# Visual Working Memory prioritization modulates Serial Dependence beyond simple attentional effects

**Ekaterina Andriushchenko**

ekaterina.andriushchenko@phd.unipd.it

University of Padua

**Andrey Chetverikov**

University of Bergen

**Gianluca Campana**

University of Padua

**Research Article**

**Keywords:** serial dependence, visual working memory, visual perception, contextual bias

**Posted Date:** May 6th, 2025

**DOI:** https://doi.org/10.21203/rs.3.rs-6511490/v1



**Additional Declarations:** No competing interests reported.

# Visual Working Memory prioritization modulates Serial Dependence beyond simple attentional effects

Ekaterina Andriushchenko [1,*], Andrey Chetverikov[2,†], Gianluca Campana[1,3,†]
[1]Department of General Psychology, University of Padova, Via Venezia 8, Padova, 35131, Italy.
[2]Department of Psychosocial Science, Faculty of Psychology, University of Bergen, Christies gate 12, Bergen, 5015, Norway.
[3]Human Inspired Technology Research Centre, University of Padova, Via Venezia 8, Padova, 35131, Italy.

* Corresponding author: Ekaterina Andriushchenko
E-mail: ekaterina.andriushchenko@phd.unipd.it;
Contributing authors: Andrey.Chetverikov@uib.no; gianluca.campana@unipd.it;
† AC and GC contributed equally to this manuscript

**Abstract**

**Background:** Serial dependence (SD) is a contextual bias in visual processing, where current perception is influenced by past stimuli. This study explores how prioritization in visual working memory (VWM) modulates SD through three experiments.

**Results:** Experiment 1 revealed that tasks requiring active memory maintenance (thus prioritization in VWM) amplified SD, with stronger biases observed when participants retained prior stimuli for extended periods. Conversely, Experiments 2 and 3, which employed pre- and post-cueing in a dual-stimuli setup, found no significant differences in SD strength between congruent and incongruent conditions, suggesting that prioritization alone does not influence SD magnitude.

**Conclusions:** The results highlight the nuanced interplay between memory maintenance, attention, and perceptual biases, suggesting that SD arises from complex interactions beyond simple attentional mechanisms. This study advances the understanding of SD within perceptual decision-making, underscoring the roles of memory prioritization and maintenance in shaping visual judgments.

**Keywords:** serial dependence, visual working memory, visual perception, contextual bias

## Background

Vision is often considered a direct reflection of the world around us. In reality, it is inherently shaped by contextual information that helps to transform a constantly fluctuating stream of stimuli into a stable and coherent visual experience (Cicchini et al., 2024; Manassi & Whitney, 2024; van Bergen & Jehee, 2019). At the same time, context might bias perception when the information it provides is irrelevant. This paper focuses on serial dependence — a pervasive contextual bias in visual information processing — and explores how prioritization in memory modulates this phenomenon.

Serial dependence (SD) refers to the tendency of perception of a current stimulus to be biased toward previously encountered information (Cicchini et al., 2024; Manassi et al., 2023; Pascucci et al., 2023). While initially regarded as purely a perceptual phenomenon (Fischer & Whitney, 2014, but see also Fornaciai & Park, 2018; St. John-Saaltink et al., 2016), subsequent studies have revealed that it can also be affected by post-perceptual processes, including visual working memory (VWM) and decision-making (Barbosa et al., 2020; Bliss et al., 2017; Ceylan et al., 2021; Fischer et al., 2024; Hajonides et al., 2023). For example, although SD commonly manifests as attraction, where current perception is drawn toward past information to create a stabilizing effect (Fischer & Whitney, 2014; Kiyonaga et al., 2017), the direction of SD can be either attractive or repulsive depending on how the stimulus is maintained in VWM (Chen & Bae, 2024). Numerous studies have highlighted the role of visual VWM in modulating the strength of SD (see Pascucci et al., 2023, for a review) as well as the impact of memory reactivations in enhancing biases (Barbosa et al., 2020; Barbosa & Compte, 2020; Fischer et al., 2024). Recent research has further deepened this understanding by uncovering the direct neural signature of SD in VWM and emphasizing the role of later processing stages in VWM representations (Fischer et al., 2024; Shan et al., 2024). While these findings underscore the pivotal role of VWM maintenance in SD, the precise mechanism through which VWM influences this bias remains unclear and warrants further investigation.

VWM studies showed that prioritization of one item among those held in memory can affect both how precisely it is remembered and the extent to which it biases the perception of new stimuli (Gayet et al., 2017; LaRocque et al., 2016; Mei et al., 2019; Myers et al., 2017; Saito et al., 2024; Wang et al., 2025; Zhang & Lewis-Peacock, 2023a, 2023b). When a cue predicts the location of the probed stimulus, error variance is reduced, showing an increased precision of VWM representations (Bays, 2014; Jakob & Gershman, 2023; Yoo et al., 2018; Zhang & Lewis-Peacock, 2023a, 2023b). Further support for the heightened precision of prioritized representations comes from decoding studies demonstrating that prioritized memories are more stable and accessible than unprioritized ones (LaRocque et al., 2016; Rose et al., 2016; Saito et al., 2024; Sprague et al., 2016; Wolff et al., 2017). This flexible mechanism allows memories to be prioritized without necessarily compromising other stored items (Myers et al., 2018; Wang et al., 2025).

Prioritization in VWM also affects the magnitude of biases observed in the observers' responses. Some studies suggested that prioritization can reduce the likelihood of catastrophic memory loss (fewer "swap errors") but amplify attraction effects toward the distractors (Zhang & Lewis-Peacock, 2023a, 2023b). Others have found that prioritization in VWM can decrease the attraction towards distractors (Saito et al., 2024). The conflicting findings highlight the need for further investigation to clarify how prioritization influences the direction of VWM effects on biases.

When it comes to SD, researchers manipulated the priority of to-be-reported stimuli by cueing their locations or other features. For example, Fischer & Whitney (2014) found that when nine stimuli appear simultaneously, only the cued one affects the perceived orientation in the subsequent trial. This effect was echoed by Fritsche & De Lange (2019) and Fischer et al. (2020), who observed that attention to previous stimulus features significantly increased attraction bias. Similarly, Czoschke et al. (2019) and Fischer et al. (2020) found that when participants encoded the motion direction of two sequentially presented random dot stimuli, SD occurred between the second one on trial *N* and the first one on trial *N*+1 only when the former was cued for reporting. Later, Fischer et al. (2024) confirmed a serial dependence effect specifically toward the motion direction of the previously cued target but not the non-cued one. Additionally, they demonstrated that the direction of the cued target could be successfully reconstructed from MEG data, whereas the non-cued target could not be reconstructed during the retro-cue phase. Hajonides et al. (2023) obtained the same result for orientation also using MEG. All these results suggest that attentional prioritization should be taken into account in explaining the SD mechanism.

***Theoretical perspectives on the role of prioritization in SD***

While SD has led to a large amount of empirical work, its theoretical understanding remains underdeveloped. Most of the proposed models remain at descriptive or mechanistic levels (in Marr's, 1983, classification), limiting their capability to explain how prioritization can affect SD. However, two normative computational models, the Bayesian model (Cicchini et al., 2018; van Bergen & Jehee, 2019) and the Demixing Model (Chetverikov, 2023a), provide predictions for serial biases that depend on uncertainty or noise. As prioritization can manipulate the amount of internal noise, it can provide a valuable framework to test these models' predictions.

According to the Bayesian model, perception is a hypothesis about the external world shaped by sensory input (Vincent, 2015). However, as mentioned earlier, sensory information is often uncertain, prompting the brain to integrate prior and present inputs to enhance behavioral precision (van Bergen & Jehee, 2019). This strategy is adaptive in natural environments where the visual inputs are relatively stable across time. However, in the experimental conditions when previous and current trials are unrelated, it leads to biases. Regarding prioritization, the Bayesian model proposes that the brain puts more weight on more reliable information. So, when a current stimulus is less uncertain compared to the previous one, the magnitude of serial dependence decreases, and vice versa. This approach allows observers to make more accurate decisions by prioritizing information that is more certain, enhancing decision accuracy (van Bergen & Jehee, 2019). However, while standard Bayesian models explain the attractive biases in SD by integrating noisy sensory measurements of current stimuli, they fail to account for repulsive biases effectively (Fritsche et al., 2020). Extensions of Bayesian models — such as those incorporating efficient encoding and Bayesian decoding — have been proposed to capture both short-term attraction and longer-term repulsion patterns more accurately (Fritsche et al., 2020). Nevertheless, these models fail to fully account for SD variability across different contexts, suggesting that the complexity of real-world perception introduces additional, sometimes contradictory, mechanisms.

The Demixing Model (Chetverikov, 2023a) suggests that SD, like other contextual biases, results from the observer's attempt to separate neural signals related to different stimuli in the environment. In the case of SD, these are the signals related to the current and the previous items. Whether the bias is attractive or repulsive depends on factors such as item similarity, the level of sensory noise, and the observer's assumptions about the environment. Prioritization is assumed to decrease sensory noise, which, according to the model, should decrease the bias magnitude from previously encountered stimuli. However, Chetverikov (2023a) noted that in the case of biases between sequentially presented items, the pattern might be more complex. This is because the observer has already encoded and reported the previous item when they encounter the current one. Accordingly, the initial noise level for the previous item might differ from the one at the time when the current item is perceived. As an illustration, a perception of a noisy low-contrast Gabor patch might be transformed into a representation of a single line, either real (e.g., a response bar) or imaginary, that might have a lower noise level. Nevertheless, we reasoned it would be interesting to explore the model's predictions for serial dependence to see how they compare with the predictions of a Bayesian observer.

In sum, understanding SD within the framework of VWM — especially with attention to noise parameters and prioritization — could be essential for developing comprehensive models of perceptual processing.

## Computational models

### *Bayesian observer*

***Model***. To illustrate the Bayesian model predictions, we simulated the behavior of a Bayesian observer as described by van Bergen & Jehee (2019).

