Fast Capture, Slow Shift: How Working Memory Guides Perception

Hyung-Bum Park[1] Weiwei Zhang[2]

[1]Institute for Mind and Biology, University of Chicago

[2]Department of Psychology, University of California, Riverside

Abstract

The top-down influence of working memory (WM) can manifest as attentional capture and a “tinted lens” that alters perceptual appearance. Yet it remains unclear whether these effects arise from a common mechanism or reflect functionally and mechanistically distinct processes. Across two experiments, we embedded a perceptual estimation task during WM maintenance. Hierarchical Bayesian mixture modeling revealed robust bidirectional attraction between memory and perception. More importantly, time-resolved analyses of mouse trajectories showed that attentional capture emerged as a fast, transient deviation linked to movement initiation, whereas perceptual bias reflected a slower, sustained shift shaping the final perceptual judgment. Notably, the prospective influence of WM on perception engaged both effects, whereas the retrospective influence of perception on WM involved only the sustained shift. These findings indicate that WM deploys distinct top-down signals that operate over different timescales to dynamically shape our perceptual experience.

Keywords: working memory, perceptual appearance, attentional capture, mouse tracking, cognitive penetrability

Working memory (WM) is a core cognitive system for actively maintaining task-relevant information to guide behavior (Cowan, 2001). Beyond short-term storage, WM dynamically configures perceptual processing in anticipation of upcoming sensory input, serving as a predictive control state that connects moment-to-moment perception with higher-order cognition (Myers et al., 2017; Nobre & Stokes, 2019). In this view, WM is not just a record of recent experiences but an active state that prepares the perceptual system for what is likely to occur next. By maintaining goal-relevant information active, WM tunes perceptual processing in advance of incoming input, linking memory, attention, and perception within a predictive control framework (Friston, 2018; Kiyonaga et al., 2017; Summerfield & de Lange, 2014).

This top-down influence manifests in at least two fundamental ways. First, goal-relevant features instantiated in WM act as attentional templates that bias sensory gain toward matching inputs (Desimone & Duncan, 1995; Wolfe, 1994). Behaviorally, this often appears as attentional capture, so that attention is automatically drawn to memory-matching input, even when this is detrimental to the task at hand (Olivers et al., 2006; Soto et al., 2008; Park & Zhang, 2024; Won & Zhang, 2025). Second, and more controversially, WM may directly alter perceptual appearance, acting as a "tinted lens" that pulls new sensory information toward remembered features (Teng & Kravitz, 2019; Teng et al., 2023). This tinted-lens account posits that WM content can pull nearby sensory representations toward itself, producing reliable distortions in perceived color, orientation, or shape. This account aligns with sensory-recruitment theories, which propose that sensory cortices contribute to both perception and memory maintenance (Harrison & Tong, 2009; Serences et al., 2009; Rademaker et al., 2019), as well as attractor-dynamics models in which stable memory states act as attractor basins that bias the readout of sensory evidence (Burak & Fiete, 2009; Wimmer et al., 2014).

Historically, these two processes, *attentional capture* and *perceptual bias*, have been investigated as separate phenomena, often with different paradigms. Studies of attentional guidance typically rely on visual search tasks to measure reaction time costs as an index of shifts in attentional selection priority. Studies of perceptual bias, in contrast, use fine-grained discrimination or estimation tasks to quantify systematic distortions in appearance, indexing representational change. This separation in methodology leaves a fundamental question unanswered: Are attentional capture and perceptual bias simply different manifestations of a single underlying mechanism, or do they reflect functionally and mechanistically distinct computations that unfold on different timescales? For instance, is attentional capture merely the initial, transient stage of a process that culminates in a perceptual bias, or do they operate in parallel, each serving a separate functional role?

Answering these questions requires a method capable of tracking how WM influences perception over time within a single trial. Conventional measures such as accuracy and reaction time are insufficient, as they conflate early attentional events with later representational changes. Neurophysiological evidence suggests that WM's influence can indeed operate on multiple timescales. For instance, event-related potential studies have shown that WM-matching items modulate early visual components (e.g., P1/N1, around 100–200 ms post-stimulus), consistent with rapid attentional modulation, as well as later components associated with conflict resolution and response mapping, which occur over 500 ms (Luck et al., 2000; Tan et al., 2014). These findings suggest temporarily distinct pathways, yet a clear behavioral signature of this dissociation has remained elusive.

To address this gap, we developed a dual-task paradigm that embeds a perceptual color-matching task within the maintenance interval of a WM task. Using continuous report measures

and time-resolved mouse cursor trajectory analyses (Park, 2025), we tracked the moment-to-moment evolution of perceptual decisions. This approach allows us to operationalize and dissociate two potential influences of WM. We define *attentional capture* as an early, transient deviation in the movement trajectory toward the WM-matching feature (Theeuwes et al., 2005) that is not explained by the final judgment, and *representational shift* as a sustained deviation that is consistent with the final, biased report. This decomposition provides a continuous behavioral window into how WM shapes both the dynamics and outcomes of perception. Across two experiments, we show that WM content can simultaneously capture spatial attention and bias perceptual appearance, but that these effects emerge on distinct temporal scales. Capture arises rapidly and dissipates, reflecting transient priority competition, whereas bias develops gradually and persists, reflecting sustained assimilation toward the maintained memory representation. Together, these findings reveal how distinct top-down signals from WM jointly shape visual experience in real time.

General Method

Participants

Thirty-three and twenty-one college students at the University of California, Riverside participated in Experiment 1 and 2, respectively, for course credit. All participants reported normal or corrected-to-normal vision and normal color vision. Informed consent was obtained prior to participation. The experimental protocol was approved by the University of California, Riverside Institutional Review Board.

For Experiment 1, the sample size was determined based on previous work adopting a closely matched experimental procedure (Park & Zhang, 2024). A priori power analysis in G*Power 3.1 (Faul et al., 2007) for a two-tailed paired *t*-test with $\alpha = .05$ and a predicted effect

size of *dz* = 0.50 to achieve approximately 80% power suggested a sample size of 34 participants. The sample size for Experiment 2 was guided by the empirical benchmark from Experiment 1, where the primary condition contrast in perceptual matching error produced a substantially larger effect than anticipated (*d* = 1.29). This robust within-subject effect suggested that a smaller sample could still provide adequate power to detect the predicted reciprocal bias between perceptual and WM. It should be noted that the primary analyses were conducted using hierarchical Bayesian modeling and resampling-based permutation tests for statistical inference (see *Data Analysis* for detail). With 21 participants, the Bayesian analyses yielded decisive posterior credible intervals for the population-level parameter of key interest, and the permutation tests detected the predicted dissociation effects, supporting the adequacy of the sample size.

Stimuli and Procedure

Stimuli were presented on a 60 Hz LCD monitor calibrated with an X-Rite i1Pro spectrophotometer, with a grey background (15.1 cd/m²) at a viewing distance of 60 cm. Task procedures were identical in Experiments 1 and 2, except for a test array (Figure 1). Each trial began with central fixation (800 ms), followed by a sample display (100 ms) of a single colored square (1.6° × 1.6°), which appeared at one of four equally spaced locations on an imaginary circle (radius 4.3°) centered at fixation. The memory color was selected randomly from 180 evenly spaced hues on a circular CIELAB color space (Zhang & Luck, 2008).

