Journal of Experimental Psychology: General

Working Memory Prioritization Changes Bidirectional Interactions With Visual Inputs

Joseph M. Saito, Frida A. B. Printzlau, Yvanna Yeo, and Keisuke Fukuda

CITATION

Saito, J. M., Printzlau, F. A. B., Yeo, Y., & Fukuda, K. (2025). Working memory prioritization changes bidirectional interactions with visual inputs. *Journal of Experimental Psychology: General*. Advance online publication. https://dx.doi.org/10.1037/xge0001813

Journal of Experimental Psychology: General

ISSN: 0096-3445
https://doi.org/10.1037/xge0001813

Working Memory Prioritization Changes Bidirectional Interactions With Visual Inputs

Joseph M. Saito[1], Frida A. B. Printzlau[2, 3], Yvanna Yeo[3], and Keisuke Fukuda[2, 3]
[1] Department of Psychology, University of California, San Diego
[2] Department of Psychology, University of Toronto
[3] Department of Psychology, University of Toronto Mississauga

Items stored in visual working memory often differ in priority. Typically, observers will shift their internal attention toward items that are relevant for impending behavior and away from those that may become relevant later. These distinct states of priority are theorized to influence bidirectional interactions between memoranda and new visual inputs by modulating their susceptibility to retroactive and proactive report biases, respectively. However, prior research has produced limited and conflicting evidence on this topic due to reliance on inconsistent retrodictive cues (retro-cues) that incentivize memory prioritization. To address this, we used a double-serial retro-cue paradigm that incentivized the complete prioritization of one of two unfamiliar shape memoranda for use in a comparison with a perceptual probe before a second, independent cue instructed observers to report one of their two memories (Experiments 1–2) or the perceptual probe itself (Experiment 2). We found that observers reported robust retroactive and proactive biases, but that only retroactive biases were modulated by prioritization. Reports of prioritized memories were more precise and contained smaller attractive biases toward the probe than unprioritized memories, whereas probe reports were biased comparably toward each. These findings reveal an asymmetrical effect of prioritization on the reciprocal interactions between new and existing visual representations.

Public Significance Statement

Memories are not simply stored in mind; they are used in everyday behavior. During such use, humans will often shift their attention between different memoranda in accordance with their current task goal, prioritizing some memories over others. At the same time, memories are known to interact with new visual inputs coming in from the surrounding world, producing biases in what we report having seen then and now. In the present study, we show that these report biases vary unevenly based on which memories are prioritized at the time. While reports of prioritized memories tend to show less bias toward visual inputs than those of unprioritized memories, reports of visual inputs tend to be biased by memories equally regardless of priority. These findings suggest that attending to a memory during behavior may help protect it against interference but may not strongly change its proactive influence on new visual representations.

Keywords: memory prioritization, perceptual comparisons, memory bias, serial dependence, memory reactivation

Supplemental materials: https://doi.org/10.1037/xge0001813.supp

---

Ashleigh Maxcey served as action editor.

The data are posted publicly online at https://osf.io/9qzw7/. The data were previously disseminated as part of conference presentations at the 2022 Working Memory Symposium (https://www.wmsymposium.org/archive); the 2022 Object Perception, Attention, and Memory Conference; and the 2022 Annual Meeting of the Psychonomics Society (https://www.psychonomic.org/page/2022program). The authors have no conflicts of interest to declare that are relevant to the content of this article.

This research was supported by the National Sciences and Engineering Research Council (Grant RGPIN-2017-06866) awarded to Keisuke Fukuda and the Connaught New Researcher Award.

Joseph M. Saito played a lead role in visualization, writing–original draft, and writing–review and editing and an equal role in conceptualization, formal analysis, and methodology. Frida A. B. Printzlau played an equal role in conceptualization, formal analysis, methodology, and writing–review and editing. Yvanna Yeo played a lead role in data curation and investigation. Keisuke Fukuda played a lead role in funding acquisition, resources, software, and supervision and an equal role in conceptualization, methodology, and writing–review and editing.

Correspondence concerning this article should be addressed to Joseph M. Saito, Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, San Diego, CA 92093, United States. Email: josaito@ucsd.edu

Humans are capable of maintaining information about a small number of visual stimuli within their visual working memory (VWM) system for use in ongoing behavior (Adam et al., 2017; Awh et al., 2007; Bays & Husain, 2008; van den Berg et al., 2012; Wilken & Ma, 2004; W. Zhang & Luck, 2008). During such use, the priority of stored items is subject to change based on their relevance to ongoing behavior. For example, when searching for a pair of friends in a crowded shopping mall, you may prioritize looking for one friend over another when you hear the sound of their voice close by. In this situation, the sound of your friend's voice serves as a retrodictive cue (or retro-cue) about the relevance of their memory representation for completing your task relative to the friend that you did not hear. In order to implement this state of prioritization, prominent theory and empirical evidence suggest that observers shift their internal focus of attention (FOA) to the cued representation, which imparts an enhanced state of accessibility and precision (Griffin & Nobre, 2003; Nobre & Stokes, 2019; Oberauer & Hein, 2012; Olivers et al., 2011; Souza & Oberauer, 2016). Upon doing so, other unprioritized memoranda are subsequently stored outside the FOA, either in a separate state of VWM (e.g., region of direct access) or in long-term memory (e.g., activated long-term memory; Oberauer, 2002, 2007, 2009).

In the brain, prioritized representations tend to remain decodable from their sustained activity across the cortex, including in sensory regions that are responsible for processing new visual inputs (LaRocque et al., 2013, 2017; Lewis-Peacock et al., 2012; Sprague et al., 2016). By contrast, neural activity tied to unprioritized representations is significantly reduced or eliminated, suggesting reliance on a different type of storage mechanism instead (e.g., transient synaptic weighting; Mongillo et al., 2008; Rose et al., 2016; Wolff et al., 2017; but see also Barbosa et al., 2020; Christophel et al., 2018; Iamshchinina et al., 2021; Wan et al., 2020; Yu et al., 2020). From a theoretical perspective, these distinct states of representation may help to facilitate perceptual comparisons between prioritized memories and new visual inputs, as this represents a ubiquitous demand across visual cognition, including when searching for friends in a crowd (Adam et al., 2022; Olivers & Roelfsema, 2020; Rademaker et al., 2019; Xu, 2020).

Different neural codes may also imply that memories interact differently with visual inputs based on their behavioral priority. New and existing visual representations are known to interact in a bidirectional manner, often producing retroactive and proactive biases in their respective behavioral reports that are thought to indicate subtle interference between their representations. In the retroactive direction, encountering a new percept during VWM maintenance tends to produce a reliable attraction bias in observers' subsequent memory reports (e.g., Huang & Sekuler, 2010; Mallett et al., 2020; Teng & Kravitz, 2019; Van der Stigchel et al., 2007; Wildegger et al., 2015), especially when the percept is relevant to ongoing behavior (e.g., Rademaker et al., 2015; Saito et al., 2022, 2024). In the proactive direction, observers' reports of new visual inputs show similar biases with respect to the existing contents of VWM (e.g., Kang et al., 2011; Teng & Kravitz, 2019). Interestingly, these proactive biases can arise even when memoranda are no longer relevant to ongoing behavior. For example, memory reports on a given trial tend to show attractive biases toward memoranda from the immediately preceding trial (Barbosa & Compte, 2020; Fischer & Whitney, 2014; Kiyonaga et al., 2017; Liberman et al., 2014). Findings of this kind have led to considerable debate regarding the exact nature and time course of these bidirectional interactions. At present, it remains unknown whether these biases occur at the level of the sensory representations, the decisions that observers make about the representations, or some combination of the two (Hajonides et al., 2023; Kiyonaga et al., 2017; Lorenc et al., 2018; Manassi et al., 2023; Moon & Kwon, 2022; Rademaker et al., 2019; Saito et al., 2023; Sheehan & Serences, 2022).

Despite this clear evidence of both retroactive and proactive biases in behavior, it is unclear how memory priority modulates these interactions. In the retroactive direction, prioritization effects are highly inconsistent and have led to competing theories of how attention modulates perceptual interference in memory. On the one hand, some studies suggest that prioritization confers modest protection of memory against retroactive interference by new visual inputs (Barth & Schneider, 2018; Makovski & Pertzov, 2015; Makovski & Jiang, 2007; Schneider et al., 2017; Souza et al., 2016; van Moorselaar et al., 2015). For example, when cueing which of multiple memoranda is to be reported at the end of a trial with absolute certainty (i.e., 100% cue validity), the cued item becomes more resistant to interference by task-irrelevant visual distraction. This set of findings has been viewed as supporting a *protective attention account* of prioritization (Oberauer, 2002; Oberauer & Lin, 2017). On the other hand, other studies have suggested an opposite effect of prioritization on retroactive interference. In these studies, cueing which of multiple memoranda is most likely to be reported at the end of the trial (i.e., <100% cue validity; Z. Zhang & Lewis-Peacock, 2023a, 2023b) or whose report is worth the most reward (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014, 2016) eliminates or reverses the protective benefits of the FOA. In response, some have proposed a *protective dormancy account* of prioritization instead, where unprioritized memories are better at avoiding interactions with new visual inputs by virtue of their unique storage state in the brain that is less sensory-specific than prioritized memories (Lorenc et al., 2021; Z. Zhang & Lewis-Peacock, 2023a, 2023b). While some of the inconsistency between these accounts can be attributed to a historical reliance on insensitive measures of interference, such as recognition or recall accuracy (Allen & Ueno, 2018; Hu et al., 2014, 2016; Makovski et al., 2008; Makovski & Jiang, 2007), conflicting patterns have also been observed among studies that employ more sensitive continuous measures as well (Barth & Schneider, 2018; Schneider et al., 2017; van Moorselaar et al., 2015; Z. Zhang & Lewis-Peacock, 2023a, 2023b).

In the proactive direction, studies of memory prioritization are not as much inconsistent as they are lacking. For the most part, researchers have focused exclusively on measuring attentional guidance by prioritized and unprioritized memoranda in the context of visual search. These studies suggest that while prioritized memories are capable of guiding attention toward matching visual inputs, guidance by unprioritized memories is typically transient and less reliable (Greene et al., 2015; Mallett & Lewis-Peacock, 2018; Olivers et al., 2011; Ort et al., 2018, 2019; Ort & Olivers, 2020; van Loon et al., 2017; but see also Bahle et al., 2020; Beck et al., 2012). On the one hand, these adjacent findings suggest that the priority of a memory can indeed influence how visual inputs are processed. On the other hand, because these studies do not prompt observers to report those visual inputs, their findings can only speak to the effect of prioritization on attentional allocation and not to how those attended contents are represented. In order to provide a more comprehensive theoretical account of the bidirectional interactions between new and existing visual representations, it is necessary to characterize how the accuracy of new representations changes based on the priority of the memory that interacts with them.

Here, we considered the possibility that interactions between new and existing visual representations depend upon the nature of the retro-cues that incentivize prioritization. For the protective attention account, one could argue that the protective benefits of the FOA are contingent upon dropping uncued items from VWM in response to a 100% valid cue (Kuo et al., 2012; Lorenc et al., 2021; Schneider et al., 2017; Souza et al., 2014; M. Williams & Woodman, 2012), making it unclear whether prioritization can mitigate retroactive interference when unprioritized memoranda must still be maintained. While the use of a graded certainty cue by proponents of the protective dormancy account helps to address this set size discrepancy, graded cueing may change how observers implement prioritization. For example, when cueing the relevance (or reward) of a memory with 80% certainty (or value), observers may attempt to allocate their attention proportionately by shifting their FOA between cued and uncued items (Gould et al., 2011; Gunseli et al., 2015; Wang & van Ede, 2025; M. Williams & Woodman, 2012). While an attentional bias of this kind may be sufficient to produce differences in the precision of cued and uncued memory reports (e.g., Gunseli et al., 2015; Makovski & Pertzov, 2015; Z. Zhang & Lewis-Peacock, 2023a, 2023b), protective benefits may require fully committing attentional resources to only one item. These systematic differences in cueing between studies represent a significant barrier to resolving the theoretical debate described above, as no study to our knowledge has been able to successfully control for both cue validity and memory set size within a single paradigm.

