; van Moorselaar et al., 2015), prioritized items may be particularly vulnerable to visual distraction (Hitch et al., 2018; Hu et al., 2014). Additionally, prioritized items that were associated with higher reward (more "points" earned for a correct response) also had larger distraction costs compared with unprioritized items that were associated with lower reward (Allen & Ueno, 2018, Experiments 2, 3, and 4; Hitch et al., 2018; but see Allen & Ueno, 2018, Experiment 1). In related work, prioritized working memory items that were probed in a change detection task during maintenance had larger interference costs later on, compared with unprioritized items that were not probed, in a subsequent memory recall test (Shepherdson, 2021). A recent functional magnetic resonance imaging (fMRI) study from our lab provided evidence consistent with this idea that higher priority items in working memory are more vulnerable to distraction. We found that lower priority working memory representations of faces and scenes in the ventral temporal cortex

recovered better from a distracting perceptual task compared with higher priority representations (Mallett & Lewis-Peacock, 2019). Together, convergent evidence suggests that prioritized representations in working memory, which are presumably coded in an active format, are particularly vulnerable to visual distractors compared with unprioritized representations, which are presumably coded in a passive format (for review see Lorenc et al., 2021).

In summary, prior studies examining perceptual interference effects could be generally separated into three camps: studies that support the *protection hypothesis* such that prioritization protects working memories from distraction (Makovski et al., 2008; Makovski & Pertzov, 2015; Schneider et al., 2017; van Moorselaar et al., 2015), studies that support the *vulnerability hypothesis* that prioritization renders working memories more vulnerable to distraction (Allen & Ueno, 2018, Experiments 2, 3, and 4; Hitch et al., 2018; Hu et al., 2014; Mallett & Lewis-Peacock, 2019), and studies that support the *null hypothesis* that prioritization does not moderate the vulnerabilities of working memories to distraction (Hollingworth & Maxcey-Richard, 2013; Jerke et al., 2014).

The discrepancies in existing results might be largely driven by differences in task designs. To reconcile those findings, it becomes important and necessary to systematically manipulate those task factors that could potentially modulate the perceptual interference effect of visual distraction. Upon reviewing prior studies, we isolated two main factors that could contribute to those discrepant findings: (a) the format of the retro-cue and (b) the format of the visual distraction. Most prior studies that support the vulnerability hypothesis used reward-based retro-cues (Allen & Ueno, 2018, Experiments 2, 3, and 4; Hitch et al., 2018; Hu et al., 2014). However, the effects of reward retro-cues and relevance retro-cues on distractibility have never been assessed in the same task paradigm. Thus, it remains unclear whether such findings could be generalized to relevance retro-cues, especially given that reward retro-cues and relevance retro-cues might lead to independent prioritization operations (Atkinson et al., 2018). Additionally, prior studies either stick to a passive viewing task or an active secondary task to induce the interference effect, but the two forms of visual distraction have never been assessed in the same experiment; thus, making it unclear whether they have differential effects on working memory.

In the current study, we investigated how visual distractors impact prioritized and unprioritized working memories. Across five experiments, we systematically manipulated the type of retro-cue (reward vs. relevance) and the type of distractor (passive vs. active) to directly contrast the three hypotheses regarding the perceptual interference effect. We adopted a delayed-estimation task with a method-of-adjustment response to directly measure participants' memory response errors (Mallett et al., 2020; Ma et al., 2014; Wilken & Ma, 2004; Zhang & Luck, 2008). Prior studies mainly measured the decline in recall accuracy in distraction trials as an indication of the perceptual interference effect. However, this measurement might not be sensitive enough to capture the subtle effect of distraction. Visual distractions could lead to systematic biases in participants' memory responses, without increasing memory errors in general (Rademaker et al., 2015). Those subtle differences in distraction effects could hold the key to unlock the mechanisms that underlie the perceptual interference effect. Based on findings that when a single representation is maintained in working memory, the intervening visual distractor leads to memory representations and memory responses being attracted

toward the distractor (Lorenc et al., 2018; Mallett et al., 2020; Rademaker et al., 2015; Teng & Kravitz, 2019; Van der Stigchel et al., 2007), we hypothesized that working memory representations should be generally biased toward visual distractors. If the *protection hypothesis* is true, we would expect to see memory responses for prioritized representations are unbiased or less biased compared with unprioritized representations (Figure 1a). If the *vulnerability hypothesis* is true, we would expect to see that the attraction bias toward distractor was enhanced in memory responses for prioritized items compared with unprioritized items (Figure 1b). Finally, If the *null hypothesis* is true then we expect to find prioritized and unprioritized representations are similarly vulnerable to distractions (Figure 1c).

## Method

### Participants

For the five main experiments, participants were recruited through Prolific. Demographic information was collected via free response boxes that were displayed at the beginning of the experiment. In total, 240 adults participated in the study (30 in Experiment 1, 50 in each of Experiments 2–4, 60 in Experiment 5; Age:  $M = 27.6$ ,  $SD = 9.3$ ; 88 women, 152 men). All experiments lasted approximately 70 min. Participants were provided online consent in redcap and received \$12/h in compensation. Ten participants were replaced in Experiment 5 because of a failure to comply with task instructions. For the control experiment, 27 participants were recruited from the undergraduate participant pool at the University of Texas at Austin (Age:  $M = 19.5$ ,  $SD = 1.4$ , 10 no reports; 18 women, 8 men, 1 no report). Participants were provided online consent and received course credits in compensation.

### Stimuli

Memory items were identical for all experiments. They were randomly selected from a continuous face space consisting of 80 gray-scale three-dimensional (3D) face images that varied along the dimensions of age and sex (Lorenc et al., 2014). Distances between faces were converted to a  $360^\circ$  for interpretability (distance between two faces =  $4.5^\circ$ ). Relevance cues were dark gray wheels (25.5,

25.5, 25.5). Reward cues were white numbers (255, 255, 255). Stimuli were presented in a gray background (127.5, 127.5, 127.5). Stimuli were presented using Pavlovia for online participants, and Stimuli were presented with Psychopy for in-person participants.

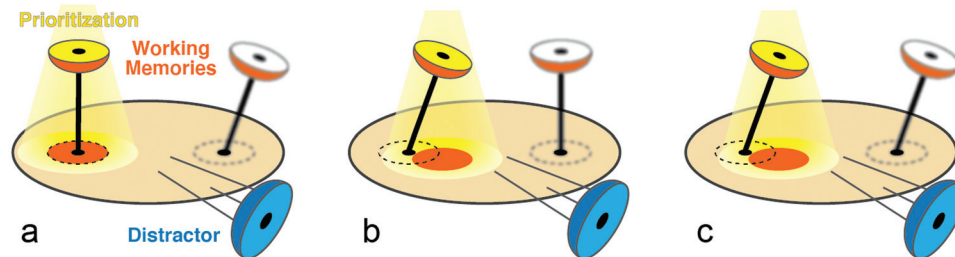
### Procedure

In all experiments, participants performed a delayed-estimation task with a method of adjustment response (see Figure 2).

#### Experiment 1: Load 2, Relevance Cue, and Passive Distraction

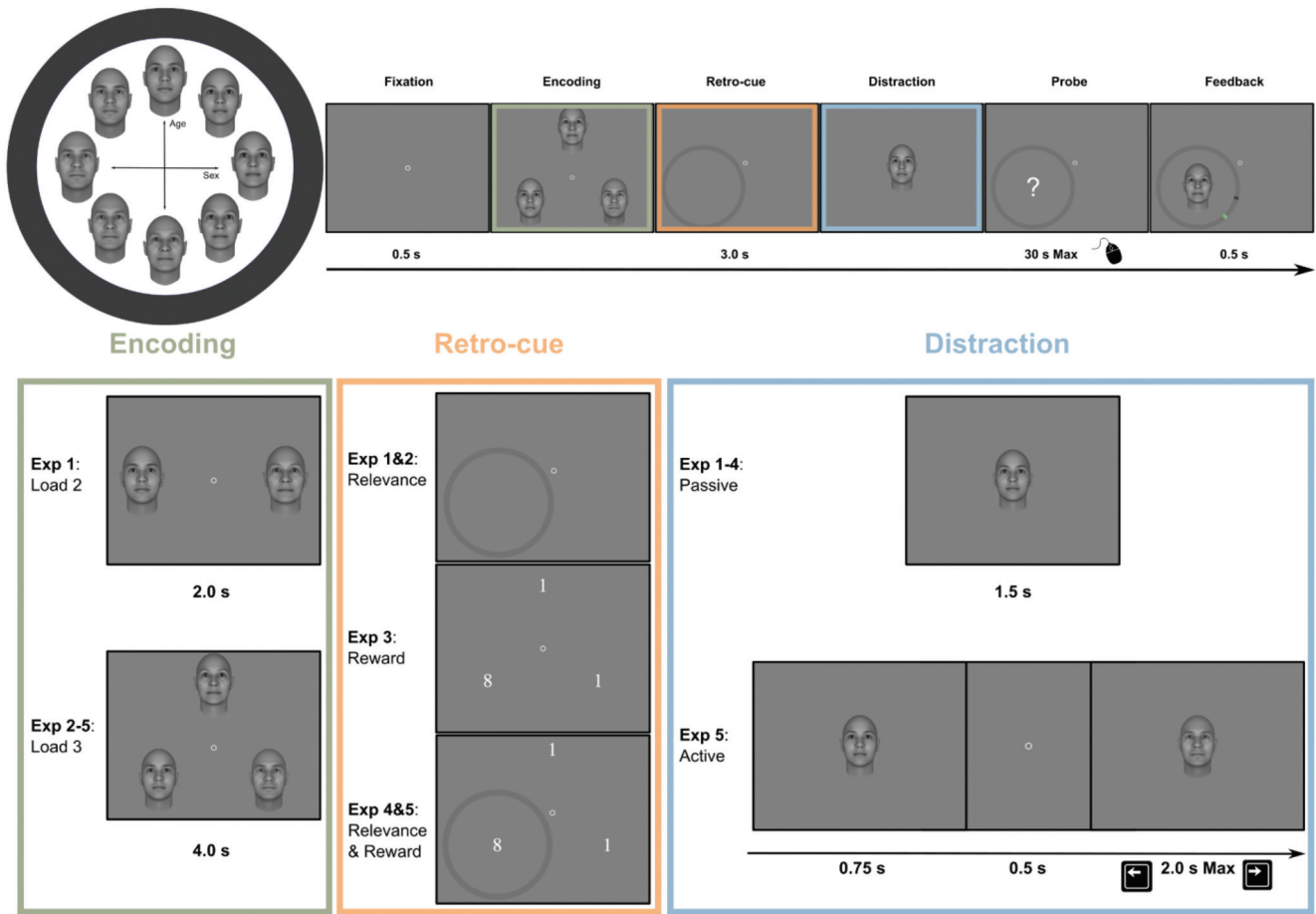
After a practice block of 24 trials, participants completed 10 experimental blocks, including 80 distraction absent trials and 160 distraction present trials. Each trial started with a central fixation (0.5 s,  $r = 20$  pixels). Then two memory faces were presented (2 s,  $w = 330$  pixels) at the left and right side of the screen, 300 pixels away from the center of the screen. The two memory faces were approximately  $157.5^\circ$  (34–36 faces) away from each other. After the encoding phase, a relevance cue was presented to be concentric to one of the memory faces (3 s,  $r$  of inner circle = 230 pixels,  $r$  of outer circle = 280 pixels). The relevance cue indicated that the cued face would highly likely be tested at the end of trial (on 75% of trials). In distraction-present trials, after the relevance cue, one distracting face would be presented at the center of the screen (1.5 s,  $w = 330$  pixels). The distracting face was clockwise (50% of trials) or counterclockwise of the tested face by approximately  $67.5^\circ$  (14–16 faces). Participants were instructed to ignore the distracting face while maintaining fixation. In distraction-absent trials, a fixation circle was presented instead. After a short delay (1 s), a response wheel was presented to be concentric to the tested face. In 75% of trials, the tested face was the cued face. In the other 25% of trials, the tested face was the uncued face. Participants first clicked on a question mark that was centric to the response wheel to start the response phase. During the response phase ( $\leq 30$  s), a face ( $w = 330$  pixels) that was centric to the response wheel morphed continuously as the cursor moved around the wheel. The orientation of the face space was rotated randomly along the wheel in each trial, and the morphable face was not presented until the cursor made contact with the wheel. Following response via mouse click, feedback was provided, with a green indicator of the correct location of the tested item on the response wheel (0.5 s). Trials ended with a 1 s blank ITI. Blocks were separated by a 15 s break.