The model starts with the assumption that in each trial, the observer obtains a sensory measurement ($x$) of stimuli ($s$) corrupted by noise:

$$p(x|s) = f_{WN}(x; s, \sigma^2) \quad (1)$$

Here, we use wrapped normal noise distribution to account for circularity in the orientation space.

The observer also assumes that the stimuli on consecutive trials are related to each other following the statistics of the natural environment:

$$p(s_t; s_{t-1}) = \ p_{same} C(s_t; s_{t-1}, \sigma_s, \gamma) + (1 - p_{same}) U(0,2\pi) \quad (2)$$

where $s_t$ is the stimulus on the current trial, $s_{t-1}$ is the stimulus on the previous trial, $p_{same}$ is the probability that there was no abrupt change in the environment, and $U(0,2\pi)$ is the circular uniform distribution. The function $C$ describes the probability of stimulus changing between the trials in the absence of abrupt changes:

$$C(s_t; s_{t-1}, \sigma_s, \gamma) = \frac{1}{Z} \exp\left(-\frac{1}{2\sigma_s^2} \left|\text{angle}(s_t, s_{t-1})\right|^{\gamma}\right) \quad (3)$$

where $Z$ is the normalization constant, $angle(s_t, s_{t-1})$ is the angular difference between the stimuli, and $\gamma$ determines the steepness of the function.

The observer then inverts this generative model (Eqs. 1-2) to infer the most likely stimulus in a given trial using the Bayes rule. In other words, the observer combines information about the current and the previous stimuli based on the measurements obtained in the two trials and the assumed relationship between the stimuli:

$$p(s_t|x) \propto p(x|s_t)p(s_t|s_{t-1}) \quad (4)$$

This final distribution represents the observer's belief about the orientation of the current stimulus based on the sensory observations about the current and the previous stimuli. The observer then uses the mean of the posterior distribution as a response.

***Simulations.*** To simulate the observer behavior, we first randomly picked the stimuli for $10^6$ trials. By adding the noise to each stimulus (with the high or low level of noise assigned randomly), we created a vector of sensory observations across trials. For each sensory observation, we then computed the likelihood using a wrapped normal distribution function, with the mean based on the observation and the variability determined by the trial noise. In alignment with our first experiment, the amount of noise ($\sigma$) is modulated by prioritization: when a target was cued, it represented a low noise condition, and when a target was non-cued, it represented a high noise condition (6 and 9 degrees, respectively, converted to radians). Only two levels of noise were used because the predictions of the Bayesian observer model are relatively straightforward and have been described before (e.g., Cicchini et al., 2018; van Bergen & Jehee, 2019). To calculate the prior distribution, we used a uniform prior for the first observation, and for subsequent observations, we used the previous posterior convolved with a transition kernel. For our simulations, we used the values of $\sigma_s = 16.9$, $\gamma = 2.6$, and $p_{same} = 0.64$ based on van Bergen & Jehee (2019). We evaluated errors in the observer’s responses by comparing these estimates to the actual orientation of the stimulus for each observation and computed the bias by multiplying errors by the sign of the distance to the previous target on each trial.

Finally, we followed the procedure used for the actual data (see Methods) to estimate SD across trials (see Figure 1). We plotted bias against orientation differences between consecutive stimuli to illustrate how SD was affected by the noise level of the previous target. The model predicted an attractive bias, with a stronger SD when the previous trial was cued (indicating higher internal noise for the current item's representation) and a weaker SD when the current trial was cued (indicating lower internal noise for the current item's representation).

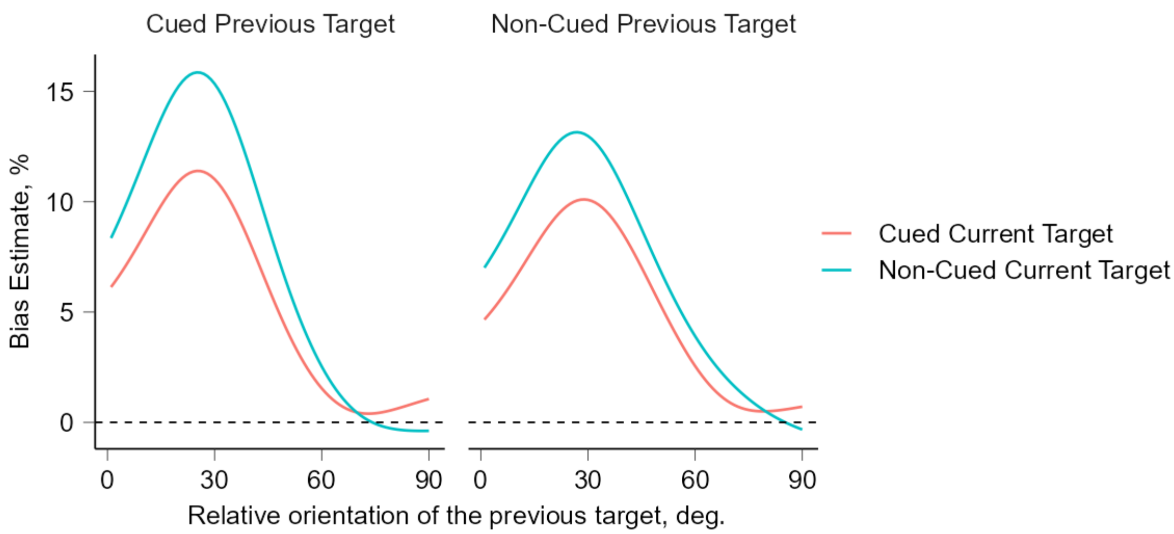


**Figure 1.** *A Bayesian Observer Model for Serial Dependence effect with two levels of cueing for the current and the previous target.* Biases in orientation estimates (in degrees) for responses as a function of dissimilarity between current and previous stimuli. The two panels represent conditions where the previous target was cued (left) or non-cued (right). The red and blue lines correspond to conditions where the current target was cued or non-cued, respectively. Positive values indicate an attractive bias.

### *The Demixing Model (DM)*

***Model.*** We derived DM predictions following the approach described by Chetverikov (2023a). DM assumes that in each trial, the observer obtains multiple sensory measurements *X*. Each sensory measurement comes from one of two sources (or components), with probabilities $\pi_1$ and $\pi_2 = 1 - \pi_1$. One component here represents a recent stimulus that the participant has to remember, while the other component represents a previous stimulus that is no longer relevant but may still influence perception.

These measurements capture information across two key perceptual dimensions. The first dimension here represents orientation. Due to its circularity, we modeled orientation using a wrapped normal distribution. The second dimension represents time or other features that help identify which stimulus to respond to. Unlike orientation, this dimension here is not circular and follows a normal distribution.

Each of these two components has its own characteristics: a mean value ($\mu_{1,j}$, $\mu_{2,j}$) and standard deviation ($\sigma_{1,j}$, $\sigma_{2,j}$) for both the orientation and temporal dimensions, where $j \in \{1,2\}$ indicates the component. The mean corresponds to the true stimuli parameters (e.g., an orientation of 45° and a specific time point), while the standard deviation quantifies the noise in neural processing due to various factors.

When modeling the probability of observing a particular measurement $x_i = \left(x_{i,1}, x_{i,2}\right)$, the probabilities from both components are combined:

$$p(x_i|\theta) = \sum_{j=1}^{2} \pi_j \left[f_{WN}\left(x_{i,1}; \mu_{1,j}, \sigma_{1,j}^2\right) \cdot f_N\left(x_{i,2}; \mu_{2,j}, \sigma_{2,j}^2\right)\right] \tag{5}$$

where $\theta = \{\pi, \mu, \sigma\}$ is the set of all parameters, $f_{WN}$ is the wrapped normal distribution (for orientation), and $f_N$ is the standard normal distribution (for the temporal dimension). The equation assumes independence between the orientation and temporal dimensions within each component, allowing the joint probability to be expressed as the product of the individual dimension probabilities.

Based on these sensory measurements, the observer determines the most likely values for the means and standard deviations of the two components through maximum likelihood estimation:

$$\theta = \text{argmax}\, L(\theta; X) \tag{6}$$

The behavioral response is then determined by selecting the estimated orientation mean ($\hat{\mu}_{1,j}$) from the component with a higher value on the temporal dimension (larger $\hat{\mu}_{2,j}$). This selection process models how observers identify the most recent or temporally relevant stimulus when making their response.

***Simulations.*** In the simulations, noise levels ($\sigma$) were manipulated to represent different prioritization conditions. When a stimulus is cued as important (high priority), it is modeled with lower noise, reflecting more precise encoding. When a stimulus is not cued (low priority), it is modeled with higher noise, reflecting less precise encoding. At the same time, we assumed that the previously shown item always has higher noise levels than the current one.

In the case of the Bayesian model, predictions related to noise have been previously described in the literature and are relatively straightforward, making the number of noise levels a less critical factor. Therefore, we used only two levels of noise, which was consistent with our experiment. In contrast, predictions from the Demixing model are more complex and have not yet been described. To explore these predictions in more detail, we introduced additional noise levels, allowing for a more comprehensive examination of the model's behavior.

Therefore, based on the preliminary exploration of the parameter space, different levels of noise in orientation perception were examined by testing five different standard deviation values ($\sigma^2_{1,1}$ ranging from 12° to 28° in 4° steps) for the first component (representing the stimulus in the current trial) and three different levels ($\sigma^2_{1,2} \in \{40°, 60°, 80°\}$) for the second component (representing the previous stimulus). These values were selected so that behavioral variability lies in the same range as the behavioral variability of the real observers and the overall direction of biases for the low-noise item remains positive, corresponding to the typical pattern of serial dependence effects. The temporal standard deviation $\sigma_2$ was assumed to be equal for both components and fixed ($\sigma_2 = \sigma_{2,1} = \sigma_{2,2}$ =20) and the discriminability in the temporal dimension was fixed as well ($d'_{temp} = \frac{\mu_{2,1} - \mu_{2,2}}{\sigma_2} = 1$). The number of observations was fixed at $N = 100$ measurements in our simulations and the probabilities of signals being caused by each component were assumed equal ($\pi_1 = \pi_2 = 0.5$). These parameters represent the 'average' case considered by Chetverikov (2023a) and do not cover the full space of potential model behavior. Our explorative analysis of other parameters suggests that they do not affect the direction of the previous item's noise effect, while the current noise effect can become positive-only or negative-only (with respect to the changes in bias) in addition to the inverted U-shaped pattern described here. Finally, the orientation difference between the two components was systematically varied across 120 steps from 0° to 90°, representing conditions where the two orientations range from identical (0° difference) to maximally different (90° difference).