To try and address these discrepancies, we conducted a pair of experiments using a double-serial retro-cue paradigm that incentivized the complete prioritization of one memory item without the ability to drop the other from VWM (e.g., Lewis-Peacock et al., 2012; Rose et al., 2016). On each trial, observers were asked to encode a pair of unfamiliar shapes into VWM in anticipation of a subsequent memory report. During maintenance, observers were retro-cued to compare one of these two memoranda to a novel probe stimulus in order to satisfy a similarity judgment. By implementing demands for an explicit comparison, the first priority cue is necessarily rendered 100% valid, and observers' ability to easily shift their attention between memoranda is constrained. From here, we then presented a second retro-cue that instructed observers to report one of the two memoranda from a continuous wheel. Because this second retro-cue was independent of the first, observers were not able to simply drop the first uncued item from VWM. To address the existing knowledge gap surrounding proactive interactions, we conducted a follow-up experiment using a modified version of the task where observers were sometimes cued to report the features of the probe stimulus itself immediately after the perceptual comparison. This allowed us to characterize proactive biases in the incoming probe while also testing for replication of the retroactive biases observed in the first experiment.

In the retroactive direction, we found that prioritizing a memory item was associated with more precise memory reports and smaller attractive report biases compared to unprioritized maintenance. When we attempted to parse the nature of these report biases using computational mixture modeling (Schurgin et al., 2020; J. R. Williams et al., 2022), we found that prioritization was associated with smaller representational shifts of the memory toward the probe and fewer misreports of the probe itself (i.e., swap errors), suggesting a reduction in both modest and catastrophic forms of interference during usage. By contrast, in the proactive direction, we found that observers' reports of the probe were attracted toward prioritized and unprioritized memoranda with comparable magnitude. These findings suggest that the FOA is capable of providing modest protection of memoranda against retroactive interference by new visual inputs during usage, consistent with the protective attention account, but that this protective benefit does not extend proactively to the inputs themselves. In our discussion, we speculate that this asymmetric influence of prioritization across retroactive and proactive interactions may reflect an important distinction between the functional memory states that are conferred by attentional selection and the corresponding representational states of those memories in the brain (Stokes et al., 2020; Sutterer et al., 2019).

Experiment 1

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The data and analysis scripts are available on the Open Science Framework (https://osf.io/9qzw7/). All analyses were conducted using MATLAB, Version R2020a (The MathWorks Inc., 2020), and the Psychophysics Toolbox extension, Version 3.0.16 (Kleiner et al., 2007). This study's design and analyses were not preregistered.

Participants

All participants in the experiment were undergraduate students at the University of Toronto Mississauga that reported normal or corrected-to-normal visual acuity and color vision. Each participant provided informed consent in accordance with the procedures approved by the research ethics board at the University of Toronto.

We conducted a series of planned t tests to compare the report biases induced by performing perceptual comparisons between new and existing visual representations. Previous studies investigating these effects report large effect sizes (i.e., Cohen's $d > 0.8$, Fukuda et al., 2022; Saito et al., 2022, 2024). However, because we were interested here in the additional modulations that are introduced by memory prioritization, we chose to rely on a more conservative estimate of effect size (i.e., Cohen's $d = 0.5$). A power calculation performed with an α level of .05 and statistical power of 0.8 indicated that we would need at least 34 participants to obtain such an effect (Faul et al., 2007).

Participants were recruited on a weekly basis until the sample size exceeded the minimum number identified by our power analysis. This resulted in the recruitment of 51 participants. Each participant reported their demographic information in a response window that appeared before the start of the experimental script. For gender identity, participants selected between male, female, and prefer not to answer. Participants submitted their age in years. Information about racial and ethnic identity was not collected. One participant was excluded for not following instructions during the first block of the task. Two participants were excluded for poor overall performance on the task, as indicated by a mean absolute report error in the baseline condition that was >2.5 *SD* above the mean of the group. Data collected from the remaining 48 participants were submitted to analysis. Among this remaining sample, 35 participants reported their gender as female and 11 as male and one chose not to report. The mean age of the sample was 19.13 years.

Apparatus and Stimuli

Participants completed the experiment remotely using a personal desktop or laptop computer. To ensure that experimental conditions were satisfactory, completion of all task procedures was monitored by researchers in real-time using Zoom video conference software (Zoom Video Communications, Inc, 2016). All stimuli were generated and presented in Inquisit 6 (Millisecond Software LLC, 2016), which was run locally on each participant's computer. Given that participants' viewing distance could not be tightly controlled using an online procedure, we fit all presented stimuli to a square whose size was fixed proportionately to the vertical axis of the monitor (see below).

We used a continuous shape space whose circular visual similarity has been empirically validated (Li et al., 2020). We opted to use this feature space because it does not contain any prototypical shapes (e.g., triangles, squares) that could allow observers to easily rely on verbalization or long-term categorical knowledge when performing the task. The target and probe items for each trial were sampled from 360 shapes within this stimulus set. Target shapes were positioned at 30% and 70% of the horizontal axis of the monitor (i.e., 20% from screen center), respectively, and the probe shape was positioned at 50% of the horizontal axis (i.e., at screen center). The shape wheel consisted of 12 equidistant shape exemplars presented in a circular arrangement with a radius set to 40% of the vertical axis of the monitor. The size of each exemplar shape on the wheel was fixed at 8% of the vertical axis. The retro-cues were sized at 5% of the vertical axis and positioned at 40% of the vertical axis (i.e., 10% above screen center).

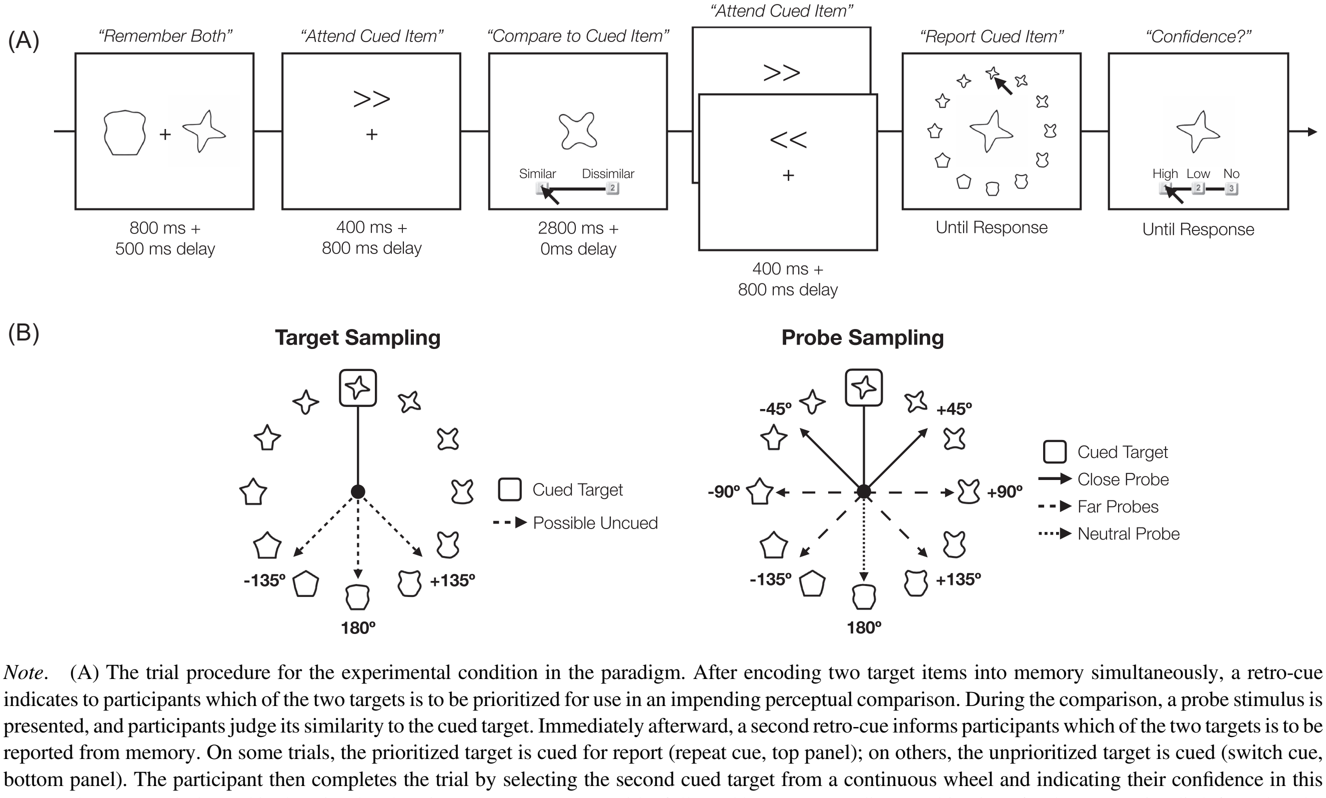
Procedure

Percentages reported in parentheses below indicate the proportion of trials comprised by a given condition.

Participants performed eight blocks of 36 trials (Figure 1A). Each trial began with two target shapes presented on the left and right sides of a centrally presented fixation cross for 800 ms. Participants were instructed to remember both of the targets as precisely as possible for the remainder of the trial. One of the targets was randomly sampled from the feature space, while the other target was sampled ±135° (67%) or 180° (33%) away from it (Figure 1B). The direction of offset was counterbalanced between clockwise and counterclockwise. The distant separation between targets was implemented to minimize their confusability during encoding and retention (e.g., Bae & Luck, 2017; Chunharas et al., 2022; Scotti et al., 2021). Separate analyses yielded moderate evidence that

Figure 1

Double-Serial Retro-Cue Paradigm With Target and Probe Sampling



Note. (A) The trial procedure for the experimental condition in the paradigm. After encoding two target items into memory simultaneously, a retro-cue indicates to participants which of the two targets is to be prioritized for use in an impending perceptual comparison. During the comparison, a probe stimulus is presented, and participants judge its similarity to the cued target. Immediately afterward, a second retro-cue informs participants which of the two targets is to be reported from memory. On some trials, the prioritized target is cued for report (repeat cue, top panel); on others, the unprioritized target is cued (switch cue, bottom panel). The participant then completes the trial by selecting the second cued target from a continuous wheel and indicating their confidence in this report. The "+" symbols used in the stimulus durations are synonymous with "followed by." (B) Sampling combinations for the target stimuli (left) and probe stimulus (right) for each trial. The stimulus highlighted in the box and connected by a solid line is used as the prioritized target in (A).

interitem biases between targets in the baseline condition were indeed negligible, repeat: $t(47) = -1.15$, $p = .257$, Cohen's $d = -0.17$, $BF_{01} = 3.44$; switch: $t(47) = -0.76$, $p = .453$, Cohen's $d = -0.11$, $BF_{01} = 4.87$. This distant sampling procedure may have also incentivized the use of verbal labels for the targets, which could have produced categorical report biases (G.-Y. Bae et al., 2015; Garside et al., 2025) and changes in report precision and forgetting (Chung et al., 2024; Overkott & Souza, 2022; Souza et al., 2021). However, because our sampling of the target and probe stimuli was independent of any underlying categorical structure in the stimulus space, any systematic contributions of verbal labeling to performance were assumed to be significantly reduced or eliminated when measuring errors relative to the probe (see the Behavioral Analysis section).

Five hundred milliseconds after the offset of the targets, a retro-cue pointing to the left or right side of the screen was presented above the centrally presented fixation cross for 400 ms. Participants were instructed to shift their internal FOA toward the target representation that was presented on the side of the screen indicated by the retro-cue. The direction of the retro-cue was randomly assigned. Eight hundred milliseconds after the offset of the first retro-cue, a probe shape was presented at the center of the screen for 2,800 ms (67%). The probe was sampled ±45° away from one of the two targets in a counterbalanced manner (Figure 1B). Below the probe, two response buttons labeled "similar" and "dissimilar," respectively, were presented horizontally to one another and connected by a line. Participants were instructed to compare the probe to the cued target and report whether the probe was similar or dissimilar to the target by clicking on the response button of choice with the cursor. The similarity judgment was subjective, and participants were not provided any specific instruction about how they should determine the similarity between the target and probe. Once participants reported their similarity judgment, the probe and response buttons were removed from the screen, and the remainder of the comparison period was filled by a blank delay. On a subset of trials in each block (33%), the probe was not presented during the judgment period, and participants did not need to make any response. We refer to this subset of trials as baseline trials.

At the completion of the judgment period, a second retro-cue pointing to the left or right side of the screen was presented above the centrally presented fixation cross for 400 ms. Participants were instructed to shift their internal FOA toward the target representation that was presented on the side of the screen indicated by the retro-cue. The second retro-cue was counterbalanced to cue the same target as the first retro-cue (i.e., repeat trial, 50%) or the opposite target (i.e., switch trial, 50%).