**Figure 1**  
*Hypothesized Models of Perceptual Distraction Effect*



*Note.* (a) Protection hypothesis: prioritized representations have high memory fidelity (sharp not fuzzy) and are protected from visual distraction (vertical not tilted). (b) Vulnerability hypothesis: prioritized representations have high memory fidelity but are vulnerable to visual distraction. (c) Null hypothesis: prioritized representations have high memory fidelity, but are similarly vulnerable to visual distraction as unprioritized representations. See the online article for the color version of this figure.

**Figure 2**  
Illustrations of Experimental Procedures for Experiments 1–5



Note. See the online article for the color version of this figure.

**Experiment 2: Load 3, Relevance Cue, and Passive Distraction**

To investigate the potential effects of working memory load on the perceptual interference effect, we increased the memory load from two items to three items. The increase in memory load should make the strategy of actively maintaining all presented items less likely and should encourage participants to prioritize cued items, as the practical benefits of prioritization were enhanced. The cue validity was increased from 75% to 80%. In order for participants to have sufficient encoding time, the duration of the encoding phase was increased to 4 s. The three memory faces were approximately 121.5° (26–28 faces) away from each other. To create an empirical baseline condition for potential prioritization effects in Experiment 2, we included 15% no-cue trials to replace the no-distractor trials from Experiment 1. In no-cue trials, no prioritization effect should be expected. This change was also applied to Experiments 3–5. In the other 85% of trials, we provided relevance cues. In 80% of relevance-cue trials, the tested face was the cued face. In the other 20% of relevance-cue trials, the tested face was one of the two uncued faces. In no-cue trials, the three memory faces were equally likely to be tested. Finally, participants completed eight experimental blocks of 24 trials after a practice block. All other experimental designs were the same as Experiment 1.

**Experiment 3: Load 3, Reward Cue, and Passive Distraction**

To generalize our findings in Experiments 1 and 2 to other prioritization manipulations, in Experiment 3, the relevance manipulation was replaced by a reward manipulation. After the encoding phase, three numbers were presented at the locations of memory faces, indicating reward points associated with each of the memory faces. In 85% of trials, there were unequal reward where one memory face was associated with eight reward points, whereas the remaining two faces were associated with one reward point. Based on prior literature, the difference in reward points would create a difference in prioritization status such that the face associated with eight reward points would be prioritized above the faces associated with only one reward point. The remaining 15% of trials had equal reward where all three faces were associated with one reward point, so they should have similar prioritization status. During the memory test, all three faces had equal probabilities of being tested, regardless of how many reward points they were associated with. During ITIs, participants received feedback about the reward points they have earned in a trial. During breaks, participants received feedback about overall reward points they have earned and the potential maximum. Participants were instructed to work on maximizing their reward points. Finally, participants completed

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.



eight experimental blocks of 24 trials after a practice block. All other experimental designs were the same as Experiment 1.

#### **Experiment 4: Load 3, Relevance and Reward Cue, and Passive Distraction**

Prior research has shown that reward and relevance retro-cues can produce additive benefits in recall responses (Atkinson et al., 2018). To further enhance potential prioritization effects and examine its potential interaction on distraction protection, in Experiment 4, we combined the relevance manipulation with the reward manipulation. In 85% of trials, one memory face was associated with both a relevance cue and a high reward of 8 points, whereas the two remaining faces were associated with only a low reward of one point. In 15% of trials, no relevance cue was presented, and all memory faces were associated with one reward point. During the memory test, in 80% of relevance cue trials, the tested face was the cued face. In the other 20% of relevance cue trials, the tested face was one of the two uncued faces. In no relevance cue trials, the three memory faces were equally likely to be tested. Finally, participants completed eight experimental blocks of 24 trials after a practice block. All other experimental designs were the same as Experiment 1.

#### **Experiment 5: Load 3, Relevance and Reward Cue, and Active Distraction**

To generalize our findings to other forms of visual distractions and test how working memory copes with interrupting tasks that are more engaging, in Experiment 5, the passive viewing of a distracting face was replaced by an active delay match-to-sample task. Following retro-cues, an initial distracting face was presented for 0.75 s. After a short delay of 0.5 s, a test distracting face was presented for 2 s. Participants needed to determine whether the test distracting face was different from the initial distracting face and report by pressing the left or right arrow keys on the keyboard to indicate the same or different judgments. In 50% of trials, the test distracting face was 99° (21–23 faces) away from the initial distracting face, whereas in the other 50% of trials, the test distracting face was identical to the initial distracting face. Finally, participants completed eight experimental blocks of 24 trials after a practice block. All other experimental designs were the same as Experiment 4.

#### **Control Experiment: Load 1, No Cue, and No Distraction**

The control experiment was conducted to derive the psychological scaling function for the Target Confusability Competition (TCC) model. Participants completed eight blocks of 24 trials. Each trial started with a central fixation (0.5 s). Then one memory face was presented (1 s,  $w = 10^\circ$ ) at the center of the screen. After a brief delay (1 s), participants reported the memory face with the continuous response wheel. Trials ended with a 1 s blank ITI.

### **Analysis**

#### **Nonparametric Analysis**

Memory errors were computed as the difference between the response and the tested face in the continuous face space (as in Mallett et al., 2020). Absolute memory errors were extracted and averaged for each participant. To quantify biases, absolute memory errors were assigned with a positive sign if they were in the same

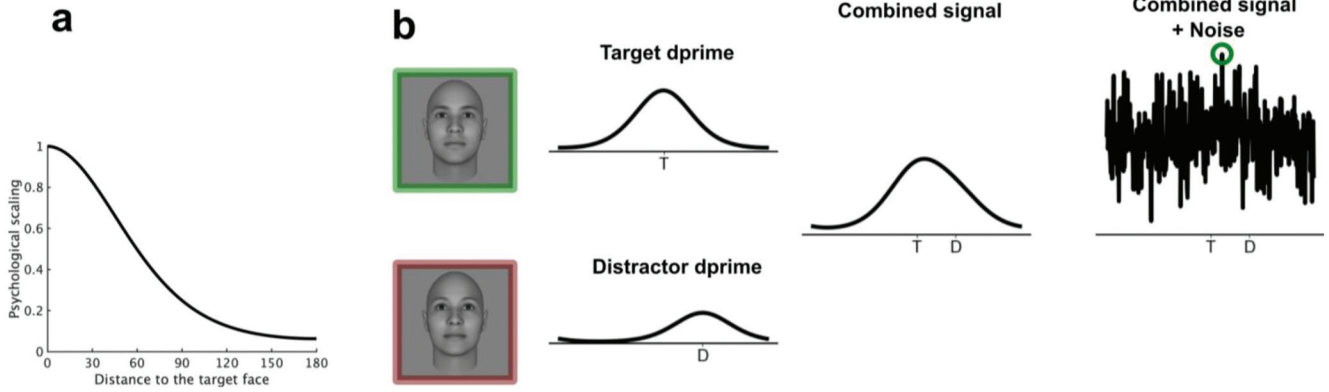
direction as distracting faces, but were assigned with a negative sign if they were in the opposite direction as distracting faces. The response bias was computed by averaging signed errors. A positive bias (attraction bias) indicates that participants' responses were biased toward the distracting faces, whereas a negative bias (repulsion bias) indicates that participants' responses were biased away from the distracting faces. To evaluate the presence of response biases, we performed one-sample *t* tests against zero for each of the tested item conditions. To compare biases between different tested item conditions, we performed repeated-measures analysis of variances (ANOVAs). All statistical tests were two-tailed. Sidak corrections were applied for multiple comparisons, and corrected *p* values were reported. For each experiment, we removed participants whose mean absolute memory errors were beyond 1.5 *SDs* from the group mean, and whose mean biases were larger than 50°. Those participant rejection procedures were used to remove participants who might not follow the task instruction in our online experimental settings. We removed participants who had fewer than 15 trials in any of the experimental conditions. In Experiment 5 specifically, we also removed participants whose accuracy in the distracting task was below 1.5 *SDs* from the mean (Experiment 1: *N* = 9, Experiment 2: *N* = 15, Experiment 3: *N* = 12, Experiment 4: *N* = 8, Experiment 5: *N* = 15).