For each combination of parameters, we simulated 10000 trials. For each trial, sensory measurements were generated according to the true model, then the Expectation-Maximization algorithm was employed to obtain the maximum likelihood estimate of the model parameters. To ensure convergence to the global optimum, 50 different random initializations were used for each simulation.

Finally, to generate predictions for Experiment 1, we estimated the SD across trials using the same procedure applied to the actual data (Figure 2). For clarity, in the left plot, current stimulus noise ($\sigma_{1,1}^2$) is fixed at 24°, and in the right plot previous stimulus noise ($\sigma_{1,2}^2$) is fixed at 60°. The main finding reveals a linear increase in SD as the noise of the previous item decreases (indicating the higher internal noise for the current item). In contrast, varying the current item noise produces a non-linear, inverted U-shaped dependence. For example, the strongest bias in Figure 2 (right panel) is observed for the intermediate level of noise in the current item ($\sigma_{1,1}^2 = 24$).

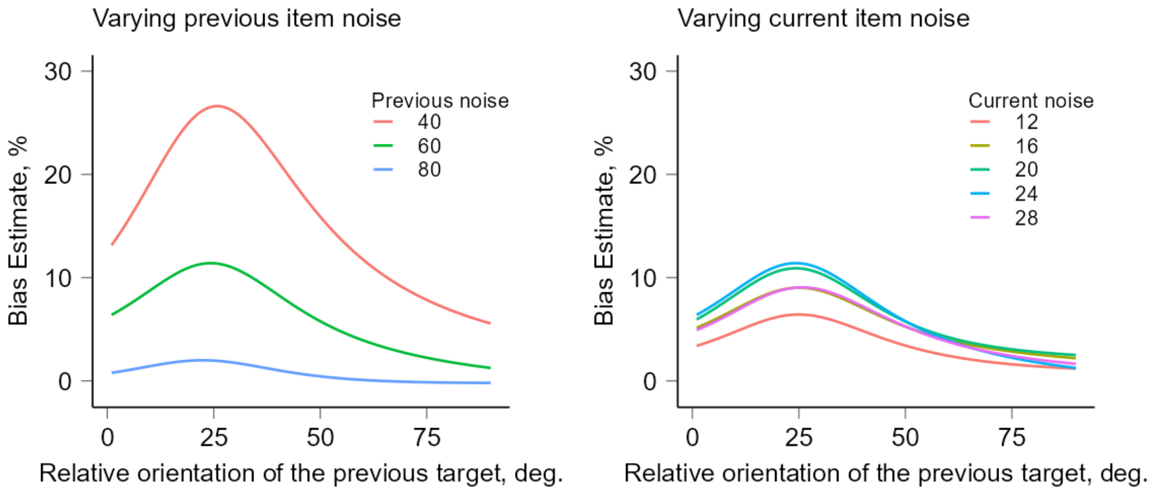


**Figure 2.** *A Demixing Model for Serial Dependence effect with varying levels of noise for the current and the previous stimuli.* The left plot shows the effect of varying previous stimulus noise while keeping the current stimulus noise fixed at 24°. The right plot shows the effect of varying current stimulus noise while keeping the previous stimulus noise fixed at 60°. Biases in orientation estimates (in degrees) for responses as a function of dissimilarity between current and previous stimuli. Positive values indicate an attractive bias.

## The current study

We explored how prioritization in memory influences SD in three experiments. In Experiment 1, we manipulated prioritization in a standard delayed report task with a single-oriented Gabor stimulus in each trial. To this end, in some trials, we presented a cue before the stimulus (‘precue’), indicating to participants that they would need to additionally report the stimulus the second time after the next trial. We hypothesized that it would increase the priority of this stimulus during the retention interval, highlighting the critical role of the VWM maintenance component. In Experiments 2 and 3, we further examined the effect of prioritization through pre- or post-cueing of one of two stimuli with a cue probabilistically indicating the item participants would have to report. Similarly, we expected that cued items should be less affected by SD from previous trials and, in turn, induce stronger SD in the following trials, as predicted by the Bayesian Observer Model. Previewing the main results, we found that in Experiment 1 additional requirement to hold information in memory for a longer period of time increased SD strength. In contrast, in Experiments 2 and 3, we found no significant differences between congruent and incongruent conditions, indicating that manipulating uncertainty through pre- or post-cueing for one of two simultaneous stimuli did not affect SD. This is despite the clear evidence of prioritization in the form of reduced error variability for prioritized items in all three studies. Our findings suggest that active memory maintenance can amplify perceptual biases beyond mere prioritization effects. This opens pathways for refining models of SD in perceptual judgment tasks.

# Results

### *Experiment 1*

**Data preprocessing.** Participants' responses were preprocessed to remove idiosyncratic orientation-dependent bias (i.e., individual variations in perception based on stimulus orientation) from reports and to identify and remove outliers for each participant using the *remove_cardinal_biases* function in the *circhelp* package in R (Chetverikov, 2023b).

We estimated biases relative to previously seen items by calculating the asymmetry in response probability density. This measure indicates that it was much more likely that participants made an error in the direction of the previous item than in the direction away from it for each angular distance between items. The *density_asymmetry* (*circhelp* package in R; Chetverikov, 2023b) function was utilized to generate a smoothed estimate of the asymmetry in error probability as a function of dissimilarity (angular difference) between the current and the previous stimuli. When estimating probabilities for each dissimilarity step, we considered not only trials with the same difference but also all trials to capture a more continuous and nuanced probability distribution of errors. However, trials that were closer in difference to the current one were given higher weight in the asymmetry estimate. Analyzing the asymmetry in probability allowed us to identify clearer patterns in participants' responses than relying solely on mean bias values.

The complete preprocessing and analysis code is available on the Open Science Framework (OSF) and can be accessed via the following link: https://osf.io/wunf8/?view_only=079afc6a31904559b93df4fa4debfc0c

**Overall performance.** We first tested the effectiveness of the pre-cue manipulation, which involved holding a stimulus in memory for cued stimuli (see Figure 3A). As expected, participants had smaller errors for the immediate report (Response 1 to Stimulus 1) of cued compared to non-cued stimuli ($M$ = 8.59, $SD$ = 7.82, vs. M = 9.24, SD = 8.89, $t(17)$ = 4.03, $p$ < .001). In contrast, Response 2 to Stimulus 2 was slightly negatively affected when observers had to hold another item in memory, with larger errors for cued compared to non-cued streaks ($M$ = 9.24, SD = 8.61, vs. $M$ = 9.08, $SD$ = 8.64, $t(17)$ = -0.93, $p$ < .001). The delayed Response 3 to Stimulus 1 resulted in larger errors compared to Response 1 and Response 2 ($M$ = 16.88, $SD$ = 19.73; $p$ < .001 for both comparisons).

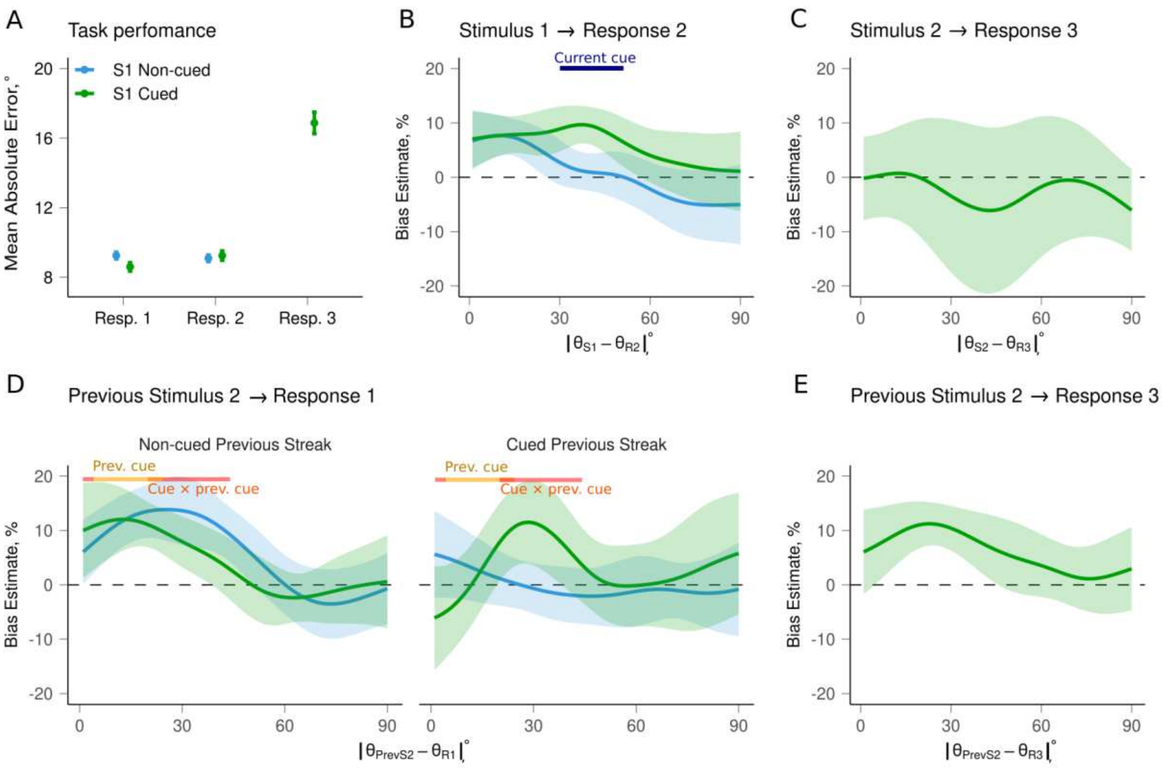


**Figure 3.** *Task Performance and Serial Dependence Plots.* **A.** Mean absolute error in orientation estimates in each condition. Bars show 95% confidence intervals (CI). **B-E.** Biases in orientation estimates for Response 1-3 as a function of dissimilarity (angular difference)

between the current and previous stimuli. Positive values correspond to attractive bias (serial dependence). Shaded regions show 95% CI. The horizontal segments above the lines indicate the dissimilarity range where the effects of conditions (labeled near the lines) were significant.