Eight hundred milliseconds after the offset of the second retro-cue, a continuous shape wheel was presented. The wheel was randomly rotated every trial to prevent participants from recoding the target shapes into locations on the wheel. Participants were instructed to report the second cued target from the wheel by moving the cursor to the part of the wheel where the remembered target was located and clicking on that part of the wheel. Once the cursor touched the wheel, a report probe was presented at the center of the screen that matched the stimulus value at the cursor location. The report probe updated continuously as the participant moved the cursor along the wheel to inform participants of the exact stimulus value that the cursor was placed on. Once participants clicked on the wheel, the wheel was removed from the screen and the report probe remained. Below the report probe, three response buttons labeled "High," "Low," and "No," respectively, were presented horizontally to one another, connected by a line. Participants were instructed to indicate how confident they were that their selection matched the second cued target by clicking the response button of choice. The accuracy of the memory and confidence reports was emphasized and was therefore performed without an imposed time limit.

To summarize, the task included two conditions—a target report following perceptual comparison (33% repeat cue, 33% switch cue) and a baseline target report with no probe presented (16.67% repeat cue, 16.67% switch cue). On trials where a probe was presented, it was always ±45° from one of the two targets, meaning it could fall at four possible distances from the to-be-reported target: ±45° (33%), ±90° (11%), ±135° (11%), or 180° (11%; Figure 1B).

Behavioral Analysis

For each trial, we computed the error in the memory report by subtracting the degree value of the memory report in the circular feature space from that of the target item that was presented originally to report. To provide directional information about the report error relative to the probe, we aligned the report errors across trials such that positive error values indicated report errors in the direction of the probe (i.e., signed report errors). To quantify the magnitude of any systematic biases in participants' memory reports, we computed the circular mean signed report error within the condition of interest. Note that this report bias could not be computed in the baseline condition due to the lack of an interfering probe being presented. In principle, this lack of a bias baseline prevents us from determining the directionality of any bias differences between prioritized and unprioritized memories. Therefore, any modulation in bias determined with this procedure is conservatively referred to as "protection" without any further specification of directional change. Report precision was calculated by computing the circular standard deviation of the raw (i.e., unsigned) report errors within the baseline condition.

To maximize power in our analyses, we report the results across all trials. When we constrained our analyses to only include trials where participants reported being highly confident in their memory report, the results were nearly identical (see the Supplemental Materials).

We report both frequentist (i.e., t values) and Bayesian (i.e., Bayes factors) statistics to allow for providing evidence in favor of null differences between conditions. Two-tailed t tests were employed to evaluate the statistical significance of any observed biases, as well as any differences in bias magnitude or report precision between conditions. The sign of the t statistic was used to determine the directionality of any differences that reached statistical significance. BF_{01} indicates evidence in favor of the null hypothesis, and BF_{10} indicates evidence in favor of the alternative hypothesis. To ensure that our results were not unduly influenced by any remaining outliers, we screened all of our dependent measures for values that were at least 3 *SD* above or below the mean. None of the results containing outliers were changed by including or excluding these outliers. Therefore, we report our results with the outliers included and flagged in the figures. In the Supplemental Materials, we include a table comparing the Bayes factors reported in our analyses when outliers are included versus excluded.

Modeling Analysis

To characterize any observed systematicity in participants' signed report errors, we constructed and fit a computational mixture model using MemToolbox (Suchow et al., 2013) that allows for interference by nontarget stimuli. The model is an extension of the target confusability competition (TCC) model proposed by Schurgin et al. (2020) fit with additional parameters that quantify the magnitude of representational shifts in the target, as well as the frequency of trials where observers misreport the probe during target reports (i.e., swap errors). We provide a brief description of the model and its construction below.

TCC Model

The TCC model is a signal detection model that accounts for variability in continuous report data by incorporating information about the global similarity structure within the stimulus space. The model assumes that encoding a target stimulus briefly enhances the familiarity of that target value in memory. However, unlike classic signal detection models, which assume that only the familiarity of the target is enhanced (Macmillan & Creelman, 2005), the TCC model assumes that this familiarity signal propagates to neighboring nontargets in the stimulus space as well. This spread in familiarity is assumed to follow a fixed psychophysical similarity function that is approximately exponential in shape when averaged across stimuli (Nosofsky, 1992; Shepard, 1987) and is able to be mapped using an independent perceptual task (Schurgin et al., 2020). When prompted to report a target value from memory, the model assumes that the observer selects the stimulus value that appears maximally familiar to them at that time. Due to noise that is accumulated during encoding and maintenance, nontarget values occasionally appear more familiar to the observer than the actual target value. As a result, memory reports tend to cluster around the location of the true target value across trials but are imperfect, nonetheless. This model has been shown to closely fit the data from a wide array of continuous report tasks, including those containing nontarget interference, and outperform most conventional models of VWM (Schurgin et al., 2020; J. R. Williams et al., 2022, 2023). For example, recent parameter recovery simulations showed that TCC is more robust against variations in memory strength and interstimulus distance when estimating swap errors, while other alternative models elicit pronounced skewing in their respective estimates (J. R. Williams et al., 2023).

When formalized, the model assumes that continuous reports along a circular wheel are equivalent to 360-alternative forced-choice tasks, where each stimulus value that occupies 1° of the wheel represents a choice. Given the distance, x, of a particular stimulus value from the target in the stimulus space, we let $f(x)$ represent how psychophysically similar the stimulus is to the encoded target and let d' represent the strength of the target in memory. From here, we let vector $(X_{-179}, \ldots, X_{180})$ represent the target distribution where each value in the vector is a normal random value with mean $d_x = d' \times f(x)$ and unit variance. Thus, the value reported by the observer, r, is as follows:

$$r \sim \operatorname{argmax}(X_{-179}, \ldots, X_{180}). \tag{1}$$

TCC Swap + Shift Model

To introduce the possibility of representational shifts and probabilistic swap errors to the TCC model, we adopted the TCC-swap model proposed by J. R. Williams et al. (2022) and appended an additional parameter that allows for shifting of the target value in the stimulus space. We refer to this as the TCC-swap + shift (TCC-SS) model. Unlike the original TCC model, which assumes that observers' reports are always based on similarity to the target, the TCC-SS model assumes that reports are sometimes based upon similarity to a nontarget—in this case, the probe. Therefore, the model is formalized to include separate familiarity distributions for the target and the probe with a parameter that determines which of the two distributions that the report will be drawn from. To account for systematic shifts in the target representation, the model allows for a global shift in the target familiarity distribution toward or away from the probe in the stimulus space.

Given the distances, x and y, of a particular stimulus value from the target and probe, respectively, we let $f(x)$ represent similarity to the target and let $g(y)$ represent similarity to the probe. We then let the normal random vector $(X_{-179}, \ldots, X_{180})$ represent the target distribution with means $d_x = d'_{\mathrm{T}} \times f(x + \theta)$ and unit variance and let $(Y_{-179}, \ldots, Y_{180})$ represent the probe distribution with means $d_y = d'_{\mathrm{P}} \times g(y)$ and unit variance. θ represents the mean of the target similarity function, which is set in terms of degrees offset from the true target value. In the case where there is no shifting in the target distribution, θ is 0. We let β represent the rate of swap errors. Thus, responses are modeled as follows:

$$\omega \sim \operatorname{Bernoulli}(\beta). \tag{2}$$

$$\begin{aligned} r \sim{} & \omega \times \operatorname{argmax}(Y_{-179}, \ldots, Y_{180}) \\ & + (1 - \omega) \times \operatorname{argmax}(X_{-179}, \ldots, X_{180}). \end{aligned} \tag{3}$$

The model freely estimates four of the above parameters (i.e., θ, β, d'_{T}, d'_{P}) using a maximum likelihood estimation procedure (Dempster et al., 1977). During parameter search, θ was allowed to vary from −180° to 180° in increments of 1°, β was allowed to vary from 0 (*i.e., 0% swap rate*) to 1 (*i.e., 100% swap rate*) in increments of 0.02, and both d' were allowed to vary from 0 to 4.5 in increments of 0.05. Note that the inclusion of separate d' parameters for the target and probe represents a departure from the original TCC-swap model, which uses a single d' to scale both distributions identically. This modification was intended to account for the sequential presentation of the target and the probe, as well as the unique task demands tied to each of them, which may have produced differences in their respective strength. Psychophysical similarity was set by a function mapped to the circular shape space by Robinson and Brady (2023) using perceptual matching data.

Results

Before measuring the retroactive influence of the probe on the remembered targets, we first sought to determine whether participants used the retro-cues reliably to prioritize their memories. As aforementioned, a common behavioral signature of attentional prioritization in memory is more precise reporting of prioritized memories than unprioritized memories (Gunseli et al., 2015; Pertzov et al., 2013; Souza & Oberauer, 2016; van Moorselaar et al., 2015; Z. Zhang & Lewis-Peacock, 2023a, 2023b). Thus, we began by measuring the precision of target reports in the baseline condition, where no probe was presented during maintenance. In doing so, we found that memory reports of prioritized targets were more precise than those of unprioritized memories (Figure 2), $t(47) = 6.91$, $p < .001$, Cohen's $d = 1.00$, $\mathrm{BF}_{10} = 1.12 \times 10^6$, consistent with previous research. This confirmed that participants prioritized their memories in accordance with the instructional retro-cues.

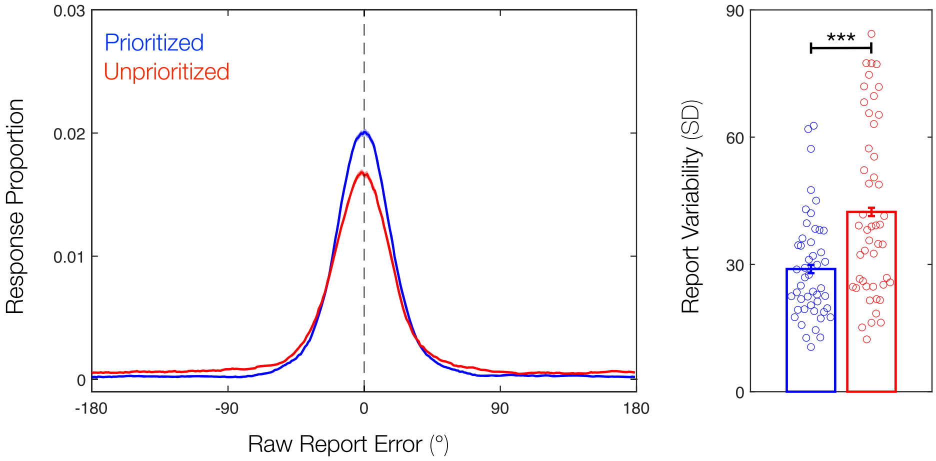
Having verified participants' use of the retro-cues, we then moved to characterize the retroactive influence of the probe on the reported target. We started by assessing how the physical similarity between the target and probe influenced the degree of interactions between them. To do this, we separated trials based on the probe's distance from the reported target in the stimulus space and measured the report bias. We grouped the probes into close (45°) and far (90°, 135°) distance bins to better balance the number of trials in our comparisons (see the Procedure section), and we removed probes that were sampled 180° opposite of the reported target since we could not recover information about the direction of report errors relative to the probe on these trials.