After an unfilled 400-ms delay, the perceptual matching display appeared. A colored disk (1.6° diameter) replaced the central fixation point, surrounded by a continuous color-wheel (radius 8.2°, thickness 2.2°) with 180 evenly distributed hues from the same color space. The target color for the perceptual matching task was offset by either 40° clockwise (CW) or

counterclockwise (CCW) from the memorized color in the sample display. This relative offset was strategically chosen to fall within the similarity range that maximizes attraction in previous serial-dependence research (Fischer & Whitney, 2014). The target item remained visible until response, and participants reported the perceived hue of the target by moving the mouse cursor from the center of the screen to the best-matching color on the color-wheel with a click. Mouse trajectories (x and y coordinates and time relative to the cursor onset) were recorded continuously. Immediately after the click, 500 ms visual feedback appeared on the color-wheel, with a cross marking the color the participants picked and an arrow marking the target color. The color-wheel was randomly rotated on every trial to prevent fixed spatial mapping of the colors to locations on the wheel. The mouse cursor was re-positioned at fixation after the visual feedback.

Following another unfilled 500 ms delay, a test display occurred, which differed by experiment. Experiment 1 adopted a change-detection, in which the task array presented a single probe at the original memory item location, and participants indicated whether the probe matched or mismatched the remembered color by pressing one of two shoulder buttons on a gamepad (e.g., L1 for “no-change”, L2 for “change”). The participants used their dominant hand (right hand for all subjects) for the mouse clicking for the perceptual matching task and the nondominant hand for the change-detection task. Change and no-change trials occurred with equal probability. On change trials, the probe color rotated 60°, 120°, or 180° in either direction from the memory color with equal probability. Because the preceding perceptual matching target was always positioned 40° CW or CCW relative to the memory color, this systematic manipulation of change magnitude allowed us to test whether the perceptual judgment made during WM maintenance influenced the subsequent representation of the memory item. Auditory

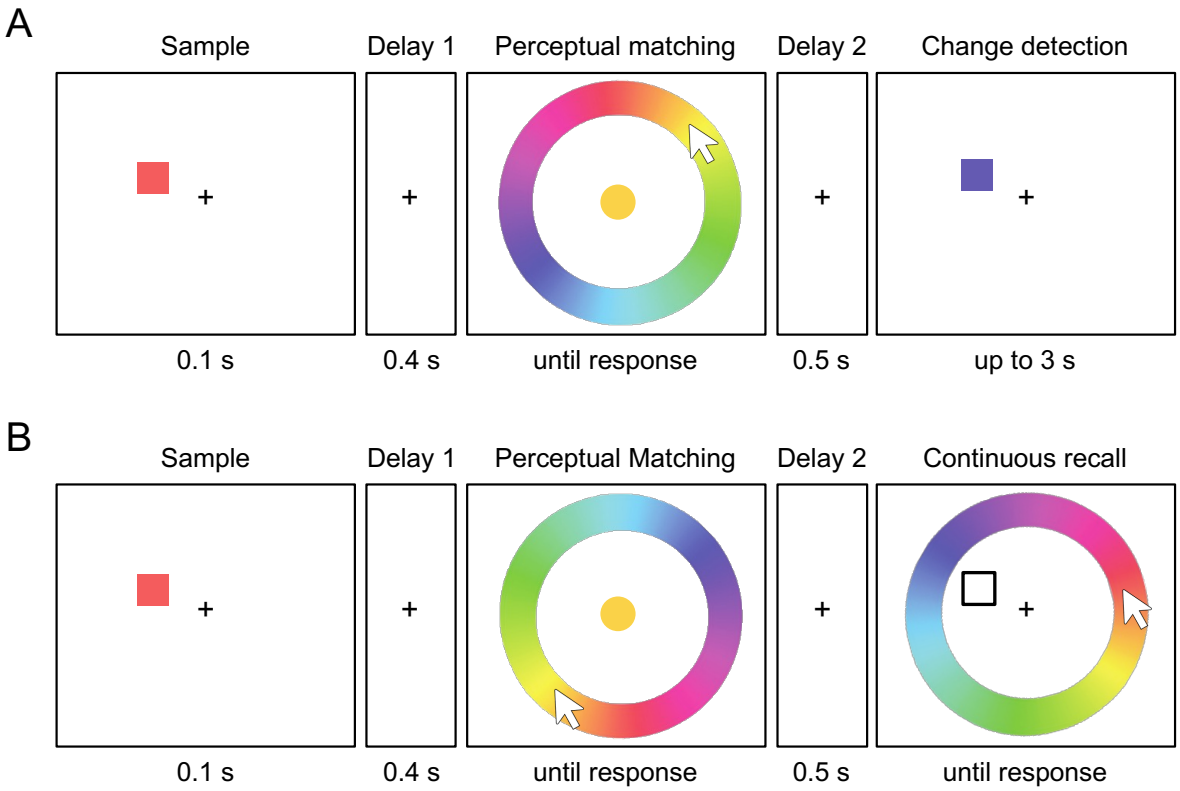

Figure 1. Task procedure for Experiment 1 (A) and Experiment 2 (B). (A) Experiment 1 embedded a perceptual matching task within the delay interval of a working memory (WM) change-detection task. After viewing a sample color and a short delay, participants reported the color of a continuously visible target using a color-wheel (perceptual matching), followed by a second delay and a change-detection, where change magnitude varied at every 60° in both directions (−180°:60°:180°). (B) Experiment 2 combined perceptual matching with continuous WM recall. After the sample and delay, participants performed a perceptual matching report, followed by a second delay and a continuous recall of the memorized color using a color-wheel.

feedback followed each response, consisting of a 500 ms low-pitch tone for correct answers and a high-pitch tone for errors.

Experiment 2 adopted a continuous recall paradigm, in which the test display consisted of an outlined square at the original location of the memory item and another color-wheel that had a circular rotation to random degrees in order to prevent a strategical mapping of stimulus to color-wheel location during the previous perceptual matching task. Participants reproduced the remembered color by moving the mouse cursor away from the center of the screen and then

clicking the best-matching color on the wheel. Immediate feedback was provided in the same manner as that for the perceptual matching task. In both experiments, the mouse cursor was automatically re-centered at fixation after each click, and all participant completed 300 trials in total, divided into three blocks with brief breaks.

Data Analysis

Hierarchical Bayesian Mixture Modeling

Perceptual matching error in Experiment 1 and 2 was defined as the signed angular distance in degrees between each reported color and the corresponding physical target color in the perceptual matching task. Similarly, memory recall error was defined as the signed angular distance in degrees between the reported and the original memory colors during WM recall in Experiment 2. These errors were modeled with a hierarchical Bayesian mixture model that extends the standard continuous estimation framework (Zhang & Luck, 2008; Park & Zhang, 2024). The model is expressed as a mixture of two distributions representing whether recall responses are based on noisy target representation or non-informative guesses (Figure 2). Specifically, the von Mises component captures responses based on noisy feature representations with mean (μ) that indexes inaccuracy or bias and concentration κ that indexes precision (reported as imprecision σ converted from κ). The uniform component captures guesses or lapses with mixture weight g, the height of the uniform probability distribution. Although guessing was expected to nearly minimal in the perceptual matching task since the target remained visible until response, the same modeling approach was applied to both perceptual matching and continuous WM recall errors in Experiment 2. This modeling approach permits a principled cross-task comparison of bias magnitude while accounting for occasional lapses or memory failures.