**Serial dependence effect within a streak.** We first looked at the average SD magnitude (i.e., the asymmetry in response probability density, see Data Preprocessing) for Response 2. The results showed that across all trials, Response 2 was attracted by Stimulus 1 only when the latter was cued ($M = 0.06$, $SD = 0.13$, $t(17) = 2.63$, $p = .012$) but not when it was non-cued ($M = 0.01$, $SD = 0.09$, $t(17) = 0.78$, $p = .438$), based on mean bias calculations.

We then analyzed the SD as a function of the angular difference between the stimuli. For each angular difference step, we computed first separate one-way *t*-tests for each condition to test if biases are present, and then a one-way ANOVA to test for the difference between conditions. This more detailed analysis by angular distance revealed a significant attractive bias in cued streaks at distance 3 – 50 degrees ($t(17) >= 2.03$, $p < .05$), as well as in non-cued streaks at distances of 1 – 25 degrees ($t(17) >= 2.19$, $p < .05$), along with a repulsive bias repulsive bias at distances of 69 – 82 degrees ($t(17) >= -2.04$, $p < .05$). This combination of opposing biases explains the null results with the non-cued streaks when the average bias is considered. In essence, in the report-and-hold-in-memory condition, the report (and hence the representation) of stimulus orientation in a given streak is biased (attractively) by the orientation of the previous stimulus when past and present stimuli are moderately similar (the strongest effect occurs at around a 35-degree difference between the two stimuli). A repeated-measures ANOVA showed significant differences between cued and non-cued streaks when the differences between the two stimuli were in the 30 – 51 degrees range ($F(1,17) >= 4.58$, $p < .05$) (see Figure 3B). Participants exhibited stronger SD, extending across larger angle differences, when required to retain previous stimuli in memory.

The results did not show any dependence between Stimulus 2 and Response 3, the representation of Stimulus 1 held in memory and reported the second time ($M = -0.02$, $SD =$

0.21, $t(17) = -0.7$, $p = .485$) (refer to Figure 3C). This suggests that stimuli with a previously given response may be less vulnerable to external visual interference.

**Serial dependence effect across different streaks.** We then analyzed the interactions across different streaks. In general, Response 1 to Stimulus 1 of the current streak was similarly attracted towards Stimulus 2 of the previous streak, regardless of whether the current Stimulus 1 was cued or not (cued: $M = 0.04$, $SD = 0.09$, vs. non-cued: $M = 0.04$, $SD = 0.09$, $t(17) = 2.56$, $p = .015$; Figure 3D). However, when considering the dissimilarity between the stimuli, the influence of cueing in the previous streak was evident with attractions of Response 1 of the current streak in the 4 – 24 degree range ($F(1,17) >= 4.68$, $p < .05$), whereas interactions between previous and current streak cueing were observed in the 1 – 4 degree and 20 – 44 degree ranges ($F(1,17) >= 4.55$, $p < .05$). Figure 3D shows a weaker bias when the current target was cued but there was no cue in the previous streak. This supports the idea that prioritization improves the resilience of the representation of the current target. However, when the previous streak was cued, the bias from Stimulus 2 was absent for non-cued current Stimulus 1, likely due to the presence of additional interfering representations (Stimulus 1) and responses (Response 3) from the previous streak, which may interfere with the representation of Stimulus 2. Interestingly, SD was observed only for cued current streaks and not for non-cued ones, which was unexpected. It was anticipated that increased prioritization of the current trial would reduce the bias, yet this result suggests otherwise in case of additional interventions. Overall, this implies that bias from the previous streak can persist, even when current stimuli are prioritized.

SD was observed not only in the immediate response to Stimulus 1 (that is, Response 1) but also in the delayed response to Stimulus 1 (that is, Response 3) ($M = 0.06$, $SD = 0.09$, $t(17) = 3.75$, $p < .001$) (Figure 3E). In other words, Stimulus 2 from the previous streak could influence the representation of subsequent Stimulus 1, and this influence persisted longer (up to Response 3).

To test whether the effects of prioritization on the magnitude of SD were due to increased attention to the cued stimulus, a second experiment was conducted using pre-cueing with two items presented simultaneously, with no delayed responses required.

***Experiment 2***

**Data preprocessing.** The analysis for Experiment 2 followed the same steps as in Experiment 1. For this experiment, we created one variable with three levels for previous and current congruence: incongruent-congruent, congruent-incongruent, and congruent-congruent. Incongruent-incongruent pairs of trials were not analyzed as they constituted a small portion of trials (around 45 for each participant).

**Overall performance.** A comparison of mean errors across conditions confirmed the effectiveness of the congruence manipulation between the pre-cue and response location, with better performance in the congruent condition during the current trial (see Figure 4A). Specifically, participants exhibited significantly smaller response errors in the congruent condition ($M$ = 11.27, $SD$ = 12.44) compared to the incongruent condition ($M$ = 15.22, $SD$ = 17.43), as indicated by the effect of current congruency ($F(1,68) = 196.7$, $p < .001$). However, no significant differences were found for previous congruency ($F(1,68) = 0.13$, $p = .713$) or their interaction ($F(1,68) = 1.89$, $p = .168$).

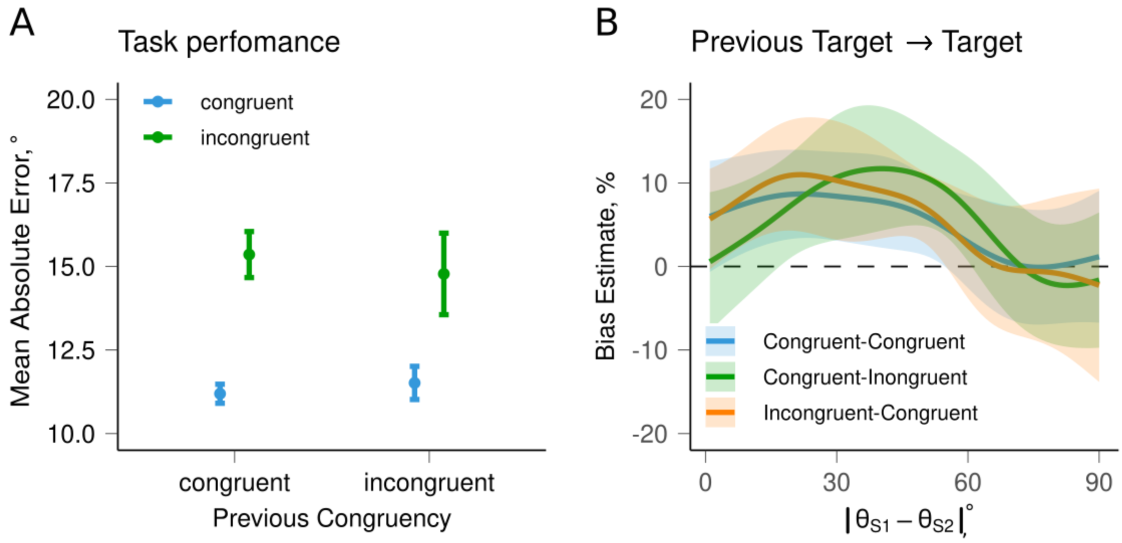


**Figure 4.** *Task Performance and Serial Dependence Plots.* **A.** The task performance plot displays the means and their corresponding confidence intervals for each condition. **B.** The bias from the Previous Stimulus to the Current Stimulus, illustrates the degree to which participants' responses deviate systematically from the target. The absence of the yellow area indicates that there are no significant differences between conditions.

**Serial dependence effect.** The analysis revealed a classic SD effect for two sequential stimuli across conditions, with participants consistently reporting the current target as being attracted to the previous target ($M = 0.05$, $SD = 0.11$, $t(17) = 4.96$, $p < .001$). We then analyzed SD in reports based on the congruency of the current and previous trials, calculating the difference between conditions for each 1 degree step of similarity (angular difference) between the current and previous targets. A one-way ANOVA showed no significant differences in the strength of SD between congruent and incongruent conditions, regardless of the dissimilarity between the current and previous targets (see Figure 4B). These results suggest that prioritization by means of increased attention to the item before encoding did not influence the magnitude of SD. To ensure the robustness of our results, we performed a Bayesian ANOVA on the average bias computed for each participant across dissimilarity levels. The Bayes factors ($BF_{10} = 0.15 \pm 1.28\%$) provided substantial evidence in favor of the baseline model (bias predicted by the random effect of participant only) over models including congruency and previous congruency, suggesting that these factors did not significantly impact bias, with the null hypothesis being more likely than the alternative.

In an exploratory analysis, we also estimated the location of the maximum bias point (a 'peak') for SD curves in the orientation dissimilarity space for each participant in each condition. A Bayesian ANOVA indicated that the peak locations were not affected by the condition ($BF_{10} = 0.51 \pm 0.62\%$).

To examine if and how prioritization by means of increased attention to encoded items in the early phase of VWM maintenance influences the magnitude of SD under post-cueing conditions compared to pre-cueing, we conducted Experiment 3 with two items presented simultaneously as in Experiment 2.