We found that memory reports were systematically biased toward the probe when it was close to the target (Figure 3A), $M = 6.91°$, $t(47) = 11.23$, $p < .001$, Cohen's $d = 1.62$, $\mathrm{BF}_{10} = 1.03 \times 10^{12}$, with moderate evidence that they were not biased when the probe was far (Figure 3A), $M = 0.61°$, $t(47) = 0.86$, $p = .392$, Cohen's $d = 0.12$, $\mathrm{BF}_{01} = 4.49$; even when we separated the 90°, $M = -0.03°$, $t(47) = -0.03$, $p = .974$, Cohen's $d > -0.01$, $\mathrm{BF}_{01} = 6.37$; and 135° conditions, $M = 1.30°$, $t(47) = 1.58$, $p = .120$, Cohen's $d = 0.23$, $\mathrm{BF}_{01} = 1.99$. When compared, we found that these reliable report biases in the close probe condition were significantly larger than those in the far probe condition as well (Figure 3A), $t(47) = 7.58$, $p < .001$, Cohen's $d = 1.09$, $\mathrm{BF}_{10} = 1.02 \times 10^7$. This modulation in report bias magnitude is consistent with the known nonlinear effects of physical similarity between stimuli in which biases increase from small to moderate distances (e.g., <75°; Rademaker et al., 2015; Teng & Kravitz, 2019) before plateauing and returning to baseline with increasing dissimilarity, as they did here (e.g., Mallett et al., 2020; Nemes et al., 2011, 2012; Saito et al., 2022; see also Lorenc et al., 2021). To then determine the effect of memory prioritization on these interactions, we separated trials in the close probe condition based on the priority of the reported target at the time of the perceptual comparison. In doing so, we found a significant report bias for both prioritized (Figure 3B), $M = 5.18°$, $t(47) = 9.55$, $p < .001$, Cohen's $d = 1.38$, $\mathrm{BF}_{10} = 6.08 \times 10^9$, and unprioritized memories (Figure 3B), $M = 9.00°$, $t(47) = 6.79$, $p < .001$, Cohen's $d = 0.98$, $\mathrm{BF}_{10} = 7.75 \times 10^5$, the prior of which was driven almost exclusively by probes that were perceived as similar to the target (see the Supplemental Materials). However, we found moderate evidence that the report bias for prioritized memories was significantly smaller than that for unprioritized memories (Figure 3B), $t(47) = -2.68$, $p = .010$, Cohen's $d = -0.39$, $\mathrm{BF}_{10} = 3.78$. These findings suggest that featural similarity alone was sufficient to produce a reliable bias in observers' memory reports, but that this bias was attenuated when the memory was prioritized.

To infer the underlying source of these behavioral report biases, we fit a computational mixture model to the data that is capable of parsing different sources of systematic error (see the Modeling Analysis section). As shown in Figure 4A, systematic biases in

Figure 2

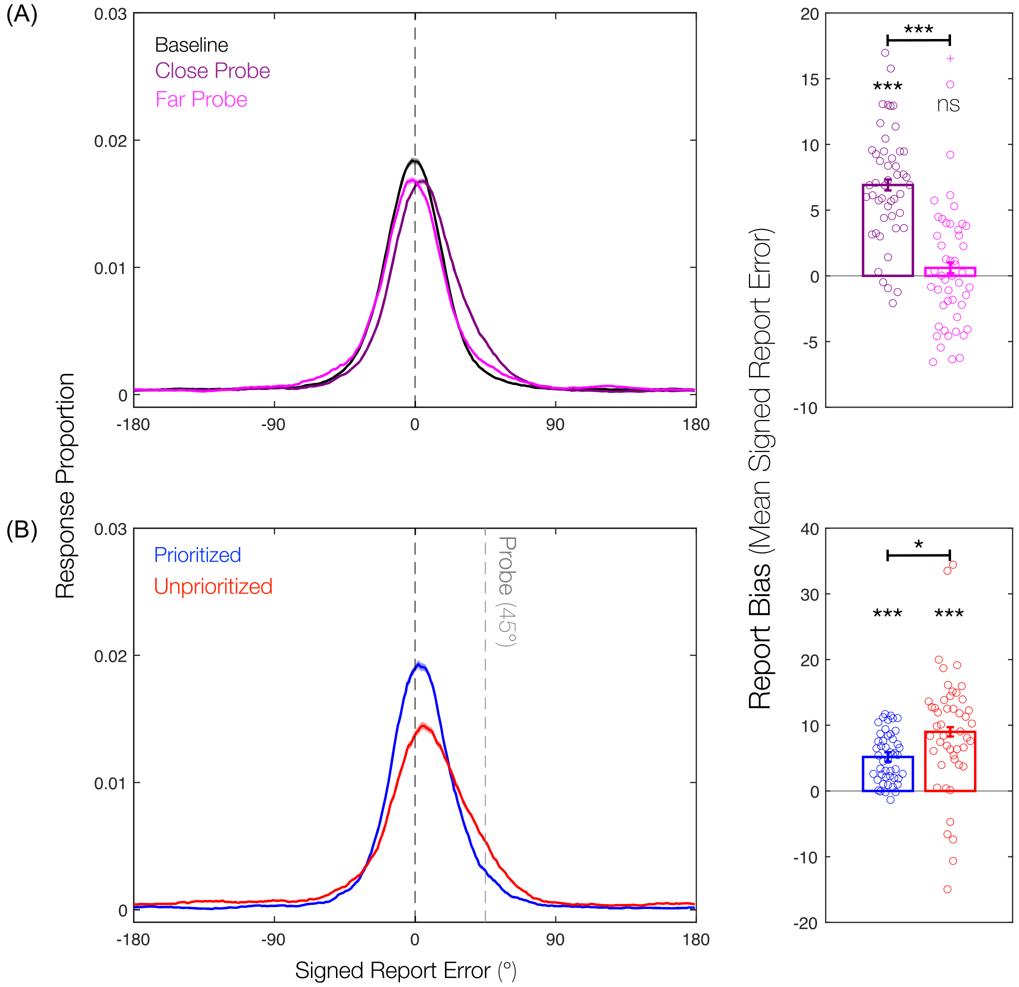
Enhanced Report Precision Following Prioritization



Note. Distributions (left) and standard deviations (right) of raw report errors in the baseline condition for prioritized and unprioritized memories. For illustration purposes, we plotted the proportion of errors for a given offset in the distribution by calculating the mean response proportion across a 30° window centered at the offset value. The vertical black dashed line indicates the location of the target zero-centered across trials. Shaded regions around the mean line indicate within-subject standard errors. Colored circles in the bar plot indicate the standard deviation for a single participant. Error bars indicate within-subject standard errors. See the online article for the color version of this figure.
$^{***}p < .001$.

Figure 3

Effect of Probe Distance and Memory Prioritization on Report Bias



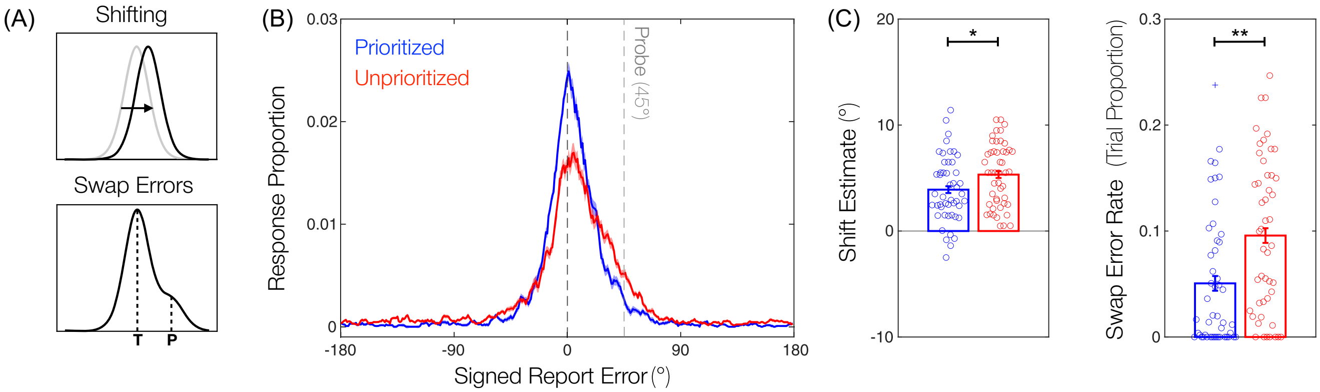
Note. Signed report error distributions for (A) close and far probe trials collapsed across cueing conditions and for (B) prioritized and unprioritized memories within the close probe condition. For illustration purposes, we plotted the proportion of errors for a given signed offset by calculating the mean response proportion across a 30° window centered at the offset value. The vertical dashed lines indicate the location of the target zero-centered across trials and the location of the close probe in (B). Shaded regions around the mean line indicate within-subject standard errors. Bar plots to the right of the distributions show the mean signed report error in each condition. Colored circles indicate the mean error for a single participant. Colored "+" symbols indicate outliers. Error bars indicate within-subject standard errors. *ns* = not significant. See the online article for the color version of this figure.
$^{*}p < .05$. $^{***}p < .001$.

continuous reports are known to stem from shifts in the target representation and from probabilistic misreports of the nontarget probe (i.e., swap errors). Importantly, these distinct sources of systematic error reflect different degrees of interaction between new and existing visual representations. Unlike representational shifts, where participants maintain correspondence with stimulus-specific information about the target stimulus, swap errors reflect a more catastrophic form of interference where participants are assumed to lose correspondence with the target and report information that is independent (Mallett et al., 2022; Pratte, 2019). Compared to prioritized memoranda, unprioritized memories may be susceptible to larger representational shifts, more frequent swap errors, or some combination of the two.

Upon fitting the model to the data, we found evidence for representational shifting and swap errors in both prioritized and unprioritized memoranda (see Figure 4B and 4C). Critically, the model identified anecdotal evidence that prioritized memoranda were associated with smaller representational shifts (Figure 4C), $t(47) = -2.23$, $p = .031$, Cohen's $d = -0.32$, $\mathrm{BF}_{10} = 1.47$, and strong evidence that they were associated with less frequent swap errors (Figure 4C), $t(47) = -3.26$, $p = .002$, Cohen's $d = -0.47$, $\mathrm{BF}_{10} = 1.49 \times 10$, compared to unprioritized memoranda. Indeed, while

Figure 4

Model Evidence for Smaller Shifts and Less Frequent Swap Errors With Prioritization



Note. (A) Schematics illustrating changes in the visual form of observers' report error distributions when contaminated by representational shifts and swap errors. When the target is shifted, the distribution exhibits a global shift away from the location of the target toward (or away from) the nontarget probe. When a swap error occurs, observers misreport the features of the nontarget probe instead of the target during memory reporting. If swap errors occur with enough frequency, misreports of the probe begin to produce a second mode in the distribution. (B) Signed report error distribution for prioritized and unprioritized memories within the close probe condition. To better illustrate any bimodality resulting from swap errors, we plotted the proportion of responses for a given signed offset by calculating the mean response proportion across a 6° window centered at the offset value. The vertical dashed lines indicate the location of the target zero-centered across trials and the location of the probe, respectively. Shaded regions around the mean lines indicate within-subject standard errors. (C) Bar plots depicting estimates of representational shift (left) and swap error frequency (right) from the target confusability competition-swap + shift model. Colored circles indicate the mean estimate for a single participant. Colored "+" symbols indicate outliers. Error bars indicate within-subject standard errors. See the online article for the color version of this figure.

$^{*}p < .05$. $^{**}p < .01$.

reductions in representational shifting were relatively small in magnitude (prioritized: $M = 3.90°$, unprioritized: $M = 5.32°$, difference: $M = 1.42°$), the incidence of swap errors was nearly halved by prioritization (prioritized: $M = 5.06\%$ trials, unprioritized: $M = 9.57\%$ trials, difference: $M = 4.51\%$). These modeling results, in conjunction with behavior, suggest that attention does confer protective benefits to memory in cases of complete prioritization and that this is true even when observers are unable to drop unprioritized memoranda from VWM.

Discussion

In Experiment 1, we used a double-serial retro-cue paradigm to address several gaps in previous literature regarding the effect of memory prioritization on retroactive interactions between new and existing visual representations. In particular, we used a combination of two independent retro-cues and a task-relevant perceptual comparison to incentivize the complete prioritization (and deprioritization) of memoranda while prohibiting the ability to drop items from VWM. In doing so, we found that prioritized memories were reported more precisely and showed weaker report biases toward new visual inputs compared to unprioritized memories. Using modeling, we show that these weaker report biases following prioritization may be underpinned by a combination of weaker representational shifts and fewer swap errors. These results support a longstanding protective attentional account of prioritization (Oberauer, 2002; Oberauer & Lin, 2017) and qualify recent protective dormancy hypotheses by limiting their scope only to circumstances where priority is uncertain (Lorenc et al., 2021; Z. Zhang & Lewis-Peacock, 2023a, 2023b). In Experiment 2, we sought to provide a more comprehensive account of the interactions between new and existing visual representations during prioritization by appending a measure of proactive report bias to the current paradigm.

Experiment 2

Method

Transparency and Openness

Same as Experiment 1.

Participants

All participants in the experiment were undergraduate students at the University of Toronto Mississauga that reported normal or corrected-to-normal visual acuity and color vision. Each participant provided informed consent in accordance with the procedures approved by the research ethics board at the University of Toronto.