The bias effects in serial-dependence and WM are typically reliable but numerically small, often about a few degrees (Barbosa & Compte, 2020; Park, 2025). Hierarchical Bayesian modeling offers key advantages for this context, because it simultaneously accounts for multiple sources of variability (e.g., across participants, conditions, and trials), thereby stabilizing estimation of small effects (Oberauer et al., 2017; see also Frischkorn & Popov, 2025). This approach allows noisy parameter estimates to be shrunk toward the population mean, leading to more robust and reliable results (Shiffrin et al., 2008). The choice of reasonably informative to non-informative priors was determined following previous studies with similar modeling approach (Oberauer et al., 2017; Park & Zhang, 2024). We took a total of 4,000 samples after 4,000 warm-ups with four chains of Markov chain Monte Carlo sampling. It yielded posterior matrices of population-level parameters sized, for example, 12,000 (samples) × 33 (subjects) × 2 (conditions) for Experiment 1. Convergence was assessed using the Gelman-Rubin $\hat{R}$ statistic, with all population-level parameters showing values equal or close to 1.0 (Gelman & Rubin, 1992). Statistical inferences are reported in terms of posterior means and the 95% highest density credible intervals ($HDI_{95\%}$; Kruschke, 2015).

Mouse Trajectory

Mouse trajectories recorded during the perceptual matching tasks, as well as the continuous WM recall in Experiment 2, were first normalized in time from movement onset (0%) to the final clicking response on the color-wheel (100%). Movement onset was defined as the first sample where the cursor moved more than five pixels in any direction from the central starting position at the center of the screen. On average, participants initiated the cursor movement at about 25.7% [$CI_{95\%}$: 22.6%, 28.8%] of the normalized time scale.

To place all trials in a common reference frame, each trial's trajectory was spatially rotated so that the true target color mapped to the top of the wheel with the x-y coordinate at (0, *r*), where *r* is the color-wheel radius. This normalization ensured that any horizontal deviation directly reflected attraction or repulsion relative to the WM-matching color during perceptual matching, or relative to the intervening perceptual target during WM recall.

The direction and magnitude of trajectory curvature was then summarized as the signed area under the curve (AUC), computed as the integral of horizontal cursor offsets across normalized time. After confirming there was no statistically significant difference in magnitude, $t(32) = 0.08$, $p = .933$, Cohen's $d = 0.01$, for trials with clockwise (CW) or counterclockwise (CCW) targets relative to the WM item, the sign for CW AUC was flipped. Consequently, positive AUC values for perceptual matching trajectory indicate attraction toward the WM representation in both experiments. Similarly, negative AUCs for WM recall in Experiment 2 indicate attraction toward the earlier perceptual decision.

To distinguish hypothesized transient attentional capture from sustained appearance shifts, we employed a decomposition method utilizing an endpoint-bias correction (Park, 2025; Park & Zhang, 2022). For each trial, a reference path was calculated by directly connecting the starting point at fixation to the final clicking point on the wheel, while preserving the same radial progress over time. Subtracting this reference path from the observed trajectory produced a *capture* component that reflects trajectory deviations from the reference path occurring during the course movement, independent of the endpoint. The *shift* component was then defined as the remainder (i.e., overall curvature minus capture), indexing the portion of curvature consistent with the endpoint displacement. Expressing overall AUC as the sum of capture ($AUC_{capture}$) and shift area (AUC_{shift}) allowed us to isolate early transient attentional capture effects from sustained

representational displacements, respectively. Comparing these components allowed us to determine whether curvature and endpoint shifts aligned or diverged across conditions, thereby distinguishing the distinct contributions of capture and shift to the overall mouse trajectory pattern.

Beyond area-based measures, we further conducted time-resolved analysis using a recently developed destination-vector transformation (DVT) to capture moment-by-moment dynamics of decision making (Park & Zhang, 2024; Figure 5A). Rather than focusing on the instantaneous position of the cursor, DVT estimates the projected location on the continuous response scale at each time point, based on the vector direction defined by two consecutive positions. For a trajectory with n samples, DVT produces $n-1$ destination-vectors in circular degrees, directly indexing the temporal evolution of the intended choice. Averaging DVT across trials and participants yielded dynamic, high-resolution profiles of how WM or recent perceptual input shaped response trajectories throughout the decision process.

Results

Experiment 1: Perceptual Matching

Endpoint Bias Toward Working Memory

Perceptual matching errors as the endpoint of the mouse trajectory showed the predicted attraction toward the WM-matching color on the response color-wheel. Mean circular error was negative in the CW condition (i.e., the perceptual target color is CW to the memory item; $M = -2.44°$ [$CI_{95\%}$: −3.03°, −1.85°]) and positive in the CCW condition (+1.89° [$CI_{95\%}$: +1.09°, +2.68°]), yielding a reliable difference between conditions, $t(32) = -7.31$, $p < .001$, Cohen's $d = -1.29$. A Hierarchical Bayesian Mixture model confirmed these attraction effects (Figure 2). Posterior distributions of the population-level bias (μ) parameter were in opposite directions for

CW and CCW trials, and their contrast credible excluded zero, $\Delta\mu_{CCW-CW}$ = +4.00° [$HDI_{95\%}$: +3.37°, +4.59°], indicating robust attraction toward the WM color. Imprecision (σ) and guess rate (g) did not differ credibly between CW and CCW trials ($\Delta\sigma$ = 0.18° [$HDI_{95\%}$: −0.12°, 0.50°]; Δg = −0.03% [$HDI_{95\%}$: −0.32%, 0.32%]).

Mouse Trajectory Deviation Toward Working Memory

Next we assess the entire mouse trajectory beyond the endpoint to reveal the temporal dynamics of the retrieval and decision process. The overall trajectory for both CW and CCW trials showed a clear early deviation toward the WM color, which peaked soon after movement onset and declined as the movement progressed (Figure 3a), which manifested as the sizable horizontal deviations (Figure 3b) and the $AUC_{overall}$ (Figure 3c, 13.8 px [$CI_{95\%}$: 10.9 px, 16.8 px]), $t(32) = 9.28$, $p < .001$, Cohen's $d = 1.64$). Moreover, $AUC_{overall}$ scaled with how fast participants initiated mouse cursor movement upon perceptual target onset. A within-subject median-split on movement onset latency revealed larger deviations on early-onset trials (AUC_{early} = 24.1 px [$CI_{95\%}$: 20.0 px, 28.2 px]), than on late-onset trials (AUC_{late} = 9.6 px [$CI_{95\%}$: 6.8 px, 12.3 px]), $t(32) = 8.05$, $p < .001$, Cohen's $d = 1.42$. This speed contingency suggests that rapid movements are subject to a brief capture toward the WM-matching item, thereby magnifying overall trajectory curvature. In contrast, movement onset latency did not influence the final perceptual matching errors which showed a nonsignificant difference between early-onset trials (2.32° [$CI_{95\%}$: 1.64°, 3.00°] and late-onset trials (1.88° [$CI_{95\%}$: 1.21°, 2.55°]) did, $t(32) = 1.23$, $p = .229$, Cohen's $d = 0.22$. This null effect rules out the possibility that endpoint bias merely reflects any under-corrected motor deviation from initial capture, and instead suggests a genuine shift in perceptual appearance. Because $AUC_{overall}$ conflates these influences,