### *Experiment 3*

**Data preprocessing.** The data analysis for Experiment 3 was identical to Experiment 2.

**Overall performance.** Performance was better in the congruent condition, demonstrating the effectiveness of congruency manipulation (see Figure 5A). The performance was better in the congruent condition ($M$ = 15.19, $SD$ = 16.15) compared to the incongruent condition ($M$ = 17.37, $SD$ = 18.15), as shown by the analysis of current congruency ($F(1,68) = 41.53$, $p < .001$). There were no significant differences found for previous congruency ($F(1,68) = 1.76$, $p = .184$) or their interaction ($F(1,68) = 0.001$, $p = .972$).

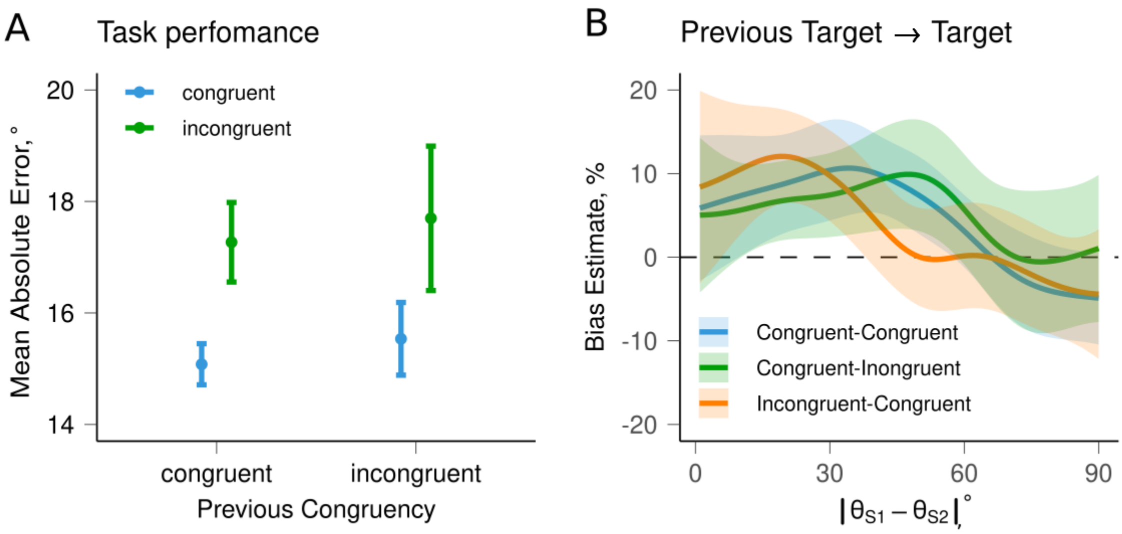


**Figure 5.** *Task Performance and Serial Dependence Plots.* **A.** The task performance plot displays the means and their corresponding confidence intervals for each condition. **B.** The bias from the Previous Stimulus to the Current Stimulus, shows the degree to which participants' responses deviate systematically from the target. The yellow area represents significant differences between conditions.

**Serial dependence effect.** The analysis revealed a SD for two sequential stimuli across conditions ($M$ = 0.04, $SD$ = 0.11, $t(17) = 4.33$, $p < .001$). A two-way repeated measures ANOVA at each step of angular difference between the current and previous targets, did not show significant differences between conditions (Figure 5B). Overall, these results support the previous suggestion that attention to the previously encoded item does not significantly influence the strength of SD. The results of the Bayesian ANOVA on the average bias computed for each participant across dissimilarity levels also supported the baseline model ($BF_{10}$ = 0.16 ± 0.61%) over models that include the effects of congruency and previous congruency, indicating that the null hypothesis is more likely than the alternative. In an exploratory analysis of "peak" locations for SD curves within the orientation dissimilarity space for each participant and condition, Bayesian ANOVA revealed that these peak locations were not influenced by condition ($BF_{10}$ = 0.34 ± 0.77%).

## Discussion

We conducted three experiments to investigate how the prioritization in VWM content affects the magnitude of serial dependence (SD). In Experiment 1, we manipulated prioritization by instructing participants to make extra effort to hold information in memory as it would be needed for a later report. This manipulation increased SD towards memorized stimulus: active memory maintenance of the previous target led to a stronger bias in the current response. In Experiment 2 and 3 we manipulated attentional prioritization through pre- or post-cueing. Despite clear changes in memory fidelity, we found no significant differences in bias strength between congruent and incongruent conditions, suggesting that attentional prioritization through pre- or post-cueing in a two-stimuli setup does not impact SD.

### Does prioritization affect serial dependence strength?

We found that prioritization via an instruction to keep an item in VWM, but not via simple attentional cueing, leads to stronger SD. The lack of effect from prioritization through attention diverges from the findings of Fischer & Whitney (2014) and others (Fritsche & De Lange, 2019; Kim et al., 2020; see Manassi et al., 2023, for a review), who reported a significant impact of attention on SD. This divergence in results indicates that while prioritization may modulate perception, its influence on SD might depend on the nature of prioritization itself. Specifically, whereas active memory maintenance strengthened bias in Experiment 1, the results from Experiment 2 and 3 suggest that attentional prioritization alone does not fully account for this effect. Instead, additional factors related to VWM maintenance likely contribute to the strength of SD, highlighting a more complex interaction between prioritization and memory processes.

However, our findings also suggest that prioritization does not consistently protect against SD, as SD also occurred when the current item was prioritized with an additional memory requirement (Experiment 1) or with an attentional cue (Experiment 2 and 3). In fact, results of Experiment 1 showed significant interactions between the cueing of previous and current streaks for the bias from Stimulus 2 of the previous streak on Response 1 of the current streak. Specifically, SD was observed in three out of four conditions, with the bias absent for non-cued current trials when the previous streak was cued. While this could be due to the presence of additional interference from a maintained representation of a previous Stimulus 1 and its corresponding report (Response 3), SD still occurred when the current trial was cued, regardless of cueing in the previous streak. The results of Experiment 2 and 3 confirmed this finding, as attentionally cueing of the current item did not reduce the bias from the previous item (irrespective of whether the previous item was cued). In other words, prioritizing the current stimulus does not always protect against bias. This result highlights the interplay of attentional and memory-based mechanisms in SD, suggesting a complex interaction that requires further investigation.

Overall, our findings that active memory maintenance intensifies SD align with prior studies showing that focusing on memorizing past stimuli can enhance attraction effects toward them (Fischer et al., 2020; Fischer & Whitney, 2014). However, we highlight the specific contribution of active memory maintenance, distinguishing it from mere attention, which was not clearly differentiated in previous studies. In contrast, our results diverge from studies proposing that giving more attention to the current stimulus offers better protection of items held in VWM from interference (Barth & Schneider, 2018; van Moorselaar et al., 2015a). Our results are partially consistent with Vergauwe et al. (2023), who found that prioritizing visual memories does not consistently make them more vulnerable or resilient to perceptual interference. We observed partial protection only for information held in memory that had undergone decision-making (no SD from Stimulus 2 to Response 3 in Experiment 1). In contrast, Little & Clifford (2025) found that SD remained unaffected by either decisional or stimulus uncertainty of prior stimuli, including differences in stimulus or noise contrast. This difference suggests that attentional focus or uncertainty alone does not fully account for the observed effects, implying that additional factors influence whether the items in memory are protected from interference.

## Comparison of Results to Model Predictions

The Bayesian model predicted an attractive bias, with a larger SD when the previous trial was cued (indicating higher internal noise for the current item's representation) and a smaller SD when the current trial was cued. This partly aligns with our findings: we did observe a stronger SD when the previous trial was cued (see Figure 3B). However, we did not observe a significantly smaller SD when the current trial was cued, and the previous was non-cued as the model predicted (see Figure 3D for the non-cued previous streak). The key difference between this model and our data is that, in our experiment, only S1 was cued or not, and we had conditions involving interference with R3. In contrast, the model allows for each stimulus to be cued or not.

The Demixing model as well predicted an attractive bias, with a linear increase in SD as the noise of the previous item decreased (indicating higher internal noise for the current trial). In contrast, variations in the current item's noise did not predict this linear relationship, which aligns with our findings from Experiment 1. Similar to the Bayesian model, the smallest bias was predicted when the current trial had the least noise. However, as said before, we did not observe a significantly smaller SD when the current trial was cued.

As a result, while both the Bayesian and Demixing models can account for some aspects of our findings, they do not fully predict the unexpected absence of SD under certain interactions between cueing conditions and noise levels. This highlights the need for refined models that better explain the empirical findings.

## The susceptibility to biases and the format of representations in VWM

Previous studies suggest that VWM prioritization could lead to different representational states: prioritized items are actively coded in VWM, while unprioritized items might rely on a different mechanism that does not require sustained neural activity (Bettencourt & Xu, 2016; Lewis-Peacock et al., 2012; Rose et al., 2016; Wolff et al., 2017; Zhang & Lewis-Peacock, 2023a, 2023b). Consequently, the way information is maintained may shape its vulnerability to cognitive biases, with potentially being more subject to interference, whereas actively coded items might be more resistant ('protection' hypothesis; Makovski & Pertzov, 2015; van Moorselaar et al., 2015b). Alternatively, the active state might lead to higher susceptibility to biases ('vulnerability' hypothesis; Mallett & Lewis-Peacock, 2019). Both of these ideas are not without contest, as some previous results suggest that an item's prioritization in VWM does not affect its susceptibility to distraction (Zhang & Lewis-Peacock, 2023b) and that task-irrelevant (unprioritized) information from the previous trial is not maintained exclusively in an activity-silent manner (Bae & Luck, 2019).

Our results also provide a complicated picture. Speculatively, the absence of SD from Stimulus 2 to Response 3 in Experiment 1 suggests that actively coded information held in VWM may become less susceptible to external visual interference over time, highlighting the importance of the memory retention stage in explaining SD. However, we did not have a control condition where Simulus 1 would be uncued but still reported twice, making it difficult to make any strong claims. We also found that the previous streak (S2) influences the current one (both R1 and R3) only when the latter is prioritized, supporting the 'vulnerability' hypothesis. At the same time, in Exps. 2 and 3 we did not find any effect of cueing, further complicating the matter.