Participants were recruited on a weekly basis until the sample size exceeded the minimum number identified by the power analysis performed in Experiment 1 ($n = 34$). This resulted in the recruitment of 40 participants. Each participant reported their demographic information in a response window that appeared before the start of the experimental script. For gender identity, participants selected between male, female, and prefer not to answer. Participants submitted their age in years. Information about racial and ethnic identity was not collected. One participant was excluded for failing to complete at least 90% of the similarity judgments. Two participants were excluded for poor overall performance on the task, as indicated by a mean absolute report error in the baseline condition that was >2.5 *SD* above the mean of the group. One participant was excluded for failing to report the probe stimulus reliably when cued, as indicated by a mean absolute report error on probe report trials that was >2.5 *SD* above the mean of the group. Data collected from the remaining 36 participants were submitted to analysis. Among this remaining sample, 27 participants reported their gender as female and nine as male. The mean age of the sample was 19.00 years.

Apparatus and Stimuli

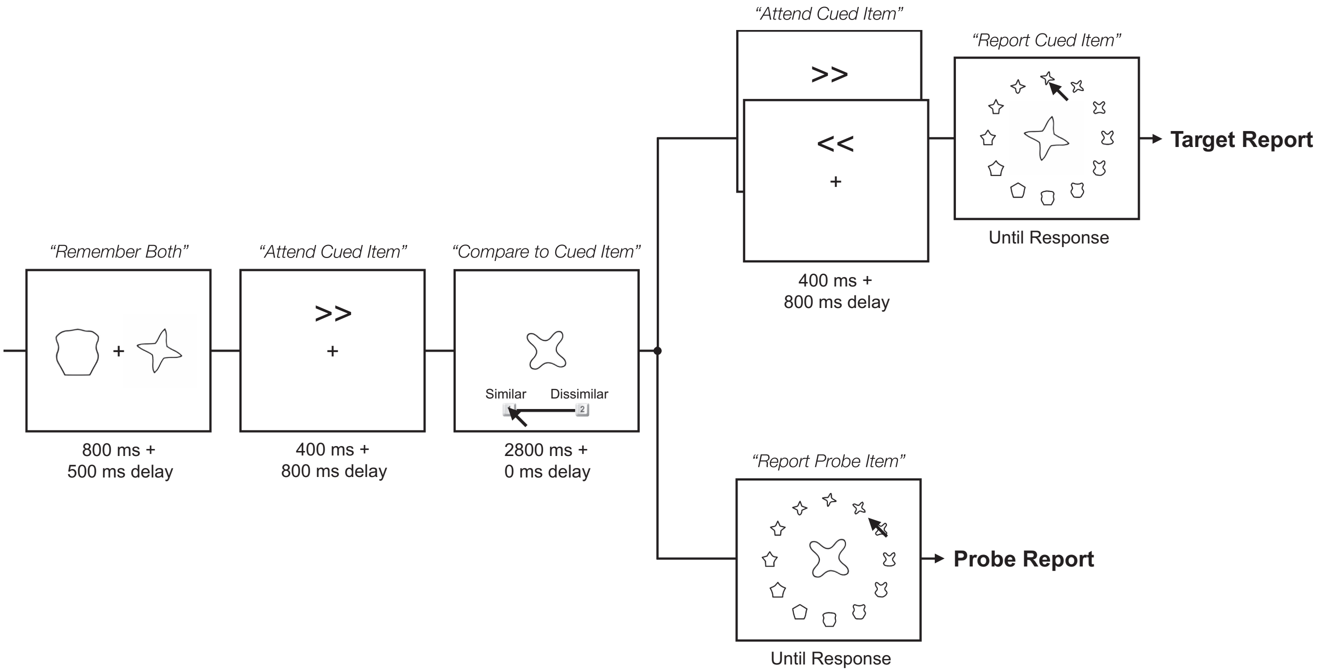
Identical to Experiment 1.

Procedure

The task procedure in Experiment 2 is shown in Figure 5. It was similar to that of Experiment 1, with the following exceptions:

1. The possible interitem distance of 180° between targets was removed, meaning that the distance between targets was always fixed to ±135° and counterbalanced in the clockwise and counterclockwise directions.
2. Due to the above change in target sampling, the possible probe distance of ±135° from one of the targets was eliminated.
3. To ensure that none of the findings in Experiment 1 could be attributed to variability in the duration of the probe stimulus, which was set to offset at the time of the similarity judgment, the probe in Experiment 2 remained onscreen for the entire 2,800-ms presentation period regardless of response.
4. On 25% of trials, participants were prompted to report the probe rather than one of the targets (Figure 5). The probe report was designed with two goals in mind. First, we wanted to ensure that the probe stimulus was presented for the same duration across trials untethered to the duration of the report. Second, we wanted to minimize the time that participants held the probe in VWM, since this could evoke additional interactions between the target and probe that were not the focus of the present study (e.g., maintenance repulsion; Bae & Luck, 2017; Chunharas et al., 2022; Scotti et al., 2021). To do this, we presented the probe stimulus for the same duration specified above (i.e., 2,800 ms) and then cued the probe report by presenting the shape wheel immediately at the time of probe offset without a second retro-cue. Note that removing the probe stimulus before the report forced observers to recall the probe representation from mind, thus making it impossible to determine whether any reported biases emerged at the time of perception or postperceptually. Therefore, biases in the probe report are conservatively referred to as

Figure 5
Double-Serial Retro-Cue Paradigm With Target and Probe Reporting



Note. The trial procedure for the experimental conditions in the paradigm. After encoding two target items into memory simultaneously, a retro-cue indicates to participants which of the two targets is to be prioritized for use in an impending perceptual comparison. During the comparison, a probe stimulus is presented, and participants judge its similarity to the cued target. In 75% of trials, a second retro-cue informs participants which of the two targets is to be reported from memory (Target Report, top row). In the remainder of trials, the continuous wheel immediately onsets after the probe, which cues participants to report the probe instead (Probe Report, bottom row). In each type of trial, the participant selects the cued item from the wheel with the mouse. The "+" symbols used in the stimulus durations are synonymous with "followed by."

proactive report biases without any further specification of their timing.

5. Given that the results of Experiment 1 were preserved regardless of observers' confidence in their memory report (see the Supplemental Materials), the confidence report was removed.

To summarize, the task included three conditions—target report following perceptual comparison (25% repeat cue, 25% switch cue), probe report following perceptual comparison (25%), and target report with no probe presented (12.5% repeat cue, 12.5% switch cue). On trials where a probe was presented, it was always ±45° from one of the two targets, meaning it could fall at three possible distances from the to-be-reported target: ±45° (37.5%), ±90° (18.75%), and 180° (18.75%).

Behavioral Analysis

To ensure once again that our results were not unduly influenced by any remaining outliers, we screened all of our dependent measures for values that were at least 3 *SD* above or below the mean. None of the results containing outliers were changed by including or excluding these outliers, except for one comparison where the outlier was removed and explicitly noted by a footnote. All other results are reported with outliers included and flagged in the figures. In the Supplemental Materials, we include a table comparing the Bayes factors reported in our analyses when these outliers are included versus excluded.

Modeling Analysis

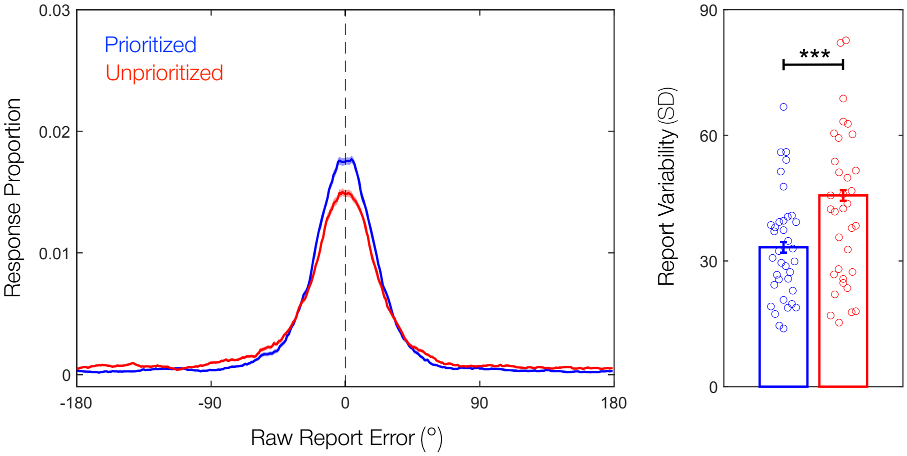
Identical to Experiment 1.

Results

We began by measuring the precision of the observer's target reports in the baseline condition to confirm their reliable use of the retro-cues to prioritize their memories. In doing so, we again found that memory reports of prioritized targets were more precise than those of unprioritized targets (Figure 6), $t(35) = -4.90$, $p < .001$, Cohen's $d = -0.82$, $BF_{10} = 9.93 \times 10^2$. From here, we then sought to determine whether the retroactive interactions between the target and probe that were observed in Experiment 1 were observed again in Experiment 2. We again separated trials based on the probe's distance from the reported target in the stimulus space and measured the report bias. As before, we found that memory reports were systematically biased toward the probe when it was close to the target (Figure 7A), $t(35) = 10.54$, $p < .001$, Cohen's $d = 1.76$, $BF_{10} = 4.02 \times 10^9$, with moderate evidence that they were not when the probe was further away (Figure 7A), $t(35) = -1.08$, $p = .286$, Cohen's $d = -0.18$, $BF_{01} = 3.25$. When directly compared, report biases observed in the close probe condition were significantly larger than those in the far probe condition, $t(35) = 6.70$, $p < .001$, Cohen's $d = 1.12$, $BF_{10} = 1.58 \times 10^5$. When we separated trials in the close probe condition based on the priority of the reported target at the time of the perceptual comparison, we found a significant report bias in both prioritized (Figure 7B), $t(35) = 7.60$, $p < .001$, Cohen's $d = 1.27$, $BF_{10} = 1.93 \times 10^6$, and unprioritized memories (Figure 7B), $t(35) = 9.32$, $p < .001$, Cohen's $d = 1.55$, $BF_{10} = 1.89 \times 10^8$, the prior of which was again driven almost exclusively by probes that were perceived as similar to the target (see the Supplemental Materials). However, as in Experiment 1, the report bias for prioritized memories was significantly smaller than that for unprioritized memories, $t(35) = -4.08$, $p < .001$, Cohen's $d = -0.68$, $BF_{10} = 1.09 \times 10^2$. When fit with the TCC-SS model, we found strong evidence that prioritized memories were

Figure 6

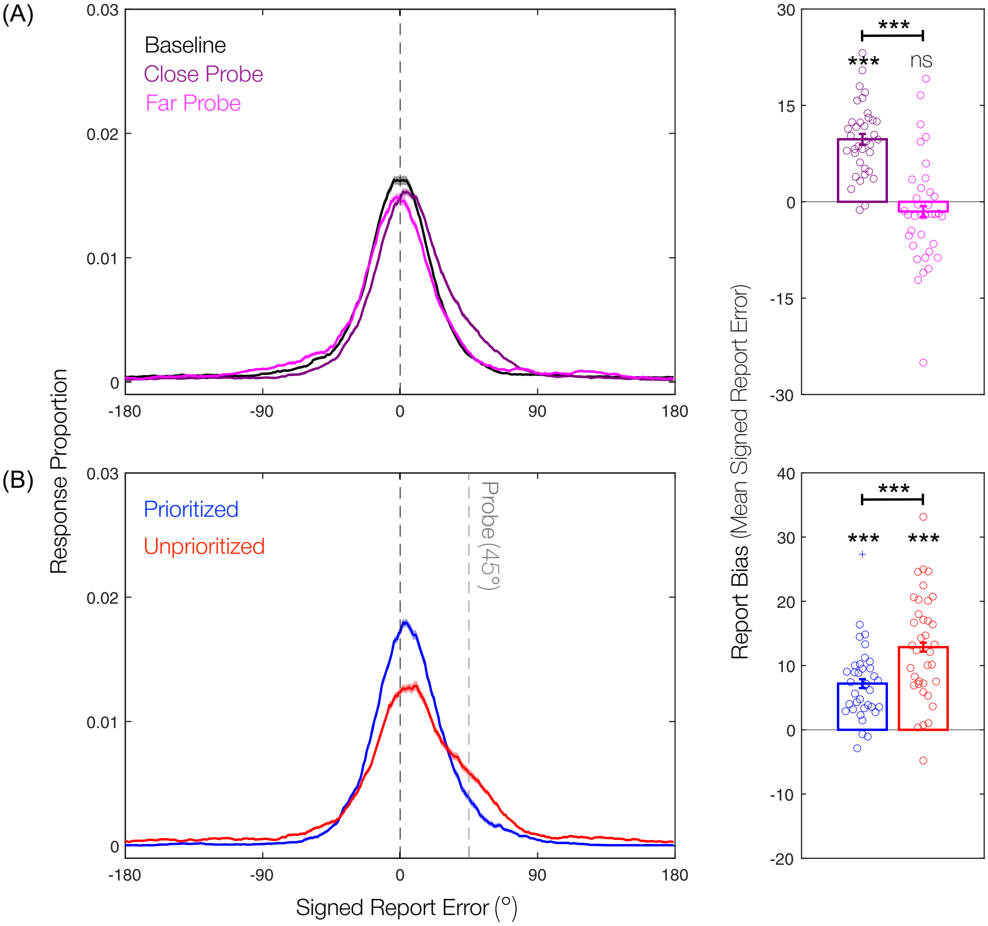
Replication of Enhanced Report Precision Following Prioritization



Note. Distributions (left) and standard deviations (right) of raw report errors in the baseline condition for prioritized and unprioritized memories. For illustration purposes, we plotted the proportion of errors for a given offset in the distribution by calculating the mean response proportion across a 30° window centered at the offset value. The vertical black dashed line indicates the location of the target zero-centered across trials. Shaded regions around the mean line indicate within-subject standard errors. Colored circles in the bar plot indicate the standard deviation for a single participant. Error bars indicate within-subject standard errors. See the online article for the color version of this figure.
*** $p < .001$.