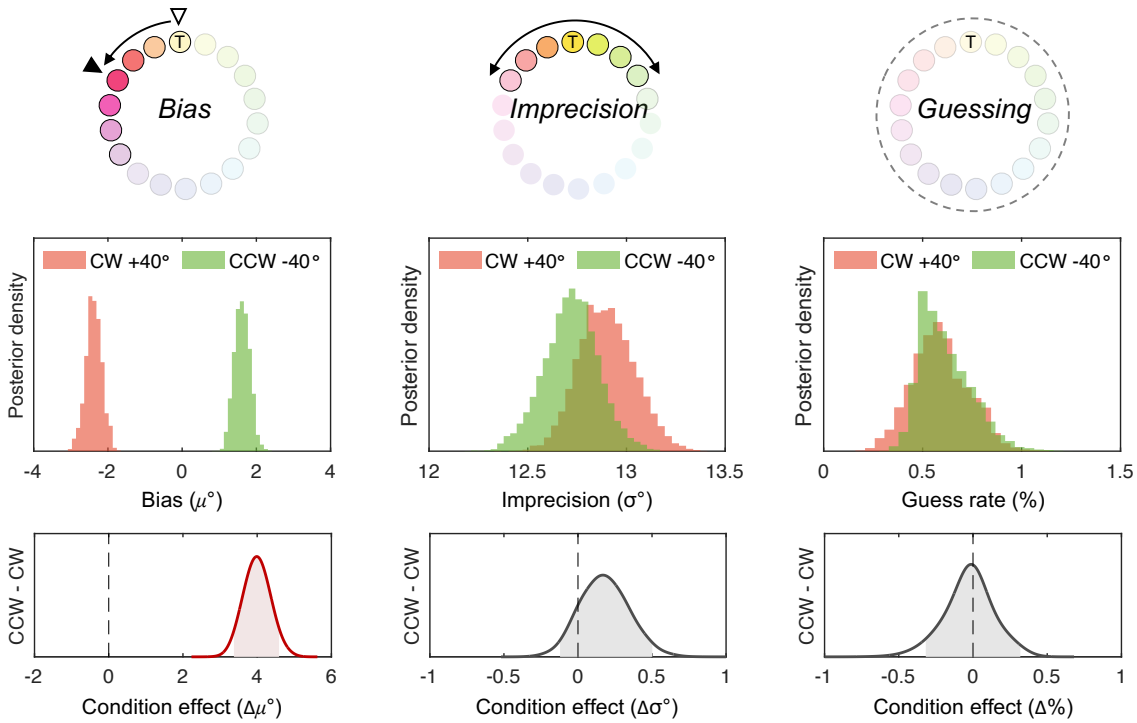

Figure 2. Hierarchical Bayesian parameter estimates for Experiment 1. Posterior distributions of the Mixture model population-level parameters are shown separately for CW (+40°) and CCW (−40°) conditions for bias (μ), imprecision (σ), and guess rate (g). The lower panels show posterior distributions of condition effects (CCW − CW) for each parameter. A credible condition difference was observed for bias (Δμ), with 95% highest density intervals ($HDI_{95\%}$) beyond zero. Shaded area represent $HDI_{95\%}$.

it is essential to decompose the measure to isolate the distinct contributions of attention capture and appearance change.

Mechanistic Dissociation: Capture versus Shift

To disentangle transient WM-driven capture from sustained representational shift, overall curvature was decomposed with an endpoint-bias correction (Figure 4A). For each trial, a reference line connecting the start and endpoint was subtracted from the observed path, yielding the *capture* component (deviations not explained by the endpoint). The *shift* component was defined as the residual (i.e., overall = capture + shift). We computed AUC for each component

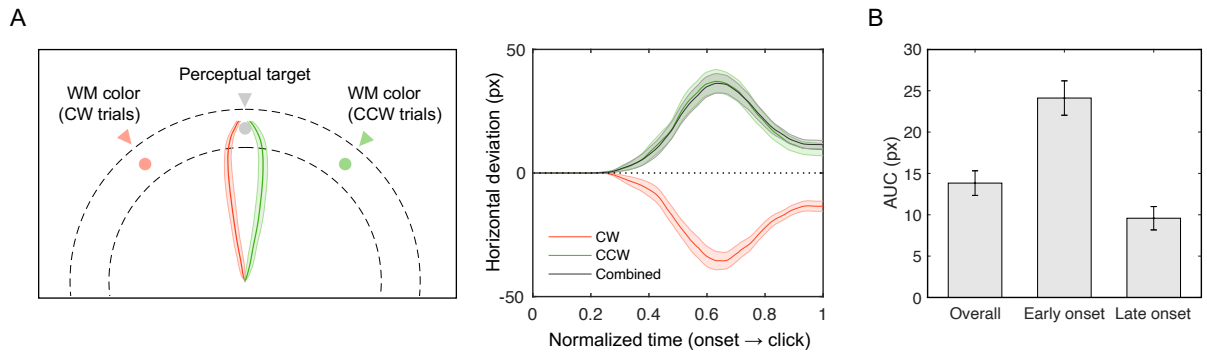

Figure 3. Mouse-tracking evidence for WM-induced perceptual bias in Experiment 1. (A) Average trajectory paths (left) and average horizontal trajectories (right), plotted separately for CW (+40°) and CCW (−40°) conditions. Shaded bands indicate 95% confidence intervals. (B) Area-under-the-curve (AUC) analyses of trajectory curvature. Early and late onset represent trials median split of movement onset latency. Error bars standard errors of the mean.

and conducted a set of individual-differences tests relating these components to movement onset latency and to the magnitude of the final perceptual shift.

This decomposition revealed a clear double dissociation (Figure 4B). The movement onset latency predicted $AUC_{capture}$, $r(31) = -.59$ [$CI_{95\%}$: −.77, −.30], $p < .001$, but not AUC_{shift}, $r(31) = -.08$ [$CI_{95\%}$: −.41, .27], $p = .654$. In contrast, the final perceptual bias predicted AUC_{shift}, $r(31) = .86$ [$CI_{95\%}$: .74, .93], $p < .001$, but not $AUC_{capture}$, $r(31) = .17$ [$CI_{95\%}$: −.18, .49], $p = .336$. These results indicate that WM-driven capture shaped evolving trajectories , more so for responses with faster initiation, whereas sustained representational shift determined altered perceptual appearance at the endpoint.