One of the key points in this study is the effect of prioritization on memory maintenance. However, it raises an important question: at which stage does this prioritization influence the representation of stimuli? Does it affect encoding, or does it shape stimulus representation during the maintenance phase? Our findings suggest that the impact of prioritization emerges specifically during the maintenance stage. A stronger SD in the next trial—but not a weaker SD in the current trial—suggests that encoding remains unaffected by prioritization. Additionally, Experiments 2 and 3 showed no effect of prioritization, suggesting that the observed phenomenon is related to a late phase of maintenance rather than encoding.

## Conclusion

In summary, our findings from three experiments highlight the nuanced effects of prioritization of a representation in memory, visual interference, and their impact on SD. Our results both confirm and challenge prior research, revealing that active memory maintenance leads to stronger SD, suggesting a heightened bias inherent to these conditions beyond mere attentional effects. Additionally, the prioritization of the current stimuli does not always significantly influence its susceptibility to bias; rather, biases from previous streaks can persist even when current stimuli in VWM are prioritized. These insights contribute to a deeper understanding of how memory and attention shape perceptual judgments and biases in sequential decision-making tasks.

## Methods and Procedure

### *Power analyses*

We conducted simulation-based power analysis for both experiments to determine the required sample sizes for detecting significant effects. Power was defined as the proportion of simulations where the null hypothesis was rejected. For each scenario, we simulated 1000 datasets with 18 participants each (18 was considered a minimum size based on the conventions in the field) and 648 trials for each participant, assuming a 1-degree difference in the magnitude of a SD effect (based on the previous literature: Ceylan et al., 2021). The noise and the baseline magnitude of SD in the observer's responses were estimated based on Houborg et al. (2023).

Subsequently, a statistical test was conducted for each simulated sample (t-test for Experiment 1 and repeated-measures ANOVA for Experiment 2 and 3, both at an alpha level of 0.05). Based on the results, a sample size of 18 participants was sufficient for detecting significant effects in both experiments (Experiment 1: power = 0.94; Experiment 2 and 3: power = 0.99). The code used for the power analysis is available on the Open Science Framework (OSF) and can be accessed via the following link: https://osf.io/wunf8/?view_only=079afc6a31904559b93df4fa4debfc0c

### *Experiment 1*

#### *Participants*

Eighteen volunteers (10 women; $M_{age}$ = 25.9 years, $SD_{age}$ = 4.3 years) participated in the experiment in exchange for monetary compensation. All of them had normal or corrected-to-normal vision. The research protocol for this and the following studies was approved by the local Ethics Committee (Human Inspired Technology Research Centre - HIT, protocol number 2023_236R2). Prior to the experiment, all participants provided written informed consent and were informed about the general purpose of the study and the experimental procedures.

#### *Stimuli, Design, and Procedure*

The procedure consisted of two identical sessions conducted on separate days, each comprising 324 trials (for a total of 648 trials). The experiment began with 6 practice trials, followed by the main experimental phase, which included 9 blocks of 36 trials each. Instructions were displayed on the computer screen at the beginning of each session. Each session took approximately 1.5 hours, with participants allowed to take breaks between blocks. Stimulus presentation and response collection were managed using PsychoPy software v.2023.2.0 (Peirce et al., 2019), using an HP p1230 screen (85 Hz, 1920 × 1440 resolution). Participants were positioned approximately 60-65 cm away from the screen.

Each trial consisted of the presentation of a Gabor patch followed by an adjustment task (see Figure 6). The Gabor patches had a diameter of 5.5 degrees of visual angle (dva) and a frequency of 2 cycles/dva, with RGB values of 0 and 255 for min and max luminance, respectively. They were displayed at the center of the computer screen against a gray background (RGB 128). The orientation of the Gabor patches was randomized and varied between 0 and 180 degrees. To clarify the design in the following text, we use the term "streak" to describe two consecutively presented stimuli: the first Gabor patch (Stimulus 1) was either cued or non-cued, while the second (Stimulus 2) was always non-cued (see below).

The experiment included two conditions: a standard report condition (67% of streaks) and a report-and-hold-in-memory condition (33% of streaks). The two conditions were randomly interleaved.

In the standard report condition, we instructed participants to report the orientation of each Gabor patch immediately after its presentation (respectively Response 1 and 2). Throughout all intervals, a fixation point (a white cross with a size of 0.05 dva) appeared at the center of the screen. The streak started with a 2000 ms fixation cross, followed by the first Gabor patch displayed for 500 ms. Then, a 2000 ms fixation cross was presented. Participants then had 4500 ms to complete the adjustment task, in which they saw a circle with a diameter of 2 dva containing a bar that they had to rotate to match the orientation of the Gabor patch using the right and left arrow keys. Pressing the spacebar confirmed their response. After another 1000 ms fixation cross, we presented the second Gabor patch for 500 ms, followed by a 2000 ms fixation cross and a second response period of 4500 ms.

In the report-and-hold-in-memory condition, participants additionally saw a cue (red circle, 1.2 dva diameter, 1000 ms) before the first Gabor was displayed (Figure 3). This cue indicated that they would need to report the orientation of the first Gabor twice: once immediately after its presentation (Response 1) and again (Response 3) after reporting the orientation of the second Gabor (Response 2), with a 1000 ms pause in between. During Response 3, an additional instruction reminded participants that they needed to report the orientation of the cued stimulus. This task required participants to retain the first stimulus in memory while perceiving and reporting the second stimulus.

Throughout the experiment, participants were instructed to fixate on the center of the screen and report the orientation of each stimulus as accurately as possible.

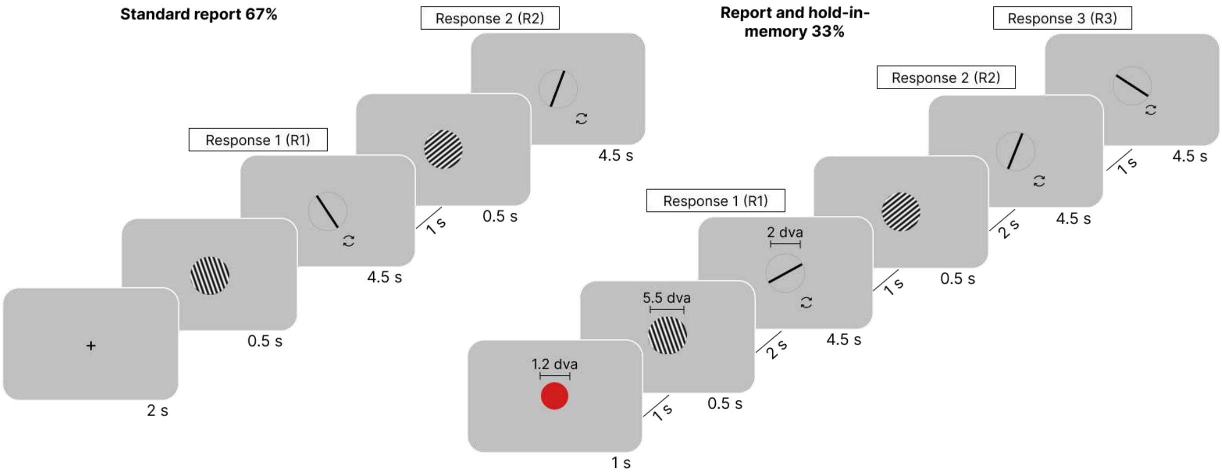


**Figure 6.** *Design of Experiment 1.* In the standard procedure (66% of trials), participants viewed a Gabor patch and were required to report its orientation immediately after the presentation. In the report-and-hold-in-memory condition (33% of trials), participants first received a cue before the initial Gabor patch and reported the orientation of the first Gabor twice: once immediately after its presentation (Response 1) and again (Response 3) after reporting the orientation of the second Gabor (Response 2). The stimuli depicted in the figure are not drawn to scale. The fixation cross was also shown during the intervals between stimuli and reports (not drawn for conciseness). The circular arrows shown near the response bar were not part of the actual experiment.

### *Experiment 2*

### *Participants*

18 participants (7 women; $M_{age}$ = 26.8 years, $SD_{age}$ = 5.1 years) took part in Experiment 2 (four participants were excluded due to low accuracy, defined as a circular standard deviation of response errors greater than 30 degrees, and replaced with new ones). The study was conducted online through the Prolific platform in exchange for a monetary reward. All participants had normal or corrected-to-normal vision and provided online informed consent before participating in the study.

***Stimuli, Design, and Procedure***

A credit card adjustment procedure was used to control the size of the visual stimuli in this online experiment (Li et al., 2020). Participants were instructed to position themselves at a distance of approximately 60-65 cm from the computer screen.

The experiment was conducted in a single session consisting of 720 trials. It began with a practice part of 12 trials, followed by the main experimental part divided into 10 blocks of 72 trials each. The entire procedure lasted approximately 1.5 hours, with participants allowed to take breaks between blocks. The experiment was developed using PsychoPy software v.2023.2.0, and responses were collected via the online platform Pavlovia.org (Peirce et al., 2019).

Each trial began with a white fixation cross, sized at 0.05 dva, at the center of the screen, which remained present throughout the experiment. Participants fixated on this cross for 1000 ms before a red pre-cue circle, with a diameter of 0.8 dva, randomly appeared on either the left or right side of the screen for another 1000 ms. This pre-cue indicated which of the upcoming Gabor patches participants needed to memorize. After the pre-cue, two Gabor patches (diameter: 3.3 dva, spatial frequency: 8 cycles/dva; min and max RGB values of the Gabor patches: 0 and 255, respectively) were simultaneously displayed for 1000 ms on both sides of the screen, centered 4 dva from the screen's center. One Gabor patch served as the target and the other as a non-target, with their orientations independently randomized between 0 and 180 degrees. The stimuli were presented against a gray background (128 RGB).