Figure 7

Replication of Retroactive Report Bias Based on Probe Distance and Memory Prioritization



Note. Signed report error distributions for (A) close and far probe trials collapsed across cueing conditions and for (B) prioritized and unprioritized memories within the close probe condition. For illustration purposes, we plotted the proportion of errors for a given signed offset by calculating the mean response proportion across a 30° window centered at the offset value. The vertical dashed lines indicate the location of the target zero-centered across trials and the location of the close probe in (B). Shaded regions around the mean line indicate within-subject standard errors. Bar plots to the right of the distributions show the mean signed report error in each condition. Colored circles indicate the mean error for a single participant. Colored "+" symbols indicate outliers. Error bars indicate within-subject standard errors. *ns* = not significant. See the online article for the color version of this figure.
$^{***}p < .001$.

associated with smaller representational shifts (Figure 8), $t(35) = -3.88$, $p < .001$, Cohen's $d = -0.65$, $BF_{10} = 6.61 \times 10$, and less frequent swap errors (Figure 8), $t(34) = -3.61$, $p < .001$, Cohen's $d = -0.61$, $BF_{01} = 3.24 \times 10$,[1] than unprioritized memories. These model findings, in conjunction with behavior, replicate the retroactive interactions observed in Experiment 1.

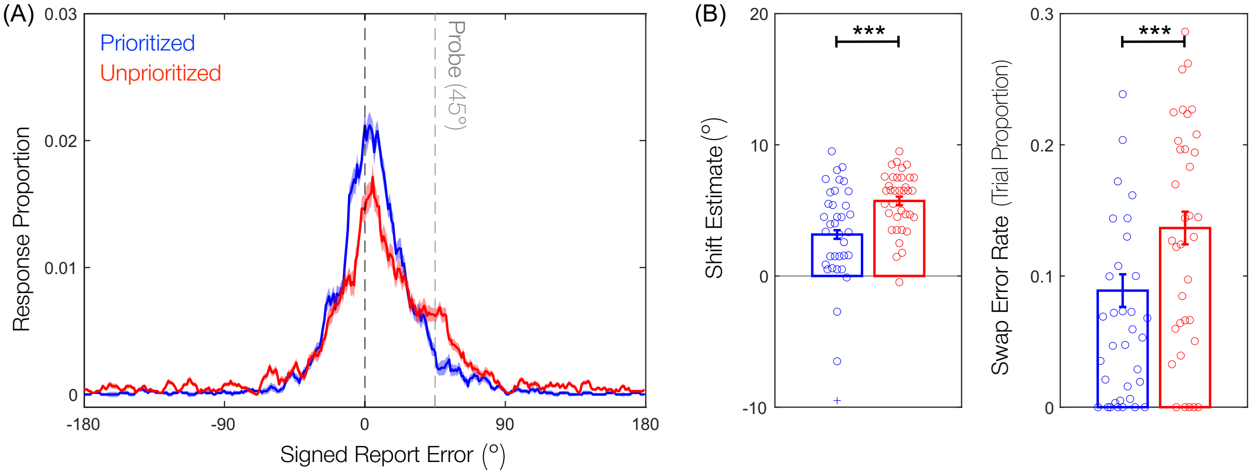
We then moved to characterize the proactive interactions between memories and new visual inputs by measuring systematic biases in observers' reports of the probe. Much like the retroactive biases observed in target reports, we found that probe reports were biased toward targets that were prioritized (Figure 9A), $t(35) = 5.05$, $p < .001$, Cohen's $d = 0.84$, $BF_{10} = 1.53 \times 10^3$, and those that were unprioritized (Figure 9A), $t(35) = 3.29$, $p = .002$, Cohen's $d = 0.55$, $BF_{10} = 1.50 \times 10$. However, unlike retroactive biases, we found moderate evidence that proactive report biases were comparable between prioritized and unprioritized memories, $t(35) = 1.03$, $p = .311$, Cohen's $d = 0.17$, $BF_{01} = 3.43$, suggesting that memory priority did not meaningfully modulate their magnitude. To rule out the trivial possibility that observers might have frequently misreported the prioritized target instead of the probe because they were unsure if a second retro-cue was presented, we fit the TCC-SS model (see Figure 9B). In doing so, we

---

[1] Outlier screening identified one participant with an extremely high swap rate in the prioritized condition (=63% of trials, $z = 4.43$) that was not detected by our baseline screening procedure. When appended to the data, this outlier was extreme enough to mask the clear difference in swap rate between prioritized and unprioritized targets, $t(35) = -1.92$, $p = .063$, Cohen's $d = -0.32$, $BF_{01} = 1.08$. As such, this outlier was removed from the comparison reported in the main text.

Figure 8

Replication of Smaller Retroactive Shifts and Less Frequent Swap Errors With Prioritization

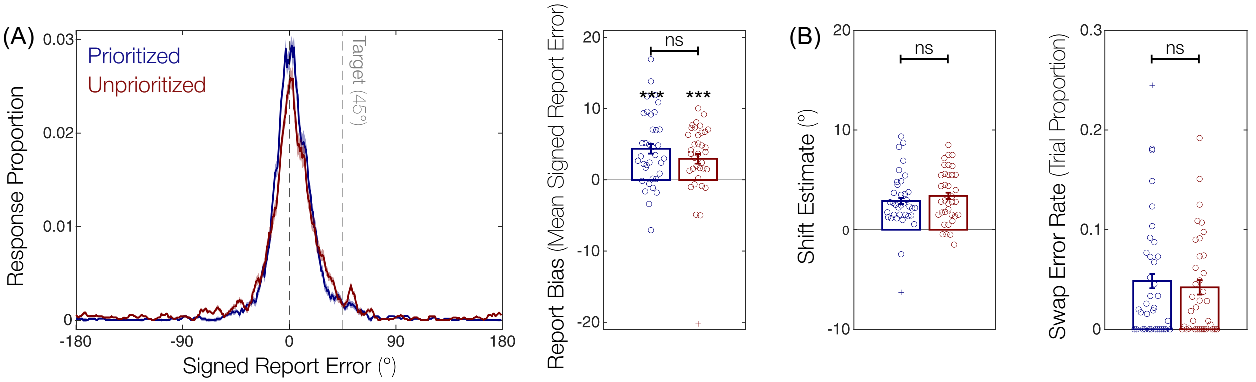


Note. (A) Signed report error distribution for prioritized and unprioritized memories within the close probe condition. To better illustrate any bimodality resulting from swap errors, we plotted the proportion of responses for a given signed offset by calculating the mean response proportion across a 6° window centered at the offset value. The vertical dashed lines indicate the location of the target zero-centered across trials and the location of the probe, respectively. Shaded regions around the mean lines indicate within-subject standard errors. (B) Bar plots depicting estimates of representational shifting (left) and swap error frequency (right) from the target confusability competition-swap + shift model. Colored circles indicate the mean estimate for a single participant. Colored "+" symbols indicate outliers. Error bars indicate within-subject standard errors. One outlier in the prioritized swap rate condition did not fit in the figure (=0.63). See the online article for the color version of this figure.
$^{***} p < .001$.

found moderate evidence that report biases toward prioritized and unprioritized targets were underpinned by comparable representational shifts, $t(35) = -0.83$, $p = .415$, Cohen's $d = -0.14$, $\mathrm{BF}_{01} = 4.07$, and comparable swap error rates, $t(35) = 0.44$, $p = .665$, Cohen's $d = 0.07$, $\mathrm{BF}_{01} = 5.11$. These findings suggest that prioritized and unprioritized memories exerted a comparable proactive influence on observers' reports of the incoming probe despite clear differences in their respective susceptibility to retroactive interactions by that probe.

Figure 9

Model Evidence for Comparable Proactive Interactions Regardless of Memory Priority



Note. (A) Signed report error distributions for probes within the close target condition separated by prioritized and unprioritized targets. To better illustrate any bimodality resulting from swap errors, we plotted the proportion of responses for a given signed offset by calculating the mean response proportion across a 6° window centered at the offset value. The vertical dashed lines indicate the location of the probe zero-centered across trials and the location of the target, respectively. Shaded regions around the mean lines indicate within-subject standard errors. Bar plots to the right of the distributions show the mean signed report error in each condition. Colored circles indicate the mean error for a single participant. Error bars indicate within-subject standard errors. (B) Bar plots depicting estimates of representational shifting (left) and swap error frequency (right) from the target confusability competition-swap + shift model. Colored circles indicate the mean estimate for a single participant. Colored "+" symbols indicate outliers. Error bars indicate within-subject standard errors. *ns* = not significant. See the online article for the color version of this figure.
$^{***} p < .001$.

General Discussion

Research suggests that memories and new visual inputs interact bidirectionally, producing retroactive and proactive biases in their respective behavioral reports (Barbosa & Compte, 2020; Fischer & Whitney, 2014; Kiyonaga et al., 2017; Liberman et al., 2014; Mallett et al., 2020; Rademaker et al., 2015; Saito et al., 2022; Teng & Kravitz, 2019). In this study, we sought to characterize how these interactions change when memories are prioritized for use in behavior. We found that prioritization modulated retroactive interactions between memories and a perceptual probe but not their proactive interactions. Behavioral reports of prioritized memories were more precise and contained smaller attractive biases toward the probe than those of unprioritized memories, but behavioral reports of the probe were biased comparably toward each of them. These findings corroborate a modest protective role of attention over memory (Oberauer, 2002; Oberauer & Lin, 2017) and elucidate a clear behavioral asymmetry in the reciprocal interactions between new and existing visual representations.

Previous studies have found inconsistent evidence regarding the effects of prioritization on retroactive interference. Some suggest that prioritized memories are partially shielded against interactions with visual inputs via top–down attention (Barth & Schneider, 2018; Makovski & Jiang, 2007; Makovski & Pertzov, 2015; Schneider et al., 2017; Souza et al., 2016; van Moorselaar et al., 2015), while others suggest that unprioritized memories are more likely to evade interactions with new visual inputs by relying on a separate neural substrate (Allen & Ueno, 2018; Hitch et al., 2018; Hu et al., 2014, 2016; Lorenc et al., 2021; Z. Zhang & Lewis-Peacock, 2023a, 2023b). While some of these inconsistencies can be attributed to differences in report procedures between studies (e.g., discrete recognition vs. continuous recall), we focused here on systematic differences in the types of retro-cues that were used to incentivize memory prioritization. Those cues demonstrating a protective benefit were often fully valid but may have allowed observers to discard unprioritized memories from mind (Kuo et al., 2012; Lorenc et al., 2021; Schneider et al., 2017; Souza et al., 2014; M. Williams & Woodman, 2012), while those cues demonstrating susceptibility were often uncertain and may have influenced the reliability of prioritization (Gould et al., 2011; Gunseli et al., 2015; Wang & van Ede, 2025; M. Williams & Woodman, 2012). By using a double-serial retro-cue paradigm in the present study that incentivized the complete prioritization of one memory item for use in an explicit comparison, we provide a potential resolution to these conflicting findings. First, we show that, when prioritized with complete certainty, prioritized memoranda are indeed relatively shielded against perceptual interference compared to unprioritized memoranda, consistent with protective attentional accounts (Oberauer & Lin, 2017). Second, because observers were unable to simply discard the unprioritized target from VWM, we confirm that these protective benefits are not attributable to changes in memory set size, as they occur even when unprioritized memoranda continue to be maintained.