Time-Resolved Decision Dynamics: Destination-Vector Series

The analyses of movement latency and component dissociation suggested that WM influences manifest as both transient attention capture and sustained representational shift. To track how these influences unfolded over time, we used DVT to reconstruct the latent decision variables throughout the movement trajectory (Figure 5A). The decomposition (Figure 5B) mirrored the AUC findings. The capture time series ($DV_{CAPTURE}$) rose sharply at movement onset

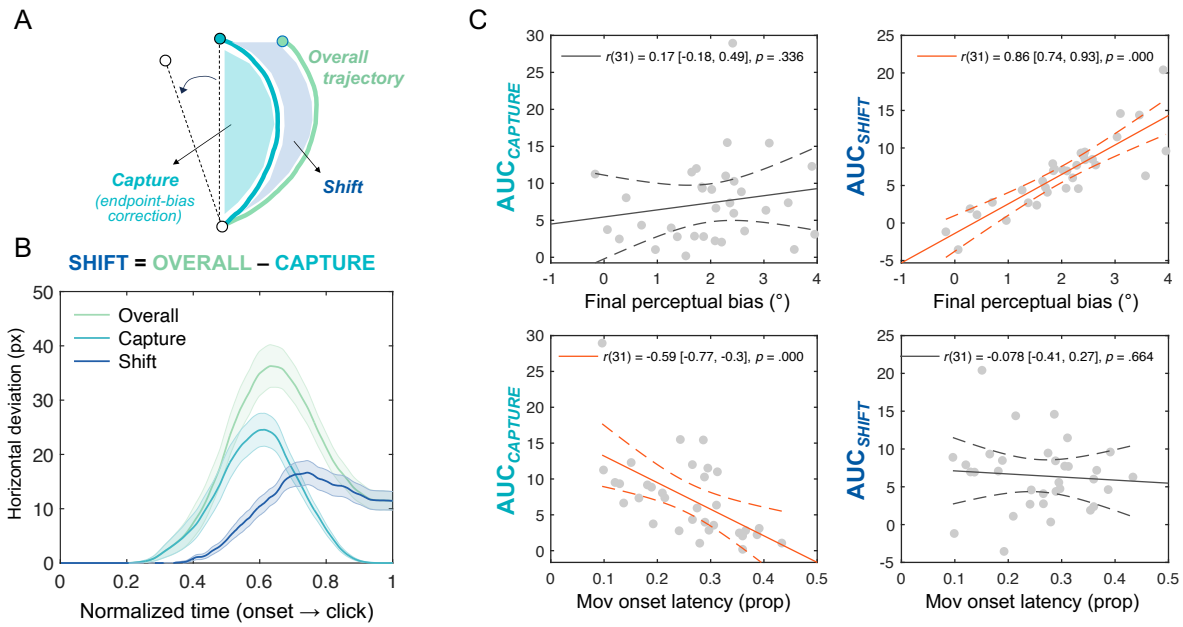

Figure 4. Decomposition of trajectory bias into capture and shift components in Experiment 1. (A) The schematic illustration of the component decomposition. The overall trajectory (green) can be decomposed into a capture component (cyan), reflecting endpoint-bias correction, and a shift component (blue), reflecting residual displacement. (B) The average time course of horizontal deviation. Shaded bands indicate 95% confidence intervals. (C) Correlations of different types of area-under-the-curve (AUC) measures with final perceptual bias and movement onset latency. The solid and dashed lines represent the best linear fit and 95% confidence intervals, respectively, with red lines indicate statistically significant effects.

and returned toward baseline before the response was completed, consistent with transient attentional capture. The shift component (DV_{SHIFT}) remained positive throughout the trajectory, emerging early and persisting until the response, consistent with a steady representational shift.

To directly compare these temporal dynamics, we examined slope differences between the capture and shift effects. Cluster-based permutation testing (10,000 iterations) revealed an early interval (5% to 15% normalized time) where capture increased more steeply than shift (i.e., capture slope > shift slope, $p_{cluster} < .05$), followed by a late period (25% to 95% normalized time) where the difference reversed in sign (i.e., capture slope < shift slope, $p_{cluster} < .05$; Figure 5C). These results indicate that WM exerts dissociable influences on decision dynamics. A rapid

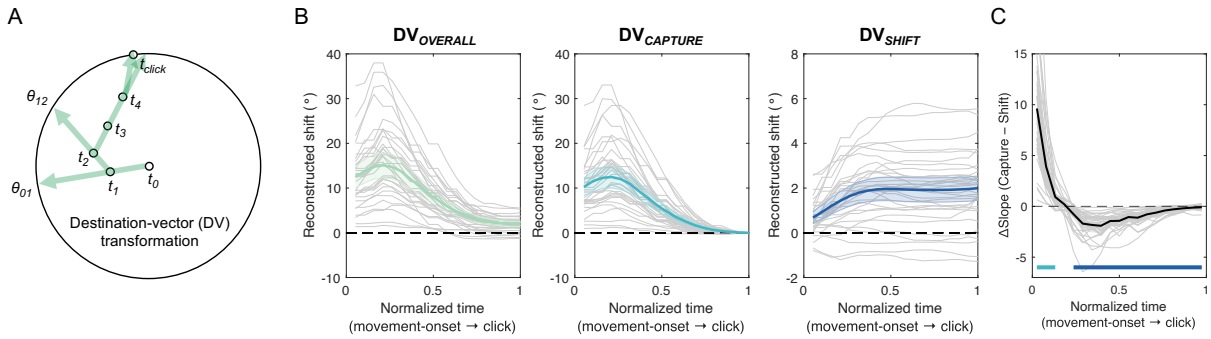

Figure 5. Destination-vector transformation and decomposition analyses in Experiment 1. (A) Schematic illustration of the destination-vector (DV) transformation method, which represents each movement segment as an angle (θ) relative to the target based on two consecutive moment points, allowing reconstruction of intended feature values across time (Park & Zhang, 2024). (B) Reconstructed DV dynamics for overall deviation (green), capture (cyan), and shift (blue). Thin gray lines show individual participant DV series and thick lines with shaded bands indicate group means and 95% confidence intervals. (C) Slope differences between capture and shift DV series. Capture showed a steeper early slope, whereas shift dominated later in the movement. Colored horizontal bars denote time windows with significant cluster-based permutation effects ($p < .05$).

capture process dominates early in the movement and a slower, sustained shift persists throughout and aligns with the endpoint bias.

Experiment 1: Working Memory Change Detection

The WM change-detection task involved systematic manipulations of change magnitude, varying memory-to-probe color offset every 60° steps. To test whether perceptual matching influenced WM content retrospectively, we modeled the psychometric function relating the probability of "same" responses as a function of change magnitude, where 0° represents no-change trials (Figure 6A). The group-aggregated curves were fit with the Gaussian function to estimate the point of subjective equality (PSE, μ_{Gauss}), slope (σ), amplitude (*Amp*), and *lapse* rate, separately for CW and CCW curves. This revealed a reliable shift of the PSE (μ_{Gauss}) between CW (+2.59°) and CCW (−6.48°) conditions. For statistical inference, subject-level bootstrap resampling (10,000 iterations) provided $\text{CI}_{95\%}$ and two-sided p values for the differences in these estimates across the conditions. The PSE (μ_{Gauss}) showed a reliable condition difference between