The study design consisted of two conditions: a congruent condition (75% of trials) and an incongruent condition (25% of trials). We instructed participants to report the orientation of the cued Gabor patch. In the congruent condition, the adjustment bar subsequently appeared on the same side as the cued Gabor patch, and participants had to report the orientation of that Gabor. In the incongruent condition, the adjustment bar subsequently appeared on the opposite side of the cued Gabor patch, and participants had to report the orientation of the non-cued Gabor (e.g., if the right Gabor was cued, but they were asked to report the orientation of the left Gabor, see Figure 7). Throughout the experiment, we instructed participants to maintain their gaze on a fixation cross at the center of the screen and report the stimulus orientation as accurately as possible. In each trial, congruent and incongruent conditions were selected using a weighted random process, with congruent conditions occurring three times more frequently. In contrast to Experiment 1, no predefined streaks were imposed.

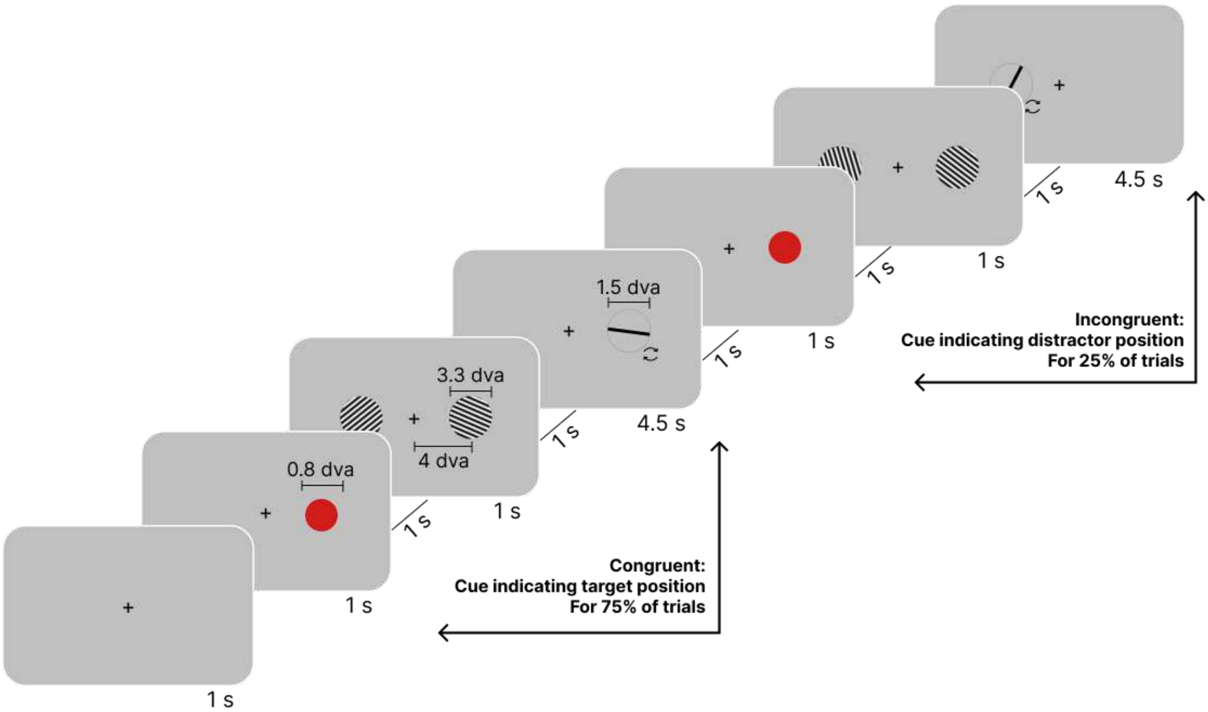


**Figure 7.** *Design of Experiment 2.* Participants reported the orientation of a Gabor patch at the location matching the adjustment bar location. This Gabor patch was either cued (congruent condition, 75% of trials) or non-cued (incongruent, 25% of trials). The stimuli depicted in the figure are not drawn to scale. The fixation cross was also shown during the intervals between stimuli and reports (not drawn here). The circular arrows shown near the response bar were not part of the actual experiment.

***Experiment 3***

***Participants***

18 participants (7 women; $M_{age}$ = 25.38 years, $SD_{age}$ = 3.97) took part in Experiment 3 online through the Prolific platform in exchange for a monetary reward (twelve participants were excluded due to low accuracy based on the same criteria used in Experiment 2 and

replaced with new ones). All participants had normal or corrected-to-normal vision and provided online informed consent before participating.

### *Stimuli, Design, and Procedure*

The design of Experiment 3 was identical to Experiment 2, with the key difference being the use of a post-cue instead of a pre-cue.

The post-cue appeared 500 ms after the stimuli presentation and lasted for 100 ms. After an additional 400 ms, the adjustment task began. After the task was completed, there was an inter-trial interval of 3000 ms. This timing was designed to maintain a 1000 ms pause between stimuli and response, and a 3000 ms interval between the response and the next stimulus, in accordance with the timing used in Experiment 2.

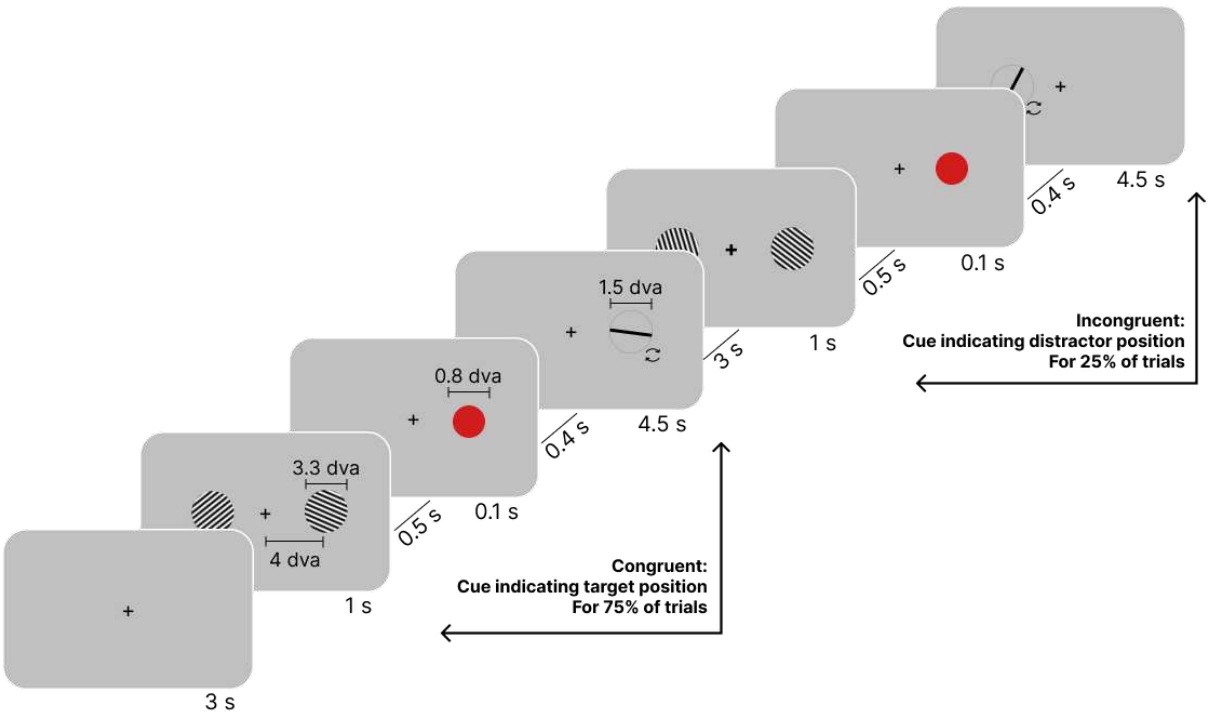


**Figure 8.** *Design of Experiment 3.* Participants reported the orientation of a Gabor patch at the location matching the adjustment bar location. This Gabor patch was either cued (congruent condition, 75% of trials) or non-cued (incongruent, 25% of trials). The stimuli depicted in the figure are not drawn to scale. The fixation cross was also shown during the intervals between stimuli and reports (not drawn for conciseness). The circular arrows shown near the response bar were not part of the actual experiment.

## Declarations

### Ethics approval and consent to participate

The study was approved by the local Ethics Committee (Human Inspired Technology Research Centre - HIT, protocol number 2023_236R2). Prior to the experiment, all participants provided written informed consent and were informed about the general purpose of the study and the experimental procedures.

### Consent for publication

Not applicable.

### Availability of data and materials

The datasets and data analysis supporting the conclusions of this article are available in the Open Science Framework (OSF) and can be accessed via the following link: https://osf.io/wunf8/?view_only=079afc6a31904559b93df4fa4debfc0c

### Competing interests

The authors declare that they have no competing interests.

### Funding

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no 101034319 and from the European Union — NextGenerationEU.

### Authors' contributions

**EA** conceptualized the study, conducted the investigation, curated and analyzed the data, and wrote the original draft. **AC** supervised the study, contributed to conceptualization, methodology, investigation, formal analysis, and manuscript review and editing. **GC** supervised the study, contributed to conceptualization, methodology, and investigation, participated in manuscript review and editing, and managed the project.

### Acknowledgments

Not applicable.