#### **TCC Model With Signal Intrusion**

We adopted a revised version of target confusability competition model (TCC) to better quantify memory strength of targets and intruded signal strength of distractors. The distractor signal intrusion model is primarily based on recent evidence that error distribution from continuous reports can be quantified by a single parameter of memory strength (target  $d'$ ) when the shape of the memory signal is fixed by psychological similarity function of the feature space (For full description of the TCC model, see Schurgin et al., 2020). Based on the sensory recruitment hypothesis and associated findings that working memory representations overlap with sensory representations (D'Esposito, 2007; D'Esposito & Postle, 2015; Harrison & Tong, 2009; Pasternak & Greenlee, 2005; Postle, 2006, 2015; Serences et al., 2009), we added a distractor signal in the original TCC model to induce interference between target and distractor signals (see Figure 3b). We made the assumption that the distractor signal should be determined by the same similarity scaling function as the target signal when they are from the same feature space. Therefore, in our distractor signal intrusion model, there is one parameter of target  $d'$  to quantify strength of target signal, and one parameter of distractor  $d'$  to quantify strength of distractor signal. The original TCC model utilizes a psychological similarity function (Laplace distribution) generated from similarity judgment responses to describe the probabilistic memory signal (Schurgin et al., 2020). However, recent modeling comparisons across multiple studies have shown that TCC models with a von Mises distribution as the signal function have superior fits to error distributions compared with the original TCC model that used a Laplace distribution (Oberauer, 2021). Therefore, we replaced the psychological similarity function (Laplace distribution) with a von Mises distribution with a parameter  $k$  to describe the activation function of both the memory signal and the distractor signal in the current model. The parameter  $k$  (concentration of the signal) is determined by the property of the feature space should not be varied by subjects. Consistent with this

**Figure 3**

(a) Target Confusability Competition (TCC) Model



*Note.* (a) Psychological scaling function estimated from the control experiment. (b). Illustration of the signal component of the Target Confusability Competition (TCC) model. The transformation from signal function to response distribution is based on the signal detection rule. In a given trial, the face that has the strongest familiarity signal (green circle) is selected as the target face based on the distractor intrusion model. Face stimuli was adapted from “Expertise for upright faces improves the precision but not the capacity of visual working memory,” by E. S. Lorenc, M. S. Pratte, C.F. Angeloni, & F. Tong, 2014, *Attention, Perception & Psychophysics*, 76(7), 1977. Copyright 2014 by Psychonomic Society, Inc. Adapted with permission. See the online article for the color version of this figure.

argument, recent model comparisons showed parameter  $k$  can be fixed across experimental conditions (Oberauer, 2021; Schurgin et al., 2020). Therefore, we created a TCC model with only the target signal component to fit to the group data from the control experiment without cues or visual distractions, and estimated the parameter  $k$ . We fixed the parameter  $k$  for all models we applied here (see Figure 3a for the scaling function,  $k = 1.383$ ).

Based on this distractor signal intrusion model, on each trial, the to-be-remembered face is boosted by a familiarity signal (target  $d'$ ), and the similarity signal decreases roughly exponentially as a function of distance in the face space (Fechner’s law). Therefore, faces that were close to the to-be-remembered one get larger boosts in familiarity signals compared with faces that were further away. Similarly, due to the overlap between sensory representations of distractions and working memory representations, the distracting face also gets a boost in familiarity signals (distractor  $d'$ ), which decreases along the face distance axis as do the target signals. Therefore, in the model, the familiarity signal that guides reports is a combination of a weighted target signal and a weighted distractor signal. Finally, familiarity signals are corrupted by random noise. Formally, the distractor signal intrusion model can be described by the following equations.

$$T(x) = d'_T \frac{\exp(k \cos(x - \theta))}{2\pi I_0(k)} \quad (1)$$

$$D(x) = d'_D \frac{\exp(k \cos(x - \theta))}{2\pi I_0(k)} \quad (2)$$

$$S = T(x) + D(x) \quad (3)$$

Here  $x$  is the potential feature value of the activation functions  $T(x)$  and  $D(x)$ .  $T$  is the activation function of the target signal, and  $D$  is the activation function of the distractor signal. Both signal functions are von-Mises distribution functions with the mean  $\theta$  ( $\theta = 0$  in error distributions) and concentration  $k$ . These functions

are multiplied by signal strength  $d'$ . The final signal is the sum of target and distractors signals.

$$r = \arg \max(S + \epsilon), \text{ with } \epsilon \sim N(0, 1) \quad (4)$$

Noise was drawn from a standard normal distribution and added to the signal. Then the signal distribution was transformed into a response distribution via a signal-detection rule. That is, in a given trial, the face that has the strongest “memory + noise” signal will be selected as the final response  $r$ .

Model fitting was performed with MemToolbox (Suchow et al., 2013) and custom MATLAB scripts. The distractor signal intrusion model was fitted separately to the group data from each of the experiments to derive estimation of the two parameters, target  $d'$  and distractor  $d'$ . For each model fit, we used the 15,000 postconvergence samples to calculate the 95% credible interval (CI). The 95% CI indicates that the true parameter value has a 95% probability to be within this interval. To compare posteriors between conditions, we compute differences between posterior samples and then calculate the 95% CI. Credible differences between conditions are found when the 95% CI of the difference posterior does not overlap with zero.

### Model Parameter Recovery

Data simulations were performed with MemToolbox (Suchow et al., 2013) and custom MATLAB scripts. We used the estimated probabilistic signal (von Mises distribution,  $k = 1.383$ ) to construct memory and distractor signals. Both memory  $d'$  and  $d'$  distractor  $d'$  were randomly sampled between 0 and 1.5 with a step of .1 for each simulation. For each target-distractor distance between  $10^\circ$  and  $180^\circ$  with a step of  $10^\circ$ , we randomly sampled  $d'$ s 1,000 times. For each of the 1,000 simulations, we simulated 1,000 trials of data. We then fit the simulated data back to a TCC model with signal intrusion and estimated both the target and distractor  $d'$ . Biases in parameter recovery were computed as the differences between the recovered parameters and the true parameters.

**Results**

**Absolute Memory Errors**

The absolute memory errors across all experiments are shown in Figure 4. Retro-cue benefits were consistently found in all experiments except for Experiment 3 where the reward manipulation was employed. Specifically, in Experiment 1, absolute memory errors were smaller for prioritized items compared with unprioritized items, leading to a main effect of tested item in a repeated-measures ANOVA with factors of distraction (distraction-present trials, distraction-absent trials), and tested target (prioritized item, unprioritized item),  $F(1, 21) = 6.00, p = .023, \eta_p^2 = .22$ . No significant main effects of distraction or interactions were found. In Experiment 2, a main effect of tested item was found,  $F(2, 68) = 4.34, p = .017, \eta_p^2 = .11$ . Follow-up  $t$  tests revealed that absolute memory errors for prioritized items were smaller compared with unprioritized items,  $t(34) = -2.41, p = .063, d = .41$ . Absolute memory errors for neutral control items were also smaller compared with unprioritized items,  $t(34) = -2.32, p = .078, d = .39$ . In Experiment 3, no main effect of the tested item was found,  $F(2, 74) = 2.04, p = .137$ . In Experiment 4, a main effect of tested item was found,  $F(2, 82) = 12.90, p < .001, \eta_p^2 = .24$ . Absolute memory errors for prioritized items were smaller compared with unprioritized items or neutral controls,  $t(41) = -4.82, p < .001, d = .74; t(41) = -2.71, p = .029, d = .42$ . Absolute memory errors for neutral controls were also smaller compared with unprioritized items,  $t(41) = -2.52, p = .046, d = .39$ . Similarly, in Experiment 5, a main effect of tested item was found,  $F(2, 68) = 18.54, p < .001, \eta_p^2 = .35$ . Absolute memory errors for prioritized items were smaller compared with unprioritized items or neutral controls,  $t(34) = -4.94, p < .001, d = .83$ ;

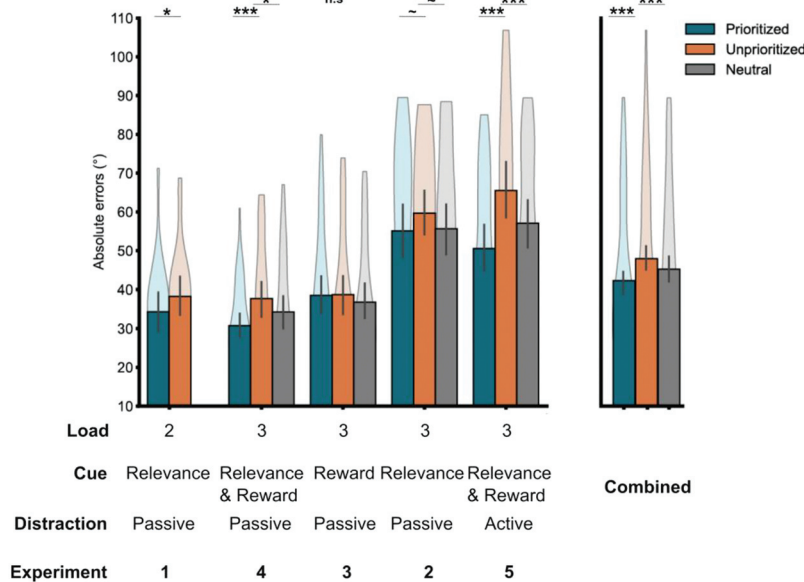
$t(34) = -2.93, p = .018, d = .50$ . Absolute memory errors for neutral controls were also smaller compared with unprioritized items,  $t(34) = -4.19, p < .001, d = .71$ .