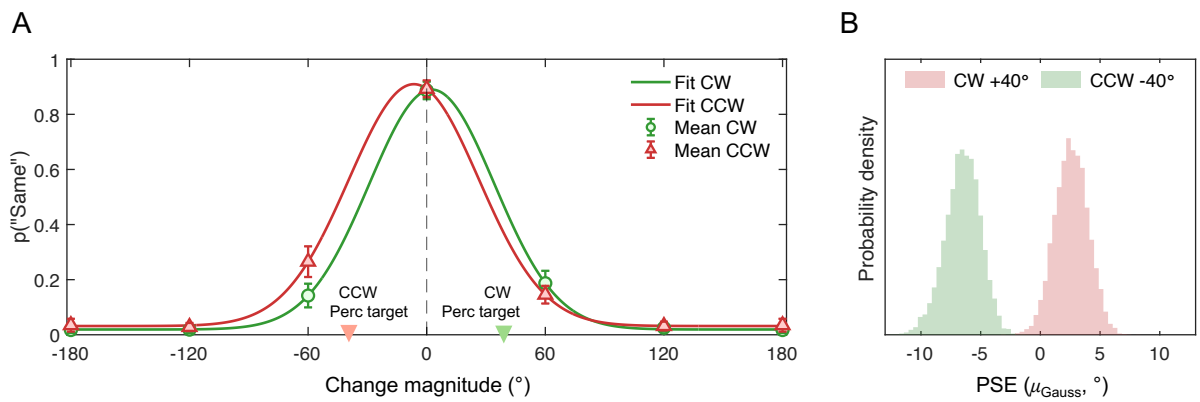

Figure 6. Working memory change-detection performance in Experiment 1. (A) Proportion of "same" responses as a function of change magnitude, with Gaussian fit curves (solid lines) shown separately for CW (+40°) and CCW (−40°) trials. Markers with error bars indicate group means and 95% confidence intervals. Inverted triangles along the x-axis denote the corresponding locations of the preceded perceptual target features. (B) Posterior distributions of the point of subjective equality (PSE, μ_{Gauss}) estimated separately for CW and CCW conditions.

the CW and CCW condition (9.11° [$CI_{95\%}$: 4.67°, 13.42°], $p_{boot} < .001$). None of other parameters showed reliable differences (for *Amp*, −0.01 [$CI_{95\%}$: -0.03, 0.02], $p_{boot} = .529$; for σ, −1.25° [$CI_{95\%}$: −4.12°, 1.68°], $p_{boot} = .384$; for *lapse*, −0.012 [$CI_{95\%}$: −0.026, 0.001], $p_{boot} = .071$).

The change-detection results demonstrate that WM judgments are reliably biased in the direction of the preceding perceptual decision. Experiment 2 replaced a binary change-detection with a continuous WM recall procedure following the perceptual matching task (Figure 1B), allowing a direct comparison between attraction observed during perceptual matching and that during mnemonic recall.

Experiment 2

Endpoint Bias between Perceptual Matching and Working Memory

Hierarchical Bayesian Mixture modeling of continuous-report errors provided converging evidence for reciprocal attraction between WM and perception (Figure 7). Posterior distributions of the μ parameter revealed opposite biases across the responses. During perceptual matching, μ

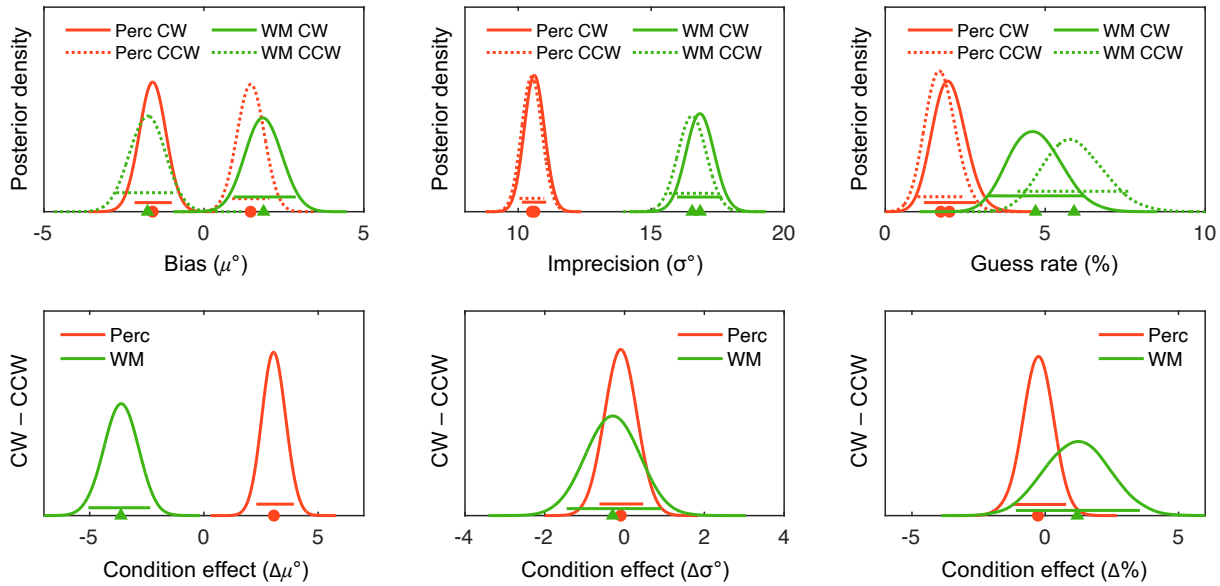

Figure 7. Hierarchical Bayesian Mixture modeling of continuous-report errors in Experiment 2. (Top row) Posterior distributions of population-level parameters for perceptual matching (Perc; red) and working memory recall (WM; green), shown separately for clockwise (CW; solid lines) and counterclockwise (CCW; dotted lines) conditions. (Bottom row) Posterior distributions of condition effects (CW − CCW) for each parameter. Markers and horizontal bars along the x-axis represent posterior means and 95% highest density intervals, respectively.

shifted toward the WM color, replicating Experiment 1 ($\Delta\mu_{CCW-CW}$ = 3.06° [$HDI_{95\%}$: 2.29°, 3.94°]). Critically, during subsequent WM recall, μ shifted in the opposite direction, toward the just-seen perceptual target ($\Delta\mu_{CCW-CW}$ = −3.63° [$HDI_{95\%}$: −5.06°, −2.34°]). Although opposite in sign, their absolute magnitudes were similar, indicating reciprocal influences of comparable strength. By contrast, imprecision (σ) did not differ credibly between CW and CCW trials in either phase (perceptual matching: $\Delta\sigma$ = −0.10° [$HDI_{95\%}$: −0.63°, 0.47°]; WM recall: $\Delta\sigma$ = −0.30° [$HDI_{95\%}$: −1.46°, 0.92°]). Guess rate (g) also remained low and stable (perceptual matching: Δg = −0.27% [$HDI_{95\%}$: −1.27%, 0.79%]; WM recall: Δg = 1.20% [$HDI_{95\%}$: −1.11%, 3.55%]). These results show that WM and perceptual judgments mutually biased one another with comparable magnitudes but in opposite directions, extending the WM-to-perception effect in Experiment 1 to the complementary perception-to-WM influence.