## References

1. Bae GY, Luck SJ. Reactivation of previous experiences in a working memory task. Psychol Sci. 2019;30(4):587-95. doi:10.1177/0956797619830398.
2. Barth A, Schneider D. Manipulating the focus of attention in working memory: Evidence for a protection of multiple items against perceptual interference. Psychophysiology. 2018;55(7):e13062. doi:10.1111/psyp.13062.
3. Barbosa J, Compte A. Build-up of serial dependence in color working memory. Sci Rep. 2020;10(1):10959. doi:10.1038/s41598-020-67861-2.
4. Barbosa J, Stein H, Martinez RL, Galan-Gadea A, Li S, Dalmau J, Adam KCS, Valls-Solé J, Constantinidis C, Compte A. Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. Nat Neurosci. 2020;23(8):1016-24. doi:10.1038/s41593-020-0644-4.
5. Bays PM. Noise in neural populations accounts for errors in working memory. J Neurosci. 2014;34(10):3632-45. doi:10.1523/JNEUROSCI.3204-13.2014.
6. Bettencourt KC, Xu Y. Decoding the content of visual short-term memory under distraction in occipital and parietal areas. Nat Neurosci. 2016;19(1):150-7. doi:10.1038/nn.4174.
7. Bliss DP, Sun JJ, D'Esposito M. Serial dependence is absent at the time of perception but increases in visual working memory. Sci Rep. 2017;7(1):14739. doi:10.1038/s41598-017-15199-7.
8. Ceylan G, Herzog MH, Pascucci D. Serial dependence does not originate from low-level visual processing. Cognition. 2021;212. doi:10.1016/j.cognition.2021.104709.
9. Chen K, Bae GY. Working memory flips the direction of serial bias through memory-based decision. Cognition. 2024;250:105843. doi:10.1016/j.cognition.2024.105843

10. Chetverikov A. Demixing model: A normative explanation for inter-item biases in memory and perception. 2023a. Preprint at: https://www.biorxiv.org/content/10.1101/2023.03.26.534226v2

11. Chetverikov A. circhelp: Circular analyses helper functions (Version 1.1.1) [Computer software]. GitHub. 2023b. Available from: https://github.com/achetverikov/circhelp.

12. Cicchini GM, Mikellidou K, Burr DC. The functional role of serial dependence. Proc R Soc B Biol Sci. 2018;285(1890). doi:10.1098/rspb.2018.1722.

13. Cicchini GM, Mikellidou K, Burr DC. Serial dependence in perception. Annu Rev Psychol. 2024;75(1):129-54. doi:10.1146/annurev-psych-021523-104939.

14. Czoschke S, Fischer C, Beitner J, Kaiser J, Bledowski C. Two types of serial dependence in visual working memory. Br J Psychol. 2019;110(2):256-67. doi:10.1111/bjop.12349.

15. Fischer C, Czoschke S, Peters B, Rahm B, Kaiser J, Bledowski C. Context information supports serial dependence of multiple visual objects across memory episodes. Nat Commun. 2020;11(1):1932. doi:10.1038/s41467-020-15874-w.

16. Fischer C, Kaiser J, Bledowski C. A direct neural signature of serial dependence in working memory. eLife. 2024;13:e99478. doi:10.7554/eLife.99478.

17. Fischer J, Whitney D. Serial dependence in visual perception. Nat Neurosci. 2014;17(5):738-43. doi:10.1038/nn.3689.

18. Fritsche M, De Lange FP. The role of feature-based attention in visual serial dependence. J Vis. 2019;19(13):21. doi:10.1167/19.13.21.

19. Fritsche M, Spaak E, De Lange FP. A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. eLife. 2020;9:e55389. doi:10.7554/eLife.55389.

20. Gayet S, Guggenmos M, Christophel TB, Haynes JD, Paffen CLE, Van Der Stigchel S,

et al. Visual working memory enhances the neural response to matching visual input. J Neurosci. 2017;37(28):6638-47. doi:10.1523/JNEUROSCI.3418-16.2017.

21. Hajonides JE, Van Ede F, Stokes MG, Nobre AC, Myers NE. Multiple and dissociable effects of sensory history on working-memory performance. J Neurosci. 2023;43(15):2730-40. doi:10.1523/JNEUROSCI.1200-22.2023.
22. Houborg C, Kristjánsson Á, Tanrıkulu ÖD, Pascucci D. The role of secondary features in serial dependence. J Vis. 2023;23(5):21. doi:10.1167/jov.23.5.21.
23. Jakob AM, Gershman SJ. Rate-distortion theory of neural coding and its implications for working memory. eLife. 2023;12:e79450. doi:10.7554/eLife.79450.
24. Kim S, Burr D, Cicchini GM, Alais D. Serial dependence in perception requires conscious awareness. Curr Biol. 2020;30(6):R257-8. doi:10.1016/j.cub.2020.02.008.
25. Kiyonaga A, Scimeca JM, Bliss DP, Whitney D. Serial dependence across perception, attention, and memory. Trends Cogn Sci. 2017;21(7):493-7. doi:10.1016/j.tics.2017.04.011.
26. LaRocque JJ, Riggall AC, Emrich SM, Postle BR. Within-category decoding of information in different attentional states in short-term memory. Cereb Cortex. 2016;26(6):2721-31. doi:10.1093/cercor/bhw283.
27. Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. Neural evidence for a distinction between short-term memory and the focus of attention. J Cogn Neurosci. 2012;24(1):61-9. doi:10.1162/jocn_a_00140.
28. Li Q, Joo SJ, Yeatman JD, Reinecke K. Controlling for participants' viewing distance in large-scale, psychophysical online experiments using a virtual chinrest. Sci Rep. 2020;10(1):904. doi:10.1038/s41598-019-57204-1.
29. Little Z, Clifford CWG. Effects of perceptual and decisional uncertainty on serial dependence in orientation perception. Atten Percept Psychophys. 2025.

doi:10.3758/s13414-025-03034-5.

30. Makovski T, Pertzov Y. Attention and memory protection: Interactions between retrospective attention cueing and interference. Q J Exp Psychol. 2015;68(9):1735-43. doi:10.1080/17470218.2015.1049623.
31. Mallett R, Lewis-Peacock JA. Working memory prioritization impacts neural recovery from distraction. Cortex. 2019;121:225-38. doi:10.1016/j.cortex.2019.08.019.
32. Manassi M, Murai Y, Whitney D. Serial dependence in visual perception: A meta-analysis and review. J Vis. 2023;23(8):18. doi:10.1167/jov.23.8.18.
33. Manassi M, Whitney D. Continuity fields enhance visual perception through positive serial dependence. Nat Rev Psychol. 2024;3(5):352-66. doi:10.1038/s44159-024-00297-x.
34. Mei G, Chen S, Dong B. Working memory maintenance modulates serial dependence effects of perceived emotional expression. Front Psychol. 2019;10:1610. doi:10.3389/fpsyg.2019.01610.
35. Myers NE, Chekroud SR, Stokes MG, Nobre AC. Benefits of flexible prioritization in working memory can arise without costs. J Exp Psychol Hum Percept Perform. 2018;44(3):398-411. doi:10.1037/xhp0000449.
36. Myers NE, Stokes MG, Nobre AC. Prioritizing information during working memory: Beyond sustained internal attention. Trends Cogn Sci. 2017;21(6):449-61. doi:10.1016/j.tics.2017.03.010.
37. Pascucci D, Tanrikulu ÖD, Ozkirli A, Houborg C, Ceylan G, Zerr P, et al. Serial dependence in visual perception: A review. J Vis. 2023;23(1):9. doi:10.1167/jov.23.1.9.
38. Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, et al. PsychoPy2: Experiments in behavior made easy. Behav Res Methods. 2019;51(1):195-

203. doi:10.3758/s13428-018-01193-y.

39. Rose NS, LaRocque JJ, Riggall AC, Gosseries O, Starrett MJ, Meyering EE, et al. Reactivation of latent working memories with transcranial magnetic stimulation. Science. 2016;354(6316):1136-9. doi:10.1126/science.aah7011.

40. Saito JM, Printzlau F, Yeo Y, Fukuda K. Working memory prioritization changes bidirectional interactions with perceptual inputs. 2024. Preprint at: https://osf.io/9qzw7/

41. Shan J, Hajonides JE, Myers NE. Neural evidence for decision-making underlying attractive serial dependence. 2024. Preprint at: https://www.biorxiv.org/content/10.1101/2024.11.18.624176.abstract

42. Sprague TC, Ester EF, Serences JT. Restoring latent visual working memory representations in human cortex. Neuron. 2016;91(3):694-707. doi:10.1016/j.neuron.2016.07.006.

43. van Bergen RS, Jehee JFM. Probabilistic representation in human visual cortex reflects uncertainty in serial decisions. J Neurosci. 2019;39(41):8164-76. doi:10.1523/JNEUROSCI.3212-18.2019.

44. van Moorselaar D, Battistoni E, Theeuwes J, Olivers CNL. Rapid influences of cued visual memories on attentional guidance. Ann N Y Acad Sci. 2015a;1339(1):1-10. doi:10.1111/nyas.12574.

45. van Moorselaar D, Gunseli E, Theeuwes J, Olivers CNL. The time course of protecting a visual memory representation from perceptual interference. Front Hum Neurosci. 2015b;8:1053. doi:10.3389/fnhum.2014.01053.

46. Vergauwe E, Hautekiet C, Langerock N. Distractor susceptibility in visual working memory: No evidence for particularly vulnerable mnemonic representations in the focus of attention. PsyArXiv. 2023. doi:10.31234/osf.io/4yqjt.

47. Vincent BT. A tutorial on Bayesian models of perception. J Math Psychol.

2015;66:103-14. doi:10.1016/j.jmp.2015.02.001.

48. Wang D, Chota S, Xu L, van der Stigchel S, Gayet S. The priority state of items in visual working memory determines their influence on early visual processing. Conscious Cogn. 2025;127:103800. doi:10.1016/j.concog.2024.103800.

49. Wolff MJ, Jochim J, Akyürek EG, Stokes MG. Dynamic hidden states underlying working-memory-guided behavior. Nat Neurosci. 2017;20(6):864-71. doi:10.1038/nn.4546.

50. Yoo AH, Klyszejko Z, Curtis CE, Ma WJ. Strategic allocation of working memory resource. Sci Rep. 2018;8(1):16162. doi:10.1038/s41598-018-34282-1.

51. Zhang Z, Lewis-Peacock JA. Bend but don’t break: Prioritization protects working memory from displacement but leaves it vulnerable to distortion from distraction. Cognition. 2023a;239:105574. doi:10.1016/j.cognition.2023.105574.

52. Zhang Z, Lewis-Peacock JA. Prioritization sharpens working memories but does not protect them from distraction. J Exp Psychol Gen. 2023b;152(4):1158-74. doi:10.1037/xge0001309.