**Memory Biases**

The memory bias results across all experiments are shown in Figure 5. Significant attraction biases were consistently found in distraction present trials for prioritized targets, Experiment 1:  $M = 4.1, t(21) = 2.58, p = .017, d = .55$ ; Experiment 2:  $M = 7.6, t(34) = 3.75, p < .001, d = .63$ ; Experiment 3:  $M = 3.9, t(37) = 2.32, p = .026, d = .38$ ; Experiment 4:  $M = 3.6, t(41) = 5.65, p < .001, d = .87$ ; Experiment 5:  $M = 11.6, t(34) = 10.37, p < .001, d = 1.75$ . For unprioritized items, significant attraction biases were found in Experiments 1, 2, 3, and 5, Experiment 1:  $M = 8.4, t(21) = 4.17, p < .001, d = .89$ ; Experiment 2:  $M = 4.8, t(34) = 2.19, p = .035, d = .37$ ; Experiment 3:  $M = 3.2, t(37) = 2.29, p = .028, d = .37$ ; Experiment 4:  $M = 3.0, t(41) = 1.62, p = .112, d = .25$ ; Experiment 5:  $M = 7.0, t(34) = 2.77, p = .009, d = .47$ . For neutral control items, we found significant attraction biases in Experiments 2 and 5, and marginally significant attraction biases in Experiments 3 and 4, Experiment 2:  $M = 9.1, t(34) = 3.15, p = .003, d = .53$ ; Experiment 3:  $M = 3.7, t(37) = 1.73, p = .092, d = .28$ ; Experiment 4:  $M = 3.0, t(41) = 1.92, p = .062, d = .30$ ; Experiment 5:  $M = 10.3, t(34) = 3.88, p < .001, d = .66$ .

Averaged memory biases across five experiments were computed. Significant attraction biases were found for all prioritized, unprioritized, and neutral targets, prioritized:  $t(171) = 8.98, p < .001, d = .68$ ; unprioritized,  $t(171) = 5.38, p < .001, d = .41$ ; neutral,  $t(149) = 5.40, p < .001, d = .44$ .

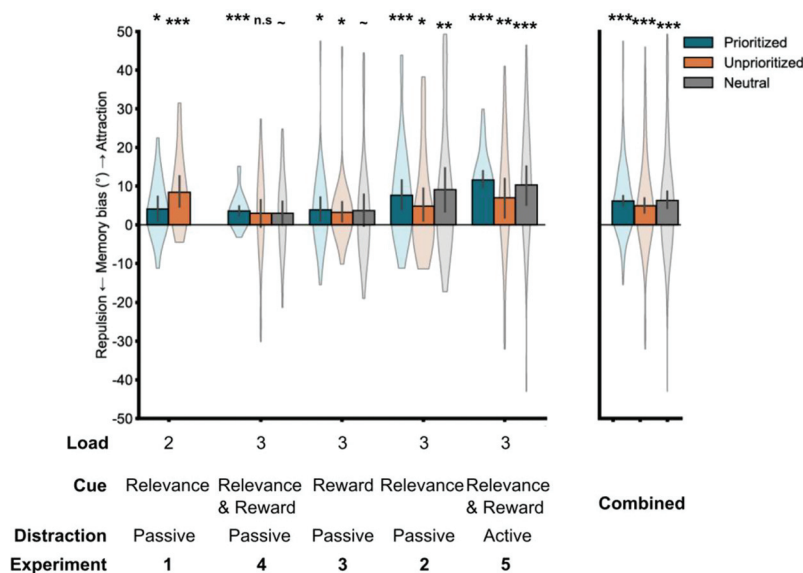
**Figure 4**  
Absolute Memory Errors From Distraction Trials in All Five Experiments



Note. Error bars indicate 95% confidence interval. Violin plots show individual mean data distributions. See the online article for the color version of this figure.  
\* indicates that  $p < .05$ . \*\*\* indicates that  $p < .001$ .

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

**Figure 5**  
Memory Biases From Distraction Trials in All Five Experiments



*Note.* Responses were biased toward distraction faces across tested targets and experiments. Error bars indicate 95% confidence intervals. Violin plots show individual mean data distributions. See the online article for the color version of this figure.

\* indicates that  $p < .05$ . \*\* indicates that  $p < .01$ . \*\*\* indicates that  $p < .001$ .

To compare differences in memory biases between tested targets, in Experiment 1, a repeated-measures ANOVA was conducted with factors of distraction (distraction present trials, distraction absent trials), and tested target (prioritized item, unprioritized item). Larger attraction biases were found in distraction present trials compared with distraction absent trials, leading to a main effect of distraction,  $F(1, 21) = 11.20, p = .003, \eta_p^2 = .34$ . No main effects of tested items or interactions were found. In Experiment 2–5, repeated measures ANOVA were conducted with a factor of tested target (prioritized item, unprioritized item, and neutral item). No significant main effects were found, suggesting that distraction led to similar level of attraction biases for working memory representations regardless of its prioritization status. Specifically, there were similar biases in memory reports for prioritized items and neutral trials (Experiment 2:  $t(34) = -.747, p = .843, BF_{01} = 4.255$ ; Experiment 3:  $t(37) = .088, p = .999, BF_{01} = 5.714$ ; Experiment 4:  $t(41) = .389, p = .973, BF_{01} = 5.587$ ; Experiment 5:  $t(34) = .489, p = .948, BF_{01} = 4.926$ ). There were similar biases in memory reports for unprioritized items and neutral trials (Experiment 2:  $t(34) = -1.577, p = .328, BF_{01} = 1.792$ ; Experiment 3:  $t(37) = -.238, p = .993, BF_{01} = 5.587$ ; Experiment 4:  $t(41) = -.001, p = 1.000, BF_{01} = 5.988$ ; Experiment 5:  $t(34) = -.967, p = .713, BF_{01} = 3.584$ ). Finally, there were similar biases in memory reports for prioritized and unprioritized items across experiments (Experiment 1:  $t(21) = -2.27, p = .067, BF_{01} = .550$ ; Experiment 2:  $t(34) = 1.41, p = .422, BF_{01} = 2.222$ ; Experiment 3:  $t(37) = .467, p = .955, BF_{01} = 5.181$ ; Experiment 4:  $t(41) = .322, p = .749, BF_{01} = 5.714$ ; Experiment 5:  $t(34) = 1.779, p = .232, BF_{01} = 1.335$ ). Bayes factors for the null hypothesis ( $BF_{01}$ ) between one and three indicate anecdotal evidence, while factors between three and 10 are considered moderate evidence for the null

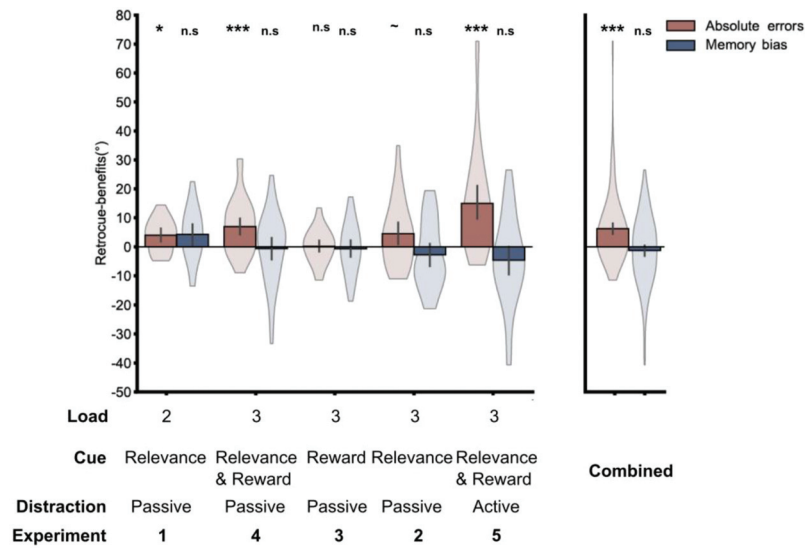
hypothesis. To increase the statistical power to detect potential differences in biases between conditions, we combined data from Experiments 2, 4, and 5 where we found reliable retro-cue benefits in absolute errors and investigated the effect of prioritization on biases. However, there was still no significant main effect of test item in biases, Experiments 2, 4, and 5:  $F(2, 214) = 1.814, p = .168, \eta_p^2 = .02, BF_{01} = 5.846$ . There were similar biases in responses for prioritized and neutral items,  $t(109) = -.22, p = .994, BF_{01} = 9.259$ ; similar biases in responses for unprioritized and neutral items,  $t(109) = -1.49, p = .361, BF_{01} = 3.226$ ; and similar biases in responses for prioritized and unprioritized items,  $t(109) = -1.63, p = .287, BF_{01} = 2.638$ .

### Separable Retro-Cue Effects in Absolute Errors and Biases

To better visualize retro-cue effects, the differences between prioritized and unprioritized items (i.e., the retro-cue benefit) in both absolute errors and biases were computed (see Figure 6). This analysis highlights the separable retro-cue effects. Relevance retro-cues consistently led to smaller absolute errors for prioritized items compared with unprioritized items even in the face of visual distractions, suggesting working memory prioritization survived through visual distractions. Internal attention was shifted to cued items and sharpened cued representations (Griffin & Nobre, 2003; Landman et al., 2003; Lepsien et al., 2005; Rerko & Oberauer, 2013; for a review, see Souza & Oberauer, 2016; ). However, no significant retro-cue effects were found in bias measurements, suggesting that prioritization did not render working memory representations particularly resistant or vulnerable to visual distractions.



**Figure 6**  
*Retro-Cue Benefits in Absolute Errors but Not in Bias Measurements*



*Note.* Consistent retro-cue benefits following relevance cues in absolute memory errors, but no retro-cue benefits in memory bias measurements. Error bars indicate 95% confidence interval. Violin plots show individual mean data distributions. See the online article for the color version of this figure.