By comparison, we found that probe reports were comparably biased toward prioritized and unprioritized memoranda, suggesting that proactive interactions were not strongly influenced by memory priority. To find an asymmetrical effect of memory priority across retroactive and proactive report biases is somewhat surprising, given that these biases are assumed to occur during a common set of reciprocal interactions between new and existing visual representations. We speculate that this asymmetry could reflect a distinction between the *functional* memory states that are conferred by behavioral relevance (Cowan, 2001; Oberauer, 2002) and their corresponding *neural* states of representation (for a detailed theoretical perspective, see Stokes et al., 2020). In the last decade, researchers have hypothesized about several different mechanisms in the brain that might be capable of implementing functionally "active" (i.e., prioritized) representations that guide ongoing processing and functionally "latent" (i.e., unprioritized) representations that are stored but unengaging (e.g., Christophel et al., 2018; LaRocque et al., 2013; Lewis-Peacock et al., 2012; Panichello & Buschman, 2021; Spaak et al., 2017). Some accounts propose a dichotomous mapping of these functional states to two qualitatively distinct mechanisms, such as recurrent neural firing and transient synaptic plasticity (Barak & Tsodyks, 2014; Mongillo et al., 2008), while others assert a single mechanism where functional states are marked by differing levels of neural activation instead (Barbosa et al., 2020; Schneegans & Bays, 2017). Most tests of these accounts have relied on neural evidence collected from tasks where observers shift their memories between active and latent functional states using top–down attentional selection. For example, in one of the first studies demonstrating correspondence between functional and neural states, the researchers trained a multivariate classifier to decode the category of stimuli stored in VWM while observers completed a double-serial retro-cue paradigm (Lewis-Peacock et al., 2012). Unlike the present study, the targets on each trial were drawn from distinct categories (e.g., words vs. oriented lines) and used in recognition judgments with probes that were identical or highly dissimilar. Prior to the first retro-cue, both memories elicited sustained, decodable activity, consistent with an active neural mechanism. When the first retro-cue instructed observers to prioritize one of the two memoranda, the prioritized memory remained decodable, but decoding of the unprioritized item dropped to chance, suggesting that it was no longer represented by the same active mechanism. Critically, however, when the second retro-cue then instructed observers to shift the unprioritized item back into a prioritized state, decoding of the previously unprioritized item was restored, suggesting that the item was maintained by a more latent neural mechanism before returning to an active neural state. In this example, and others like it, functional and neural states are yoked together by design, perhaps leading to the assumption that the two states are synonymous, such that an active neural state is universally indicative of an active functional state, and vice versa.

However, if functional and neural states do not share a perfect correspondence and are only partially dependent upon one another, this could help to explain the findings in the present study. When selected by internal attention, prioritized memories were able to be used in explicit comparisons and interact with new visual inputs, suggesting that they were active in both function and representation, consistent with previous studies. By contrast, unprioritized memories were not used in behavior, suggesting functional latency, yet still interacted with new visual inputs. One possibility is that these unprioritized representations were initially latent across both states but were reactivated at the neural level—perhaps involuntarily—by featural overlap with the probe. Indeed, previous research has already shown that latent neural representations can be reactivated by featural overlap with a cue (Printzlau et al., 2022; van Ede et al., 2020) and by other less specific forms of exogenous stimulation, such as impulses coming from magnetic stimulation (Rose et al., 2016) and irrelevant visual masks (Wolff et al., 2017). Once reactivated, these unprioritized memories may have interacted with new visual inputs in a manner unimpeded by the protective effects of top–down attention, resulting in retroactive report biases that were larger than those observed in prioritized memories. At the same time, because attentional allocation toward the perceptual probe was unrelated to attentional allocation within memory, an active neural state shared across prioritized and unprioritized memories may have produced comparable proactive report biases. That being said, observers' reports of the probe did appear slightly more precise when the probe was closer to the prioritized item than the unprioritized item (see Figure 9A), suggesting that preactivation of the prioritized item might have conferred a proactive boost in probe processing that was analogous to the probe's reactivation of the unprioritized target (Gayet et al., 2017; Serences, 2016).

A second, nonexclusive possibility is that differences in retroactive and proactive report biases emerged after interference had ended. In our modeling analyses, a significant amount of the modulation in retroactive report bias was explained by more frequent swap errors during reports of unprioritized memoranda. Unlike representational shifts, where observers' reports of the target are slightly contaminated, swap errors are assumed to reflect lost correspondence with the target entirely (Mallett et al., 2022; Pratte, 2019). Recent evidence from nonhuman primates suggests that this lost correspondence occurs primarily when observers incorrectly select information from VWM for reports (Alleman et al., 2024). This could imply that swap errors in the present study emerged after the second retro-cue when observers attempted to reorient their attention to the unprioritized target. However, this would not guarantee that retroactive interactions were comparable at the time of interference, as the model also concluded that unprioritized memories exhibited larger representational shifts as well, which may have occurred at the time of the comparison.

Some of the ostensible contradictions between the present findings and extant literature may also be aptly explained by assuming only partial dependence between functional and neural states. For example, the finding that unprioritized memoranda produce proactive biases in reports of new visual inputs appears inconsistent with the finding that they do not produce reliable attentional guidance during search (Greene et al., 2015; Mallett & Lewis-Peacock, 2018; Olivers et al., 2011; Ort et al., 2018, 2019; Ort & Olivers, 2020; van Loon et al., 2017; but see also Bahle et al., 2020; Beck et al., 2012). However, if one assumes that an unprioritized item can be reactivated at the neural level without obligatory functional reactivation (e.g., via reflexive reorienting of attention), then it is perfectly conceivable that unprioritized representations can sometimes produce report biases despite not automatically guiding search. In cases where unprioritized memories are less likely to be reactivated at the neural level, such as when visual inputs are highly dissimilar or successfully ignored (Rademaker et al., 2015; Saito et al., 2022; Teng & Kravitz, 2019), then the memory may not interact with new visual inputs at all, resulting in patterns of behavior that are more consistent with the protective dormancy hypothesis. To arbitrate among this myriad of hypotheses, which are speculative at present, future research should consider pairing the valid cue manipulation used here with a neuroimaging technique to extract direct measures of neural states during memory prioritization and interference.

Constraints on Generality

The goal of the present study was to provide a generalized understanding of how the deployment of internal attention modulates interactions between new and existing visual representations. We report the findings from behavioral tasks performed by undergraduate students from a reputable Canadian university whose cognitive abilities were likely higher than that of the average population. The observed range of performance may therefore be partially constrained by the nature of the sample and less representative of the variability present in less educated populations and in older adults. For example, due to age-related declines in both attention and working memory (Hasher & Zacks, 1988; Naveh-Benjamin & Cowan, 2023; Peich et al., 2013; Salthouse, 2016), we might expect that overall performance on these tasks will be worse across conditions in older adults. However, because our conclusions primarily rest upon the differences in performance between conditions for a given individual, we expect that the patterns critical to our research question would still hold (e.g., differences in report bias between prioritized vs. unprioritized memoranda). Future work should consider incorporating a more diverse sample of participants to demonstrate this generalizability more explicitly.

Conclusion

In the present study, we combined a double-serial retro-cue paradigm with explicit perceptual comparisons to examine the effect of memory prioritization on well-established interactions between new and existing visual representations. In doing so, we found evidence for an asymmetric pattern in observers' behavioral reports. In the retroactive direction, reports of prioritized memories were more precise and exhibited smaller attractive biases toward perceptual inputs than unprioritized memories, consistent with a protective role of internal attention. Conversely, in the proactive direction, reports of the perceptual probe showed comparable biases toward prioritized and unprioritized memoranda. We speculate that these findings, and those reported previously, can be accounted for by a theoretical framework that assumes only partial dependence between functional memory states that are conferred by top–down control and their concomitant neural states.

References

Adam, K. C. S., Rademaker, R. L., & Serences, J. T. (2022). Evidence for, and challenges to, sensory recruitment models of visual working memory. In T. F. Brady & W. A. Bainbridge (Eds.), *Visual memory* (1st ed., pp. 5–25). Routledge. https://doi.org/10.4324/9781003158134-2

Adam, K. C. S., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology*, *97*, 79–97. https://doi.org/10.1016/j.cogpsych.2017.07.001

Alleman, M., Panichello, M., Buschman, T. J., & Johnston, W. J. (2024). The neural basis of swap errors in working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(33), Article e2401032121. https://doi.org/10.1073/pnas.2401032121

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

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

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

Bae, G.-Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General*, *144*(4), 744–763. https://doi.org/10.1037/xge0000076

Bahle, B., Thayer, D. D., Mordkoff, J. T., & Hollingworth, A. (2020). The architecture of working memory: Features from multiple remembered objects produce parallel, coactive guidance of attention in visual search. *Journal of Experimental Psychology: General*, *149*(5), 967–983. https://doi.org/10.1037/xge0000694

Barak, O., & Tsodyks, M. (2014). Working models of working memory. *Current Opinion in Neurobiology*, *25*, 20–24. https://doi.org/10.1016/j.conb.2013.10.008

Barbosa, J., & Compte, A. (2020). Build-up of serial dependence in color working memory. *Scientific Reports*, *10*(1), Article 10959. https://doi.org/10.1038/s41598-020-67861-2

Barbosa, J., Stein, H., Martinez, R. L., Galan-Gadea, A., Li, S., Dalmau, J., Adam, K. C. S., Valls-Solé, J., Constantinidis, C., & Compte, A. (2020). Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nature Neuroscience*, *23*(8), 1016–1024. https://doi.org/10.1038/s41593-020-0644-4

Barth, A., & Schneider, D. (2018). Manipulating the focus of attention in working memory: Evidence for a protection of multiple items against perceptual interference. *Psychophysiology*, *55*(7), Article e13062. https://doi.org/10.1111/psyp.13062

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

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science*, *23*(8), 887–898. https://doi.org/10.1177/0956797612439068

Christophel, T. B., Iamshchinina, P., Yan, C., Allefeld, C., & Haynes, J.-D. (2018). Cortical specialization for attended versus unattended working memory. *Nature Neuroscience*, *21*(4), 494–496. https://doi.org/10.1038/s41593-018-0094-4

Chung, Y. H., Williams, L., Brady, T. F., & Störmer, V. S. (2024). *Limits of verbal labels in cognition: Category labels do not improve visual working memory performance for obfuscated objects*. PsyArXiv. https://doi.org/10.31234/osf.io/4thmd

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

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *39*(1), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743. https://doi.org/10.1038/nn.3689

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

Garside, D. J., Chang, A. L. Y., Selwyn, H. M., & Conway, B. R. (2025). The origin of color categories. *Proceedings of the National Academy of Sciences of the United States of America*, *122*(1), Article e2400273121. https://doi.org/10.1073/pnas.2400273121

Gayet, S., Guggenmos, M., Christophel, T. B., Haynes, J.-D., Paffen, C. L. E., Van der Stigchel, S., & Sterzer, P. (2017). Visual working memory enhances the neural response to matching visual input. *The Journal of Neuroscience*, *37*(28), 6638–6647. https://doi.org/10.1523/JNEUROSCI.3418-16.2017

Gould, I. C., Rushworth, M. F., & Nobre, A. C. (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. *Journal of Neurophysiology*, *105*(3), 1318–1326. https://doi.org/10.1152/jn.00653.2010

Greene, C. M., Kennedy, K., & Soto, D. (2015). Dynamic states in working memory modulate guidance of visual attention: Evidence from an n-back paradigm. *Visual Cognition*, *23*(5), 546–560. https://doi.org/10.1080/13506285.2015.1052033

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

Gunseli, E., van Moorselaar, D., Meeter, M., & Olivers, C. N. L. (2015). The reliability of retro-cues determines the fate of noncued visual working memory representations. *Psychonomic Bulletin & Review*, *22*(5), 1334–1341. https://doi.org/10.3758/s13423-014-0796-x