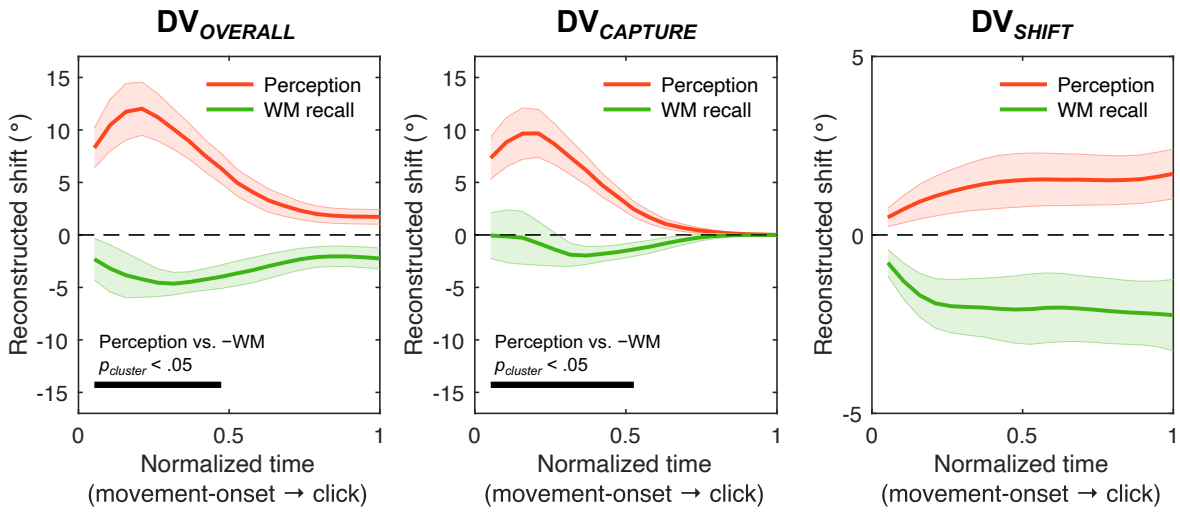

Figure 8. Reconstructed destination-vector (DV) dynamics and decomposition analyses in Experiment 2. Mean DV series with 95% confidence intervals over normalized time for perceptual matching (Perception; red) and working memory (WM) recall (green), shown separately for $DV_{OVERALL}$, $DV_{CAPTURE}$, and DV_{SHIFT} (left to right). Black horizontal bars indicate time intervals with significant cluster-based permutation effects contrasting perceptual matching and sign-flipped WM recall ($p < .05$).

Time-Resolved Decision Dynamics: Destination-Vector Series

To examine how bidirectional biases unfolded during movement, we applied the DVT to mouse trajectories for perceptual matching and WM recall, and decomposed them into the capture ($DV_{CAPTURE}$) and shift (DV_{SHIFT}) components that showed dissociable temporal dynamics for the two responses (Figure 8). Perceptual matching showed strong early deviations toward the WM color, whereas WM recall exhibited minimal early curvature. This asymmetry was primarily attributed to the capture component ($DV_{CAPTURE}$). During perceptual matching, capture rose steeply after the movement onset and declined thereafter, whereas in WM recall it was effectively minimal. In contrast, the shift component (DV_{SHIFT}) showed steadily increasing amplitude across time, but with opposite directions for the two responses that mirrors the endpoint bias.

Cluster-based permutation tests (10,000 label-swaps within subject) on the reconstructed destination-vectors confirmed this dissociation. This analysis identified a reliable early interval for capture in which perceptual matching exceeded WM recall, spanning from 5% to 55% of normalized time from movement onset to final click ($p_{cluster} < .05$), consistent with capture being expressed only during perceptual matching. For shift, no clusters survived correction (all $p_{cluster} > .05$), indicating comparable sustained bias across phases.

Together, these results demonstrate the reciprocal but asymmetrical interactions between perceptual matching and WM recall. Specifically, perceptual matching engages both capture and shift dynamics, whereas WM recall primarily shows a sustained shift, with little capture by the previous perceptual input. This dissociation suggests that WM-based attention capture for perceptual matching may likely result from the prioritization of WM contents (e.g., for maintenance and protection from interference by incoming perceptual inputs).

General Discussion

The present study examined how WM shapes perceptual appearance in relation to its established role in guiding attentional selection (see Tables 1). Across two experiments, we found robust bidirectional attraction effects between WM and perception. Perceptual judgments were systematically shifted toward the feature value held in WM, and subsequent WM recall was biased toward the preceding perceptual decision. Our findings provide time-resolved behavioral evidence that WM's top-down influence is not a single process. Instead, it consists of two temporally and mechanistically distinct components: a fast, transient attentional capture and a slow, sustained representational shift.

Three converging findings support this conclusion. First, decomposition of mouse trajectories revealed a clear double dissociation. The magnitude of early, transient curvature

(capture) correlated with how quickly a response was initiated but not with final bias. In contrast, the sustained curvature (shift) correlated strongly with final bias but not with response initiation speed. Second, time-resolved destination-vectors analyses confirmed distinct temporal patterns. The capture component peaked sharply near movement onset and quickly decayed, whereas the shift component emerged more gradually and persisted until the response was completed. Third, and most importantly, the two processes were functionally distinct across directions of influence. When WM prospectively influenced perception, both capture and shift were expressed. When perception reciprocally influenced WM, only the sustained shift was evident. This asymmetry implies different underlying mechanisms for prospective and retrospective interaction.

The fast-acting capture component can be understood as the behavioral signature of attentional guidance. It represents the transient top-down priority signal that WM generates to bias competition toward matching features. This interpretation is consistent with the biased competition model, which posits that active WM representations pre-activate corresponding feature-selective neural populations, giving any matching sensory input a competitive advantage for selection (Desimone & Duncan, 1995; Serences & Yantis, 2006). The initial deviation of the cursor toward the WM-matching color follows from this early priority signal as the system momentarily resolves competition between the target and the memory-matching feature value. The transient rise and subsequent correction are consistent with models in which attention can be captured by a salient or goal-relevant item but then quickly disengaged to reorient toward the primary task goal (Dieciuc et al., 2019). Continuous trajectory analysis therefore provides a direct behavioral measure of how this capture-and-disengage sequence unfolds within a single decision.

The asymmetry observed in Experiment 2 further clarifies under which conditions capture occurs. Capture was evident when the WM representation served a prospective role as an attentional template for the upcoming perceptual decision, but it was absent when the influence was retrospective (i.e., from a just-encoded percept back onto memory; Park et al., 2017). This suggests that capture is not an automatic consequence of any mnemonic representation but specifically reflects the priority state of WM (Park & Zhang, 2024). Neuroimaging studies support this distinction, showing that prospective WM templates engage frontoparietal networks that broadcast priority signals to sensory areas, whereas retrospective maintenance primarily involves sensory reinstatement without prioritization (Myers et al., 2017; van Loon et al., 2018). In this sense, a WM template that is actively guiding behavior is functionally distinct from a passive trace of previous percept, and only the former produces the rapid priority signal that manifests as transient motor capture toward memory-matching input.

This asymmetry arises from task relevance and representational state. During perceptual matching, the WM template remains active and relevant, supporting both transient capture and sustained appearance shift. During WM recall, the previous percept is no longer relevant and the WM trace becomes noisier, making judgments rely more on recent sensory priors. The similar magnitudes of the opposite biases suggest a unified mechanism of assimilation rather than task-specific artifacts. The dissociation between the two components strengthens this interpretation, as early capture covaried with movement initiation latency but not endpoint bias, whereas sustained shift tracked endpoint bias independently of movement initiation latency. A single decisional or motor explanation would predict shared variance across these measures and can therefore be ruled out. These patterns support a precision-weighted integration framework in

which perceptual readout combines current sensory evidence with state-dependent priors supplied by WM.

In contrast, the slow and sustained shift component provides strong evidence for a genuine change in the underlying perceptual representation itself. A key challenge in the debate over cognitive penetrability is distinguishing true perceptual changes from effects of attentional selection or post-perceptual decision processes (Firestone & Scholl, 2016). Our trajectory decomposition directly addresses this issue by isolating a sustained shift component that is independent of the initial capture effect and that directly predicts the final perceptual report. The sustained shift therefore is more consistent with a fundamental alteration in perceptual appearance rather than a decision-level artifact. From a broader theoretical perspective, the present results support a state-dependent framework in which memory acts as a flexible prior for perceptual readout. WM thus belongs to a broader class of mnemonic influences that configure perception through dynamic, context-dependent processing of environmental regularities (Girshick et al., 2011), semantic knowledge (Hansen et al., 2006; Bannert & Bartels, 2013), and episodic retrieval (Bosch et al., 2014; Won et al., 2023).