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

## Ruling out Alternative Strategies

In all our experiments, the distance between distracting faces and tested faces were fixed. This leads to the possibility that participants could learn to use the distractor information to inform their responses, such as selecting a target face that is nearby the distracting face. If participants were indeed using this strategy, we would expect that their responses would be centered around the distractor faces, which would lead to large mean biases. Our rejections of subjects whose mean bias was larger than 50 degrees should help reject participants who may have adopted this strategy, because distractors were 67.5 degrees away from the tested item. If this strategy was being used by the remaining participants, it likely would have developed across the experiment, and we would expect to see larger distractor biases in the second half of trials compared with the first half of trials. To test this idea, we split trials into first and second halves and examined the potential evolution of biases. Repeated-measures ANOVAs with factors of tested item (prioritized, unprioritized, and neutral) and trial order (first half, second half) were conducted. Across experiments, we did not find any differences in biases between the first and second half of trials. Experiment 1:  $F(1, 21) = .00$ ,  $p = .969$ ,  $\eta_p^2 = .00$ ,  $BF_{01} = 4.484$ ; Experiment 2:  $F(1, 34) = 1.80$ ,  $p = .189$ ,  $\eta_p^2 = .05$ ,  $BF_{01} = 2.427$ ; Experiment 3:  $F(1, 37) = 1.41$ ,  $p = .243$ ,  $\eta_p^2 = .04$ ,  $BF_{01} = 2.985$ . Experiment 4:  $F(1, 41) = .12$ ,  $p = .729$ ,  $\eta_p^2 = .00$ ,  $BF_{01} = 5.650$ ; Experiment 5:  $F(1, 34) = .71$ ,  $p = .888$ ,  $\eta_p^2 = .00$ ,  $BF_{01} = 5.464$ . No significant interactions between the tested item and the trial order were found. These results indicate that there were no systematic differences in biases between the first and second half of trials and it is unlikely that participants adopted a distractor-anchoring response strategy.

## Model Parameter Recovery

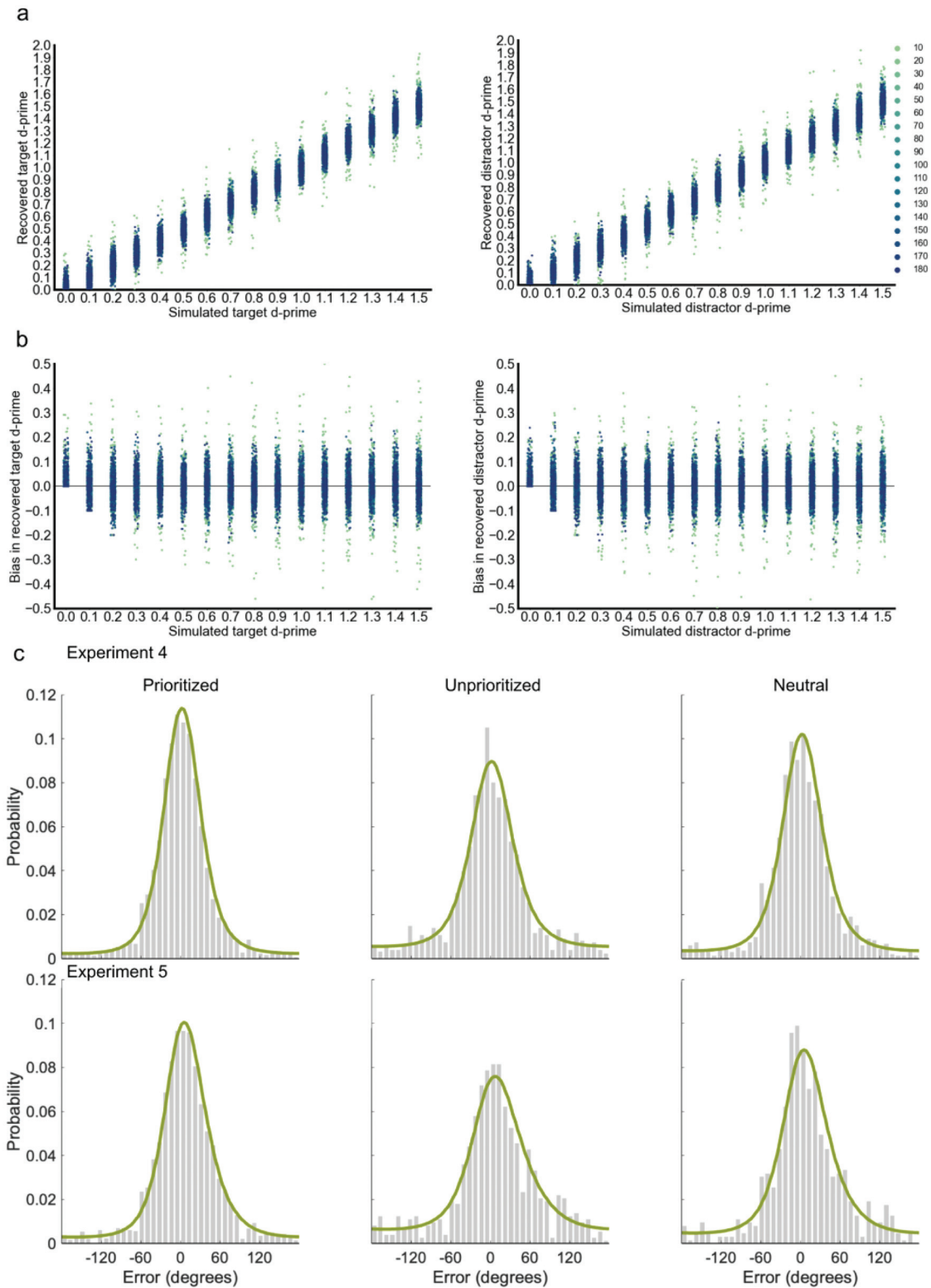
Both target  $d'$  and distractor  $d'$  that were recovered from the model were almost perfectly correlated with simulated values (see Figure 7a,  $r = .99$ ,  $p < .001$ ;  $r = .99$ ,  $p < .001$ ), indicating that the model reliably recovered these key simulated parameters. Furthermore, there was no significant correlation between the two recovered parameters, suggesting there was no systematic tradeoffs between these parameters in the model,  $r = -.01$ ,  $p = .249$ . Finally, there were no systematic biases in either of the two recovered parameters (see Figure 7b).

## Target Memory Strength

Maximum a posteriori (MAP) and 95% CI of posterior distributions were reported in Table 1. Relevance retro-cues consistently led to strengthened memory signals for prioritized items compared with unprioritized items (Experiment 1, Difference CI [.160, .324]; Experiment 2, Difference CI [.095, .267]). Memory signals for unprioritized items were weakened compared with the neutral condition (Experiment 2, Difference CI [-.020, -.254]). However, no credible difference was found between the prioritized condition and neutral condition, Difference CI [-.051, .138].

Similarly, reward retro-cues led to strengthened memory signals for prioritized items compared with unprioritized items (Experiment 3, Difference CI [.037, .157]). However, no credible difference was found between the prioritized condition and the neutral condition (Experiment 3, Difference CI [-.037, .139]), or between the neutral condition and the unprioritized condition (Experiment 3, Difference CI [-.031, .124]).

**Figure 7**  
*Model Parameter Recovery*



*Note.* (a) Recovered parameters were positively correlated with simulated parameters across target-distractor distances (b) No systemic biases were observed in recovered parameters. (c) Experiment 4 is on the top row, and Experiment 5 is on the bottom row. Line plots represent model predictions based on fitted parameters (maximum a posteriori, MAP). Bar plots represent group-level error distributions. Model predictions fit well with the empirical data. See the online article for the color version of this figure.

**Table 1**  
*Target d'*

	Relevance cue						Reward cue						Relevance and reward cue					
	Experiment 1		Experiment 2		Experiment 3		Experiment 4		Experiment 5		Experiment 4		Experiment 5		Experiment 4		Experiment 5	
	P	UP	P	N	UP	N	P	N	UP	N	P	N	UP	N	P	N	UP	N
MAP	1.141	0.892	1.120	1.067	0.942	0.980	1.041	0.980	0.938	1.163	1.343	1.163	0.987	0.973	1.164	0.973	0.799	0.799
CI	[1.095, 1.176]	[0.830, 0.966]	[1.077, 1.162]	[0.989, 1.162]	[0.864, 1.014]	[0.917, 1.054]	[0.993, 1.097]	[0.917, 1.054]	[0.906, 0.978]	[1.315, 1.379]	[1.105, 1.230]	[1.119, 1.213]	[0.926, 1.064]	[0.902, 1.054]	[1.119, 1.213]	[0.902, 1.054]	[0.721, 0.881]	[0.721, 0.881]

*Note.* MAP = maximum a posteriori; CI = 95% credible interval; P = prioritized; N = neutral; NP = unprioritized. Cuing consistently led to stronger memory signals for prioritized items compared with unprioritized items.

Combined retro-cues (relevance and reward cue) led to strengthened memory signals for cued items compared with either the unprioritized condition (see Figure 8a, Experiment 4, Difference CI [.285, .426]; Experiment 5, Difference CI [.272, .454]) or the neutral condition (Experiment 4, Difference CI [.109, .247]; Experiment 5, Difference CI [.096, .274]). Additionally, Memory signals for unprioritized items were weakened compared with the neutral condition (Experiment 4, Difference CI [-.087, -.272]; Experiment 5, Difference CI [-.069, -.294]).