Hajonides, J. E., van Ede, F., Stokes, M. G., Nobre, A. C., & Myers, N. E. (2023). Multiple and dissociable effects of sensory history on working-memory performance. *The Journal of Neuroscience*, *43*(15), 2730–2740. https://doi.org/10.1523/JNEUROSCI.1200-22.2023

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). Academic Press. https://doi.org/10.1016/S0079-7421(08)60041-9

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

Hu, Y., Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2016). Executive control of stimulus-driven and goal-directed attention in visual working memory. *Attention, Perception, & Psychophysics*, *78*(7), 2164–2175. https://doi.org/10.3758/s13414-016-1106-7

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

Huang, J., & Sekuler, R. (2010). Distortions in recall from visual memory: Two classes of attractors at work. *Journal of Vision*, *10*(2), Article 24. https://doi.org/10.1167/10.2.24

Iamshchinina, P., Christophel, T. B., Gayet, S., & Rademaker, R. L. (2021). Essential considerations for exploring visual working memory storage in the human brain. *Visual Cognition*, *29*(7), 425–436. https://doi.org/10.1080/13506285.2021.1915902

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., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial dependence across perception, attention, and memory. *Trends in Cognitive Sciences*, *21*(7), 493–497. https://doi.org/10.1016/j.tics.2017.04.011

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in psychtoolbox-3? *Perception*, *36*(14), 1–16. https://nyuscholars.nyu.edu/en/publications/whats-new-in-psychtoolbox-3

Kuo, B.-C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience*, *24*(1), 51–60. https://doi.org/10.1162/jocn_a_00087

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

LaRocque, J. J., Riggall, A. C., Emrich, S. M., & Postle, B. R. (2017). Within-category decoding of information in different attentional states in short-term memory. *Cerebral Cortex*, *27*(10), 4881–4890. https://doi.org/10.1093/cercor/bhw283

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

Li, A. Y., Liang, J. C., Lee, A. C. H., & Barense, M. D. (2020). The validated circular shape space: Quantifying the visual similarity of shape. *Journal of Experimental Psychology: General*, *149*(5), 949–966. https://doi.org/10.1037/xge0000693

Liberman, A., Fischer, J., & Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology*, *24*(21), 2569–2574. https://doi.org/10.1016/j.cub.2014.09.025

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., 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

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Lawrence Erlbaum.

Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review*, *14*(6), 1072–1078. https://doi.org/10.3758/BF03193093

Makovski, T., & Pertzov, 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., Lorenc, E. S., & Lewis-Peacock, J. A. (2022). Working memory swap errors have identifiable neural representations. *Journal of Cognitive Neuroscience*, *34*(5), 776–786. https://doi.org/10.1162/jocn_a_01831

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

Manassi, M., Murai, Y., & Whitney, D. (2023). Serial dependence in visual perception: A meta-analysis and review. *Journal of Vision*, *23*(8), Article 18. https://doi.org/10.1167/jov.23.8.18

Millisecond Software LLC. (2016). *Inquisit 6* [Computer software]. https://www.millisecond.com

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869), 1543–1546. https://doi.org/10.1126/science.1150769

Moon, J., & Kwon, O. S. (2022). Attractive and repulsive effects of sensory history concurrently shape visual perception. *BMC Biology*, *20*(1), Article 247. https://doi.org/10.1186/s12915-022-01444-7

Naveh-Benjamin, M., & Cowan, N. (2023). The roles of attention, executive function and knowledge in cognitive ageing of working memory. *Nature Reviews Psychology*, *2*(3), 151–165. https://doi.org/10.1038/s44159-023-00149-0

Nemes, V. A., Parry, N. R. A., 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

Nobre, A. C., & Stokes, M. G. (2019). Premembering experience: A hierarchy of time-scales for proactive attention. *Neuron*, *104*(1), 132–146. https://doi.org/10.1016/j.neuron.2019.08.030

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology*, *43*(1), 25–53. https://doi.org/10.1146/annurev.ps.43.020192.000325

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411–421. https://doi.org/10.1037/0278-7393.28.3.411

Oberauer, K. (2007). Activation, binding, and selective access: An embedded three-component framework for working memory. In N. Osaka, R. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 351–368). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198570394.003.0020

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 45–100). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(09)51002-X

Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current Directions in Psychological Science*, *21*(3), 164–169. https://doi.org/10.1177/0963721412444727

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

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

Olivers, C. N. L., & Roelfsema, P. R. (2020). Attention for action in visual working memory. *Cortex*, *131*, 179–194. https://doi.org/10.1016/j.cortex.2020.07.011

Ort, E., Fahrenfort, J. J., & Olivers, C. N. L. (2018). Lack of free choice reveals the cost of multiple-target search within and across feature dimensions. *Attention, Perception, & Psychophysics*, *80*(8), 1904–1917. https://doi.org/10.3758/s13414-018-1579-7

Ort, E., Fahrenfort, J. J., Ten Cate, T., Eimer, M., & Olivers, C. N. (2019). Humans can efficiently look for but not select multiple visual objects. *eLife*, *8*, Article e49130. https://doi.org/10.7554/eLife.49130

Ort, E., & Olivers, C. N. L. (2020). The capacity of multiple-target search. *Visual Cognition*, *28*(5–8), 330–355. https://doi.org/10.1080/13506285.2020.1772430

Overkott, C., & Souza, A. S. (2022). Verbal descriptions improve visual working memory but have limited impact on visual long-term memory. *Journal of Experimental Psychology: General*, *151*(2), 321–347. https://doi.org/10.1037/xge0001084

Panichello, M. F., & Buschman, T. J. (2021). Shared mechanisms underlie the control of working memory and attention. *Nature*, *592*(7855), 601–605. https://doi.org/10.1038/s41586-021-03390-w

Peich, M.-C., Husain, M., & Bays, P. M. (2013). Age-related decline of precision and binding in visual working memory. *Psychology and Aging*, *28*(3), 729–743. https://doi.org/10.1037/a0033236

Pertzov, 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

Pratte, M. S. (2019). Swap errors in spatial working memory are guesses. *Psychonomic Bulletin & Review*, *26*(3), 958–966. https://doi.org/10.3758/s13423-018-1524-8

Printzlau, F. A. B., Myers, N. E., Manohar, S. G., & Stokes, M. G. (2022). Neural reinstatement tracks spread of attention between object features in working memory. *Journal of Cognitive Neuroscience*, *34*(9), 1681–1701. https://doi.org/10.1162/jocn_a_01879

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

Robinson, M. M., & Brady, T. F. (2023). A quantitative model of ensemble perception as summed activation in feature space. *Nature Human Behaviour*, *7*(10), 1638–1651. https://doi.org/10.1038/s41562-023-01602-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.aah7011

Saito, J. M., Bae, G.-Y., & Fukuda, K. (2024). Judgments during perceptual comparisons predict distinct forms of memory updating. *Journal of Experimental Psychology: General*, *153*(1), 38–55. https://doi.org/10.1037/xge0001469

Saito, J. M., Duncan, K., & Fukuda, K. (2023). Comparing visual memories to similar visual inputs risks lasting memory distortion. *Journal of Experimental Psychology: General*, *152*(8), 2318–2330. https://doi.org/10.1037/xge0001400

Saito, J. M., Kolisnyk, M., & Fukuda, K. (2022). Perceptual comparisons modulate memory biases induced by new visual inputs. *Psychonomic Bulletin & Review*, *30*(1), 291–302. https://doi.org/10.3758/s13423-022-02133-w

Salthouse, T. A. (2016). Continuity of cognitive change across adulthood. *Psychonomic Bulletin & Review*, *23*(3), 932–939. https://doi.org/10.3758/s13423-015-0910-8

Schneegans, S., & Bays, P. M. (2017). Restoration of fMRI decodability does not imply latent working memory states. *Journal of Cognitive Neuroscience*, *29*(12), 1977–1994. https://doi.org/10.1162/jocn_a_01180

Schneider, D., Barth, A., Getzmann, S., & Wascher, E. (2017). On the neural mechanisms underlying the protective function of retroactive cuing against perceptual interference: Evidence by event-related potentials of the EEG. *Biological Psychology*, *124*, 47–56. https://doi.org/10.1016/j.biopsycho.2017.01.006

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. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, *128*, 53–67. https://doi.org/10.1016/j.visres.2016.09.010

Sheehan, T. C., & Serences, J. T. (2022). Attractive serial dependence overcomes repulsive neuronal adaptation. *PLOS Biology*, *20*(9), Article e3001711. https://doi.org/10.1371/journal.pbio.3001711

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317–1323. https://doi.org/10.1126/science.3629243

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

Souza, A. S., Overkott, C., & Matyja, M. (2021). Categorical distinctiveness constrains the labeling benefit in visual working memory. *Journal of Memory and Language*, *119*, Article 104242. https://doi.org/10.1016/j.jml.2021.104242

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: Attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1237–1256. https://doi.org/10.1037/a0036331

Souza, A. S., Rerko, L., & Oberauer, K. (2016). Getting more from visual working memory: Retro-cues enhance retrieval and protect from visual interference. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 890–910. https://doi.org/10.1037/xhp0000192

Spaak, E., Watanabe, K., Funahashi, S., & Stokes, M. G. (2017). Stable and dynamic coding for working memory in primate prefrontal cortex. *The Journal of Neuroscience*, *37*(27), 6503–6516. https://doi.org/10.1523/JNEUROSCI.3364-16.2017

Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring latent visual working memory representations in human cortex. *Neuron*, *91*(3), 694–707. https://doi.org/10.1016/j.neuron.2016.07.006

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., Fougnie, 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

Sutterer, D. W., Foster, J. J., Adam, K. C. S., Vogel, E. K., & Awh, E. (2019). Item-specific delay activity demonstrates concurrent storage of multiple active neural representations in working memory. *PLOS Biology*, *17*(4), Article e3000239. https://doi.org/10.1371/journal.pbio.3000239

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

The MathWorks Inc. (2020). *MATLAB Version: 9.8 (R2020a)*. https://www.mathworks.com

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 of the United States of America*, *109*(22), 8780–8785. https://doi.org/10.1073/pnas.1117465109

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 Ede, F., Board, A. G., & Nobre, A. C. (2020). Goal-directed and stimulus-driven selection of internal representations. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(39), 24590–24598. https://doi.org/10.1073/pnas.2013432117

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., Gunseli, 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

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

Wang, S., & van Ede, F. (2025). Re-focusing visual working memory during expected and unexpected memory tests. *eLife*, *13*, Article RP100532. https://doi.org/10.7554/eLife.100532

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

Williams, J. R., Brady, T. F., & Störmer, V. S. (2022). Guidance of attention by working memory is a matter of representational fidelity. *Journal of Experimental Psychology: Human Perception and Performance*, *48*(3), 202–231. https://doi.org/10.1037/xhp0000985

Williams, J. R., Robinson, M. M., & Brady, T. F. (2023). There is no theory-free measure of “swaps” in visual working memory experiments. *Computational Brain & Behavior*, *6*(2), 159–171. https://doi.org/10.1007/s42113-022-00150-5

Williams, M., & Woodman, G. F. (2012). Directed forgetting and directed remembering in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1206–1220. https://doi.org/10.1037/a0027389

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

Xu, Y. (2020). Revisit once more the sensory storage account of visual working memory. *Visual Cognition*, *28*(5–8), 433–446. https://doi.org/10.1080/13506285.2020.1818659

Yu, Q., Teng, C., & Postle, B. R. (2020). Different states of priority recruit different neural representations in visual working memory. *PLOS Biology*, *18*(6), Article e3000769. https://doi.org/10.1371/journal.pbio.3000769

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

Zhang, Z., & Lewis-Peacock, J. A. (2023a). Prioritization sharpens working memories but does not protect them from distraction. *Journal of Experimental Psychology: General*, *152*(4), 1158–1174. https://doi.org/10.1037/xge0001309

Zhang, Z., & Lewis-Peacock, J. A. (2023b). Bend but don’t break: Prioritization protects working memory from displacement but leaves it vulnerable to distortion from distraction. *Cognition*, *239*, Article 105574. https://doi.org/10.1016/j.cognition.2023.105574

Zoom Video Communications, Inc. (2016). *Zoom* (Version 3.5) [Software]. https://zoom.us

Received October 22, 2024
Revision received May 4, 2025
Accepted May 28, 2025 ■