**Distractor Signal Strength**

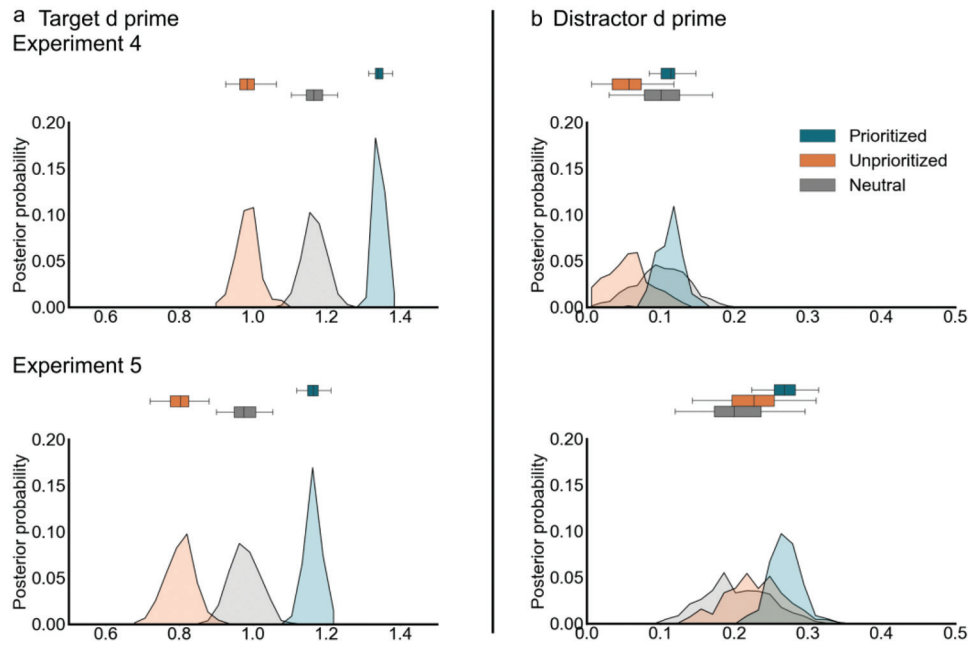
To ensure the intrusion model is sensitive to changes in distractor strength, we compared distractor *d'* between Experiment 4 where distractions were passively viewed and Experiment 5 where distractions were actively maintained in working memory. All other aspects of designs were consistent, so any changes in distractor *d'* between the two experiments should be attributed to changes in visual distractions. We found that active distractions consistently led to larger distractor *d'* compared with passive distractions across prioritization conditions (see Figure 8b; Prioritized: Difference CI [.104, .210]; Unprioritized: Difference CI [.071, .266]; No cue: Difference CI [.003, .216]), demonstrating that the distractor intrusion model was sensitive to distractor strength changes. However, across experiments, we did not find any evidence that distractor strength was modulated by prioritization (see Table 2). In other words, intrusions of distractions were similarly strong for the prioritized condition, unprioritized condition, and neutral condition.

**Discussion**

We examined the distractibility of goal-directed behaviors by systematically manipulating the prioritization of goal-relevant items in working memory and the nature of goal-irrelevant visual distraction. We consistently found that memory responses were systematically biased toward distractors for both prioritized and unprioritized representations. The finding of attraction bias toward distractors in memory responses is consistent with prior studies using a single memorandum (Lorenc et al., 2018; Mallett et al., 2020; Nemes et al., 2012; Rademaker et al., 2015). Here we show that when multiple memory items are maintained, they are biased toward visual distractors, similar to when only one item is maintained. Moreover, we found that prioritization of these memoranda does not reduce or enhance distraction biases, even when relevance and reward cues were combined to produce robust retro-cue benefits. Similarly, the distractor-intrusion model, based on the TCC model of Schurgin et al. (2020) showed that distractor strength was equally strong for prioritized, unprioritized, and neutral items. This finding is unlikely to be driven by the model being insensitive to distraction effects, because the model correctly associated active distraction (in Experiment 5) with stronger distractor intrusions compared with passive distraction (in Experiment 4). Together, these results suggest that prioritization improves the precision of memory reports, but does not modulate the vulnerability of working memories to distractions.

Relevance cues improved memory retention for cued items in the face of distractions, whereas reward cues did not. Following relevance retro-cues, absolute memory errors were reduced for prioritized items compared with unprioritized items. This effect was

**Figure 8**  
Target Memory Strength and Distractor Strength From Experiments 4 and 5



*Note.* (a) Cuing led to strengthened memory signals for prioritized faces compared with both unprioritized faces and the neutral condition in both experiments. (b) Active distractions led to stronger distractor signal strength in Experiment 5 compared with passive distractions in Experiment 4. However, distractor signal strength was indistinguishable for prioritized, unprioritized, and neutral conditions across experiments. See the online article for the color version of this figure.

confirmed by a distractor-intrusion model that identified a significant increase in target memory strength for prioritized versus unprioritized items. This finding replicates many prior studies showing recall benefits of retro-cuing (Griffin & Nobre, 2003; Landman et al., 2003; Makovski & Pertzov, 2015; Matsukura et al., 2007; Maxcey-Richard & Hollingworth, 2013; Pertzov et al., 2013; Rerko & Oberauer, 2013; van Moorselaar et al., 2015; for a review, see Souza & Oberauer, 2016). Our replication of retro-cue benefits following either passive or active visual distraction is consistent with prior research showing that the impact of prioritization in working memory is generally resistant to distraction and may not require sustained attention to achieve (Hollingworth & Maxcey-Richard, 2013; Rerko et al., 2014; Zokaei et al., 2014). Our results also suggest that prioritization based on reward cues could be mechanistically different from prioritization based on relevance cues. When reward cues were used for prioritization, we found no evidence of retro-cue benefits in absolute errors, but there was a small increase in target memory strength for prioritized items as estimated by the distractor-intrusion model. Following reward cues, it is likely that inhibiting the processing of the intervening visual distractors consumed executive control resources that were otherwise being used to maintain prioritization; thus, leading to the elimination of any retro-cue benefits. The mechanisms of prioritization for reward cues and for relevance cues may be independent of each other (Atkinson et al., 2018) and their combination may provide additive retro-cue benefits. When providing a retro-cue that conveys both reward and relevance information, we found a significant improvement in memory

for prioritized items, while memory for unprioritized items was weakened compared with uncued items.

Altogether, our findings from absolute error measurements, distraction bias measurements, and modeling are most consistent with the *null hypothesis* model of prioritization and distraction as shown in Figure 1C. Prioritization sharpens (or preserves the sharpness of) the representation of cued items compared with uncued items, but it does not alter the vulnerability of cued items to visual distractors. Across five experiments, memory bias analyses consistently found that memory responses were attracted toward visual distractors, regardless of the prioritization status of the tested item. Similarly, distractor strength estimates from the distractor-intrusion model did not reveal any systematic effects of prioritization.

Our findings are inconsistent with a group of studies showing that the disruptive effect of a visual distractor was enhanced for prioritized representations compared with unprioritized representations (Allen & Ueno, 2018, Experiments 2, 3, and 4; Hitch et al., 2018; Hu et al., 2014). There are several major design differences between our studies and theirs that might account for this inconsistency. In our study, we focused on memory bias as a key indicator for the impact of distraction on working memory. Although we found similar biases for prioritized and unprioritized items, it is possible that distraction impacted these items in other ways that were independent of memory biases. For example, prior studies supporting the vulnerability hypothesis used a combination of serial presentation of stimuli (including suffix) and a delayed match paradigm, which are likely to induce swap errors (i.e., to mistake one of untested items or even the distractor for the target item).



**Table 2**  
*Distractor d'*

Distractor	Relevance cue						Reward cue						Relevance and reward cue					
	Experiment 1 passive		Experiment 2 passive		Experiment 3 passive		Experiment 4 passive		Experiment 5 active		Experiment 4 passive		Experiment 5 active		Experiment 4 passive		Experiment 5 active	
	P	UP	P	UP	P	UP	P	UP	P	UP	P	UP	P	UP	P	UP	P	UP
MAP	0.115	0.142	0.041	0.050	0.054	0.066	0.050	0.113	0.109	0.270	0.060	0.270	0.213	0.270	0.060	0.270	0.213	0.227
CI	[0.080, 0.163]	[0.079, 0.212]	[0.008, 0.071]	[0.989, 1.162]	[0.002, 0.108]	[0.008, 0.094]	[0.014, 0.123]	[0.015, 0.071]	[0.085, 0.148]	[0.031, 0.171]	[0.007, 0.118]	[0.224, 0.317]	[0.120, 0.296]	[0.143, 0.311]	[0.007, 0.118]	[0.224, 0.317]	[0.120, 0.296]	[0.143, 0.311]

*Note.* MAP = maximum a posteriori; CI = 95% credible interval; P = prioritized; N = neutral; NP = unprioritized. No credible differences were found between prioritization conditions across experiments. Active distractors led to stronger distractor strength compared with passive distractors in Experiment 4.

Second, in almost all existing studies comparing perceptual interference effects between multiple working memory items, visual distractors were designed to be dissimilar to the memoranda (Allen & Ueno, 2018, Experiment 2, 3, and 4; Hitch et al., 2018; Hu et al., 2014). In our study they were designed to be similar to each other because, in the few studies that investigated this, the similarity of the distractor to the memoranda was shown to modulate the interference effect (Rademaker et al., 2015; Teng & Kravitz, 2019; Van der Stigchel et al., 2007). It is possible that prioritization or deprioritization only protects working memory representations from dissimilar distractors that could be easily suppressed or filtered out. In our study, using distractors that were highly similar to the memoranda might make such a protection mechanism less helpful and lead to similar memory biases for both prioritized and unprioritized items. Future research should systematically manipulate the distance between distractor and memoranda to test this possibility. Finally, in four of our five experiments, we provided relevance retro-cues to encourage participants to prioritize the cued item, whereas most prior studies on this topic relied on reward retro-cues. Providing relevance or reward information can lead to independent prioritization operations (Atkinson et al., 2018), and our results suggest that indeed visual distractors interact with these operations to produce differential effects.

Although supportive evidence for the perceptual interference effect has been consistently found in scenarios where only one item needs to be held in working memory, it was unclear whether this result could be generalized to situations when multiple working memory items are maintained. When multiple items need to be held in working memory, memory responses can show interitem biases (Bae & Luck, 2017; Chunharas et al., 2022; Golomb, 2015; Scotti et al., 2021). Such interitem biases can either be attractive or repulsive and the direction of the bias seems to be dependent on the priority status of the tested memory item (Bae & Luck, 2017; Chunharas et al., 2022). It remains unclear whether such interitem biases would modulate the perceptual interference by distractors. It is likely that low memory precision for individual items in a high working memory load setting renders these items more vulnerable to perceptual interferences. However, it is also possible that strong interitem biases obscure relatively weaker biases by perceptual distractors. Our data reveal no significant interitem biases with face stimuli (but see Haberman & Whitney, 2009; Li et al., 2016). Our data demonstrate that the perceptual interference effect does hold for higher memory loads, and the effect does not appear to be modulated by prioritization. To review, the *vulnerability hypothesis* of prioritization and distraction (Figure 1B) suggests that prioritized items are more vulnerable to distraction than are unprioritized items. The rationale is that prioritized items in working memory are maintained through persistent neural activity in the early visual cortex, which could be disrupted when the same neural network needs to process new visual inputs (Bettencourt & Xu, 2016; Hallenbeck et al., 2021). In contrast, unprioritized items are maintained by some other mechanism that does not require persistent neural activity, such as short-term synaptic plasticity (LaRocque et al., 2013; Lewis-Peacock et al., 2012; Rose et al., 2016; Wolff et al., 2017). Such “activity silent” representations could be more robust to concurrent perceptual processing in visual circuits (Lorenz et al., 2021). However, our finding that memory responses for prioritized and unprioritized items are biased similarly toward distractors is inconsistent with the *vulnerability hypothesis*.

They suggest, instead, that silent memories are not any less vulnerable to distraction than are active memories. Conversely, active memories are not any more protected from distraction than are silent memories. Our results are also inconsistent with the *protection hypothesis* of prioritization and distraction (Figure 1A).

Our findings are consistent with the *null hypothesis* of prioritization and distraction, however, whereby the prioritization status of a working memory item does not impact its vulnerability to distraction (Figure 1C). The null results might be driven by the fact that visual perception of distractors lead to disruptions in both persistent neural activity and in synaptic weights in visual cortices, which could disrupt both active and silent representations, respectively. Alternatively, control processes that are typically associated with prioritization, such as top-down inhibition and distraction filtering might protect actively coded representations that are vulnerable to visual distractions (de Vries et al., 2019; Hermann et al., 2021; Suzuki & Gottlieb, 2013; for review, see Lorenc et al., 2021). In other words, the protection effects from control processes might cancel out the vulnerability effects from active coding (for detailed discussion, see Lorenc et al., 2021). One major limitation in our study is that we have no evidence showing that prioritized representations were indeed actively coded and unprioritized representations were silently coded, although prioritization statuses are closely related to these coding schemes. In addition, we have no evidence showing that control mechanisms are similarly or differently engaged for the maintenance of prioritized versus unprioritized memories. Future research should investigate how prioritization and control interact with the perception of distractors.

In the literature, mixed results have been found regarding the effect of prioritization on the sensory interference effect. Such mixed effects can be driven by differences in task designs. Here we summarized two factors that might contribute to the mixed results and should be further investigated in future research. Most prior research showing prioritized information is more vulnerable to perceptual distractors had subjects prioritize multiple items (Allen & Ueno, 2018, Experiments 2, 3, and 4; Hitch et al., 2018; Hu et al., 2014). When multiple items need to be prioritized, there is a potential competition for limited control resources between prioritized information. It might be the competition between prioritized information or the dilution of limited cognitive resources that render prioritized information less protected. Future research is needed to look into this possibility. Additionally, the effect of prioritization on perceptual interference might be dependent on the characteristics of the distractor information. Feature similarity between the distractor and working memory contents might be one critical characteristic. When only one item needs to be maintained in working memory, the similarity of the distractor to the memory item can produce differential interference effects on multiple forms of memory errors, including memory biases, reduction of memory precision, and increases in guess responses (Gresch et al., 2021; Nemes et al., 2011, 2012; Rademaker et al., 2015; Sun et al., 2017). However, it remains unclear whether prioritization would interact differently with distractors that varied in the level of similarity with the targets.

It has been a lingering question how to separate perceptual interference from other memory errors in participants' memory reports. Our development of the signal intrusion model, together with some recent modeling work (Dubé et al., 2014; Fukuda et al., 2022; Rademaker et al., 2015; Sun et al., 2017) have provided tentative solutions for this question, with different assumptions of

mechanisms underlying the perceptual interference effect. Specifically, research that used revised mixture models (Rademaker et al., 2015; Sun et al., 2017) to separate memory biases, swap errors, and guesses implicitly assumed that perceptual distractions have multiple separable effects on memory reports. In contrast, research that used signal integration or averaging models (Dubé et al., 2014; Fukuda et al., 2022) such as our signal intrusion model assumed that the perceptual interference effect is driven by a unified mechanism of intermingling the target signal with the distractor signal during memory reports. Based on this assumption, swap errors and biases are simply caused by differences in the distractor signal strength. When the distractor signal is strong, memory responses are attracted toward distractions to a large degree, and those large biased responses are more likely to be classified as swap errors in the mixture model. When the distractor signal is weak, memory responses are attracted toward distractions to a small degree, and those minorly biased responses are more likely to be captured as memory biases instead of swap errors by the mixture model. To differentiate the two groups of models, it is necessary to elucidate mechanisms underlying seemingly different forms of errors caused by perceptual distractions, including biases, swap errors to distractions and guesses.

Despite our finding that memory responses were systematically, albeit subtly, biased toward visual distractors, we acknowledge that neither the passive perception of nor active engagement with distractors led to catastrophic memory impairment. This is consistent with repeated observations that working memory is quite robust to distractions. Recent neuroimaging work has proposed a variety of potential mechanisms supporting working memory maintenance against distraction (Bettencourt & Xu, 2016; Hallenbeck et al., 2021; Lorenc et al., 2018; Rademaker et al., 2019). Working memory representations of visual information can be decoded in both early visual cortex as well as in parietal regions, and such a parallel coding scheme could protect maintained representations from perception-related disruptions in the visual cortex (Bettencourt & Xu, 2016; Lorenc et al., 2018). Additionally, working memory representations have been shown to coexist with perceptual representation of distractors in early visual cortex (Hallenbeck et al., 2021; Rademaker et al., 2019), and if temporarily disrupted, they can quickly and effectively recover (Hakim et al., 2020, 2021; Hallenbeck et al., 2021; Mallett & Lewis-Peacock, 2019).

## Conclusions

In summary, the current study demonstrates that prioritization of a working memory item neither helps nor hurts its vulnerability to distraction. Prioritization does strengthen memory signals for prioritized items (i.e., it produces a canonical "retro-cue benefit"), but it does not provide any greater (or lesser) protection for these items against distraction.

## References

- Allen, R. J., & Ueno, T. (2018). Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Attention, Perception & Psychophysics*, *80*(7), 1731–1743. <https://doi.org/10.3758/s13414-018-1543-6>
- Atkinson, A. L., Berry, E. D. J., Waterman, A. H., Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018). Are there multiple ways to direct attention

- in working memory? *Annals of the New York Academy of Sciences*, 1424(1), 115–126. <https://doi.org/10.1111/nyas.13634>
- Bae, G. Y., & Luck, S. J. (2017). Interactions between visual working memory representations. *Attention, Perception & Psychophysics*, 79(8), 2376–2395. <https://doi.org/10.3758/s13414-017-1404-8>
- Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience*, 19(1), 150–157. <https://doi.org/10.1038/nn.4174>
- Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *The Journal of Neuroscience*, 31(25), 9315–9322. <https://doi.org/10.1523/JNEUROSCI.1097-11.2011>
- Chunharas, C., Rademaker, R. L., Brady, T. F., & Serences, J. T. (2022). An adaptive perspective on visual working memory distortions. *Journal of Experimental Psychology: General*, 151(10), 2300–2323. <https://doi.org/10.1037/xge0001191>
- D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 362, 761–772. <https://doi.org/10.1098/rstb.2007.2086>
- D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, 66(1), 115–142. <https://doi.org/10.1146/annurev-psych-010814-015031>
- de Vries, I. E. J., Savran, E., van Driel, J., & Olivers, C. N. L. (2019). Oscillatory mechanisms of preparing for visual distraction. *Journal of Cognitive Neuroscience*, 31(12), 1873–1894. [https://doi.org/10.1162/jocn\\_a\\_01460](https://doi.org/10.1162/jocn_a_01460)
- Dubé, C., Zhou, F., Kahana, M. J., & Sekuler, R. (2014). Similarity-based distortion of visual short-term memory is due to perceptual averaging. *Vision Research*, 96, 8–16. <https://doi.org/10.1016/j.visres.2013.12.016>
- Fritsche, M., & de Lange, F. P. (2019). The role of feature-based attention in visual serial dependence. *Journal of Vision*, 19(13), Article 21. <https://doi.org/10.1167/19.13.21>
- Fukuda, K., Pereira, A. E., Saito, J. M., Tang, T. Y., Tsubomi, H., & Bae, G.-Y. (2022). Working Memory Content Is Distorted by Its Use in Perceptual Comparisons. *Psychological Science*, 33(5), 816–829. <https://doi.org/10.1177/09567976211055375>
- Golomb, J. D. (2015). Divided spatial attention and feature-mixing errors. *Attention, Perception & Psychophysics*, 77(8), 2562–2569. <https://doi.org/10.3758/s13414-015-0951-0>
- Gresch, D., Boettcher, S. E. P., van Ede, F., & Nobre, A. C. (2021). Shielding working-memory representations from temporally predictable external interference. *Cognition*, 217, Article 104915. <https://doi.org/10.1016/j.cognition.2021.104915>
- 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/10.1162/089892903322598139>
- Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3), 718–734. <https://doi.org/10.1037/a0013899>
- 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>
- Hallenbeck, G. E., Sprague, T. C., Rahmati, M., Sreenivasan, K. K., & Curtis, C. E. (2021). Working memory representations in visual cortex mediate distraction effects. *Nature Communications*, 12(1), Article 4714. <https://doi.org/10.1038/s41467-021-24973-1>
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–635. <https://doi.org/10.1038/nature07832>
- Hermann, P., Weiss, B., Knakker, B., Madurka, P., Manga, A., Nárai, Á., & Vidnyánszky, Z. (2021). Neural basis of distractor resistance during visual working memory maintenance. *NeuroImage*, 245, Article 118650. <https://doi.org/10.1016/j.neuroimage.2021.118650>
- Hitch, G. J., Hu, Y., Allen, R. J., & Baddeley, A. D. (2018). Competition for the focus of attention in visual working memory: Perceptual recency versus executive control. *Annals of the New York Academy of Sciences*, 1424(1), 64–75. <https://doi.org/10.1111/nyas.13631>
- Hollingworth, A., & Maxcey-Richard, A. M. (2013). Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 39(4), 1047–1058. <https://doi.org/10.1037/a0030238>
- Hu, Y., Hitch, G. J., Baddeley, A. D., Zhang, M., & Allen, R. J. (2014). Executive and perceptual attention play different roles in visual working memory: Evidence from suffix and strategy effects. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1665–1678. <https://doi.org/10.1037/a0037163>
- Kang, M.-S., Hong, S. W., Blake, R., & Woodman, G. F. (2011). Visual working memory contaminates perception. *Psychonomic Bulletin & Review*, 18(5), 860–869. <https://doi.org/10.3758/s13423-011-0126-5>
- Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, 20(2), 228–242. <https://doi.org/10.3758/s13423-012-0359-y>
- Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, 43(2), 149–164. [https://doi.org/10.1016/S0042-6989\(02\)00402-9](https://doi.org/10.1016/S0042-6989(02)00402-9)
- LaRocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Frontiers in Human Neuroscience*, 8, Article 5. <https://doi.org/10.3389/fnhum.2014.00005>
- LaRocque, J. J., Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2013). Decoding attended information in short-term memory: An EEG study. *Journal of Cognitive Neuroscience*, 25(1), 127–142. [https://doi.org/10.1162/jocn\\_a\\_00305](https://doi.org/10.1162/jocn_a_00305)
- Lepsien, J., Griffin, I. C., Devlin, J. T., & Nobre, A. C. (2005). Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. *NeuroImage*, 26(3), 733–743. <https://doi.org/10.1016/j.neuroimage.2005.02.026>
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, 24(1), 61–79. [https://doi.org/10.1162/jocn\\_a\\_00140](https://doi.org/10.1162/jocn_a_00140)
- Li, H., Ji, L., Tong, K., Ren, N., Chen, W., Liu, C. H., & Fu, X. (2016). Processing of individual items during ensemble coding of facial expressions. *Frontiers in Psychology*, 7, Article 1332. <https://doi.org/10.3389/fpsyg.2016.01332>
- 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>
- Lorenc, E. S., Pratte, M. S., Angeloni, C. F., & Tong, F. (2014). Expertise for upright faces improves the precision but not the capacity of visual working memory. *Attention, Perception & Psychophysics*, 76(7), 1975–1984. <https://doi.org/10.3758/s13414-014-0653-z>
- Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R. E., & D'Esposito, M. (2018). Flexible coding of visual working memory representations during distraction. *The Journal of Neuroscience*, 38(23), 5267–5276. <https://doi.org/10.1523/JNEUROSCI.3061-17.2018>
- 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>
- Makovski, T., & Pertsov, Y. (2015). Attention and memory protection: Interactions between retrospective attention cueing and interference. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 68(9), 1735–1743. <https://doi.org/10.1080/17470218.2015.1049623>
- Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes.



- Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2), 369–380. <https://doi.org/10.1037/0278-7393.34.2.369>
- Mallett, R., & Lewis-Peacock, J. A. (2018). Behavioral decoding of working memory items inside and outside the focus of attention. *Annals of the New York Academy of Sciences*, 1424(1), 256–267. <https://doi.org/10.1111/nyas.13647>
- Mallett, R., & Lewis-Peacock, J. A. (2019). Working memory prioritization impacts neural recovery from distraction. *Cortex*, 121, 225–238. <https://doi.org/10.1016/j.cortex.2019.08.019>
- 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>
- Marini, F., Scott, J., Aron, A. R., & Ester, E. F. (2017). Task-irrelevant distractors in the delay period interfere selectively with visual short-term memory for spatial locations. *Attention, Perception & Psychophysics*, 79(5), 1384–1392. <https://doi.org/10.3758/s13414-017-1320-y>
- Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: Protection or prioritization? *Perception & Psychophysics*, 69(8), 1422–1434. <https://doi.org/10.3758/BF03192957>
- Maxcey-Richard, A. M., & Hollingworth, A. (2013). The strategic retention of task-relevant objects in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 760–772. <https://doi.org/10.1037/a0029496>
- Muhle-Karbe, P. S., Myers, N. E., & Stokes, M. G. (2021). A hierarchy of functional states in working memory. *The Journal of Neuroscience*, 41(20), 4461–4475. <https://doi.org/10.1523/JNEUROSCI.3104-20.2021>
- Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing information during working memory: Beyond sustained internal attention. *Trends in Cognitive Sciences*, 21(6), 449–461. <https://doi.org/10.1016/j.tics.2017.03.010>
- 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), 26–26. <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>
- Oberauer, K. (2021). Measurement models for visual working memory—A factorial model comparison. *Psychological Review*. Advance online publication. <https://doi.org/10.1037/rev0000328>
- Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32(5), 1243–1265. <https://doi.org/10.1037/0096-1523.32.5.1243>
- Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, 15(7), 327–334. <https://doi.org/10.1016/j.tics.2011.05.004>
- Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, 6(2), 97–107. <https://doi.org/10.1038/nrn1603>
- Pertsov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology: Human Perception and Performance*, 39(5), 1224–1231. <https://doi.org/10.1037/a0030947>
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139(1), 23–38. <https://doi.org/10.1016/j.neuroscience.2005.06.005>
- Postle, B. R. (2015). The cognitive neuroscience of visual short-term memory. *Current Opinion in Behavioral Sciences*, 1, 40–46. <https://doi.org/10.1016/j.cobeha.2014.08.004>
- 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>
- Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. *Nature Neuroscience*, 22(8), 1336–1344. <https://doi.org/10.1038/s41593-019-0428-x>
- Reerko, L., & Oberauer, K. (2013). Focused, unfocused, and defocused information in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(4), 1075–1096. <https://doi.org/10.1037/a0031172>
- Reerko, L., Souza, A. S., & Oberauer, K. (2014). Retro-cue benefits in working memory without sustained focal attention. *Memory & Cognition*, 42(5), 712–728. <https://doi.org/10.3758/s13421-013-0392-8>
- Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseseries, 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.aah7011>
- Schneider, D., Barth, A., & Wascher, E. (2017). On the contribution of motor planning to the retroactive cuing benefit in working memory: Evidence by mu and beta oscillatory activity in the EEG. *NeuroImage*, 162, 73–85. <https://doi.org/10.1016/j.neuroimage.2017.08.057>
- Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, 4(11), 1156–1172. <https://doi.org/10.1038/s41562-020-00938-0>
- Scotti, P. S., Hong, Y., Leber, A. B., & Golomb, J. D. (2021). Visual working memory items drift apart due to active, not passive, maintenance. *Journal of Experimental Psychology: General*, 150(12), 2506–2524. <https://doi.org/10.1037/xge0000890>
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207–214. <https://doi.org/10.1111/j.1467-9280.2009.02276.x>
- Shepherdson, P. (2021). Perceptual stimuli with novel bindings interfere with visual working memory. *Attention, Perception & Psychophysics*, 83(8), 3086–3103. <https://doi.org/10.3758/s13414-021-02359-1>
- Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, 12(9), 342–348. <https://doi.org/10.1016/j.tics.2008.05.007>
- Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception & Psychophysics*, 78(7), 1839–1860. <https://doi.org/10.3758/s13414-016-1108-5>
- 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>
- Stokes, M. G., Muhle-Karbe, P. S., & Myers, N. E. (2020). Theoretical distinction between functional states in working memory and their corresponding neural states. *Visual Cognition*, 28(5–8), 420–432. <https://doi.org/10.1080/13506285.2020.1825141>
- Suchow, J. W., Brady, T. F., Fougine, 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>
- Sun, S. Z., Fidalgo, C., Barense, M. D., Lee, A. C. H., Cant, J. S., & Ferber, S. (2017). Erasing and blurring memories: The differential impact of interference on separate aspects of forgetting. *Journal of Experimental Psychology: General*, 146(11), 1606–1630. <https://doi.org/10.1037/xge0000359>
- Suzuki, M., & Gottlieb, J. (2013). Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nature Neuroscience*, 16(1), 98–104. <https://doi.org/10.1038/nn.3282>
- Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour*, 3(8), 827–836. <https://doi.org/10.1038/s41562-019-0640-4>



- Van der Stigchel, S., Merten, H., Meeter, M., & Theeuwes, J. (2007). The effects of a task-irrelevant visual event on spatial working memory. *Psychonomic Bulletin & Review*, *14*(6), 1066–1071. <https://doi.org/10.3758/BF03193092>
- van Loon, A. M., Olmos-Solis, K., & Olivers, C. N. L. (2017). Subtle eye movement metrics reveal task-relevant representations prior to visual search. *Journal of Vision*, *17*(6), Article 13. <https://doi.org/10.1167/17.6.13>
- van Moorselaar, D., Gonseli, E., Theeuwes, J., & Olivers, C. N. L. (2015). The time course of protecting a visual memory representation from perceptual interference. *Frontiers in Human Neuroscience*, *8*, Article 1053. <https://doi.org/10.3389/fnhum.2014.01053>
- van Moorselaar, D., Theeuwes, J., & Olivers, C. N. L. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1450–1464. <https://doi.org/10.1037/a0036229>
- Wildegger, T., Myers, N. E., Humphreys, G., & Nobre, A. C. (2015). Supraliminal but not subliminal distracters bias working memory recall. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 826–839. <https://doi.org/10.1037/xhp0000052>
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 1120–1135. <https://doi.org/10.1167/4.12.11>
- Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871. <https://doi.org/10.1038/nn.4546>
- 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>
- Zokaei, N., Ning, S., Manohar, S., Feredoes, E., & Husain, M. (2014). Flexibility of representational states in working memory. *Frontiers in Human Neuroscience*, *8*, Article 853. <https://doi.org/10.3389/fnhum.2014.00853>

Received April 15, 2022

Revision received September 7, 2022

Accepted September 8, 2022 ■