The integrative and gradual nature of this shift aligns with predictive coding and Bayesian integration frameworks of perception (Ma et al., 2006; Knill & Richards, 1996). In these frameworks, perception arises from the continuous integration of sensory evidence with top-down expectations. In this context, the WM representation acts as a prior belief about which features are most likely to occur in the environment. The brain combines this prior with incoming sensory evidence (the likelihood) to form an updated percept (the posterior) that is consequently biased toward the prior. Because this computation relies on evidence accumulation, it unfolds gradually and is slower than the winner-take-all competition that characterizes

attentional selection. Within this framework, WM exerts two forms of influence on perception that differ in both timescale and computational goal, one through competitive selection where the online activation of feature codes facilitate attentional access and action, and the other through predictive integration where the mnemonic content acts as a prior that shapes perceptual inference.

In conclusion, this study demonstrates that WM functions as both a guide for attention and a lens for perception, with each role supported by distinct mechanisms operating with different temporal dynamics. A rapid priority signal produces brief capture when a template is prospectively relevant, while a slower assimilative process alters perceived appearance through predictive integration. Methodologically, by leveraging the temporal resolution of mouse-tracking, the present approach provides a process-level decomposition of decision dynamics that is sensitive to transient capture and to sustained representational change within the same trial. The approach aligns with efforts to expand the rich behavioral repertoire for studying dynamic cognition (Nobre, 2022). Together, these findings provide a dynamic and nuanced view of how internal goals shape our experience of the external world. Specifically, cognition emerges as a continuous synthesis of past and present, with WM acting as a goal-directed prior that shapes what we attend to and how the world is perceived. This process allows perception to achieve both stability and flexibility, optimizing adaptive behavior in a complex environment (Kiyonaga et al., 2017).

Conflicts of interest: All authors declare no conflicts of interest.

Funding: This research was supported by the National Institute of Mental Health (1R01MH117132).

Ethics: This research was approved by the Institutional Review Board of the University of California, Riverside.

Preregistration: No aspects of the study were preregistered

Data: All primary and processed data, and custom codes are publicly available via the Open Science Framework repository (https://osf.io/3pyjh).

References

Bannert, M. M., & Bartels, A. (2013). Decoding the yellow of a gray banana. *Current Biology, 23*(22), 2268-2272.

Barbosa, J., & Compte, A. (2020). Build-up of serial dependence in color working memory. *Scientific Reports, 10*(1), 10959.

Bosch, S. E., Jehee, J. F., Fernández, G., & Doeller, C. F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *Journal of Neuroscience, 34*(22), 7493-7500.

van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *eLife, 7*, e38677.

Burak, Y., & Fiete, I. R. (2009). Accurate path integration in continuous attractor network models of grid cells. *PLoS Computational Biology, 5*(2), e1000291.

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.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193-222.

Dieciuc, M. A., Roque, N. A., & Boot, W. R. (2019). The spatial dynamics of mouse-tracking reveal that attention capture is stimulus-driven rather than contingent upon top-down goals. *Journal of Experimental Psychology: Human Perception and Performance, 45*(10), 1285-1290.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Method, 41*(4), 1149-1160.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences, 39*, e229.

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience, 17*(5), 738-743.

Frischkorn, G. T., & Popov, V. (2025). A tutorial for estimating Bayesian hierarchical mixture models for visual working memory tasks: Introducing the Bayesian Measurement Modeling (bmm) package for R. *Behavior Research Methods, 57*(5), 144.

Friston, K. (2018). Does predictive coding have a future?. *Nature Neuroscience, 21*(8), 1019-1021.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457-472.

van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *eLife, 7*, e38677.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience, 9*(11), 1367-1368.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*(7238), 632-635.

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.

Knill, D. C., & Richards, W. (Eds.). (1996). *Perception as Bayesian inference*. Cambridge University Press.

Kruschke, J.K. (2015). *Doing Bayesian data analysis, Second Edition: A tutorial with R, JAGS, and Stan*. Burlington, MA: Academic Press

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4*(11), 432-440.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience, 9*(11), 1432-1438.

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing information during working memory: Beyond sustained internal attention. *Trends in Cognitive Sciences, 21*(6), 449-461.

Nobre, A. C. (2022). Opening questions in visual working memory. *Journal of Cognitive Neuroscience*, 35(1), 49-59.

Nobre, A. C., & Stokes, M. G. (2019). Premembering experience: A hierarchy of time-scales for proactive attention. *Neuron, 104*(1), 132-146.

Oberauer, K., Stoneking, C., Wabersich, D., & Lin, H. Y. (2017). Hierarchical Bayesian measurement models for continuous reproduction of visual features from working memory. *Journal of Vision, 17*(5), 1-27.

Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1243-1265.

Park, H. B. (2025). Process dynamics of serial biases in visual perception and working memory processes. *Psychonomic Bulletin & Review*, Advance online publication.

Park, H., & Zhang, W. (2024). The dynamics of attentional guidance by working memory contents. *Cognition, 242*, 105638.

Park, H., & Zhang, W. (2022). Trial-by-trial mouse trajectory predicts variance in precision across working memory representations: A critical reanalysis of Hao et al. (2021). *Psychonomic Bulletin & Review, 29*(6), 2181-2191.

Park, H. B., Zhang, W., & Hyun, J. S. (2017). The aftermath of memory retrieval for recycling visual working memory representations. *Attention, Perception, & Psychophysics, 79*(5), 1393-1407.

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.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20*(2), 207-214.

Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences, 10*(1), 38-45.

Shiffrin, R. M., Lee, M. D., Kim, W., & Wagenmakers, E. J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science, 32*(8), 1248-1284.

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences, 12*(9), 342-348.

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience, 15*(11), 745-756.

Tan, J., Zhao, Y., Wu, S., Wang, L., Hitchman, G., Tian, X., ... & Chen, A. (2014). The temporal dynamics of visual working memory guidance of selective attention. *Frontiers in Behavioral Neuroscience, 8*, 345.

Teng, C., Kaplan, S. M., Shomstein, S., & Kravitz, D. J. (2023). Assessing the interaction between working memory and perception through time. *Attention, Perception, & Psychophysics, 85*(7), 2196-2209.

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour, 3*(8), 827-836.

Theeuwes, J., Olivers, C. N., & Chizk, C. L. (2005). Remembering a location makes the eyes curve away. *Psychological Science, 16*(3), 196-199.

van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *eLife, 7*, e38677.

Wimmer, K., Nykamp, D. Q., Constantinidis, C., & Compte, A. (2014). Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience, 17*(3), 431-439.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202-238.

Won, B. Y., & Zhang, W. (2025). Search strategy modulates memory-driven capture. *Attention, Perception, & Psychophysics*, Advance online publication.

Won, B., Park, H., & Zhang, W. (2023). Familiarity enhances mnemonic precision but impairs mnemonic accuracy in visual working memory. *Psychonomic Bulletin & Review, 30*. 1452-1462.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233-235.

Table 1

Distinct Temporal and Mechanistic Components of Working Memory's Top-Down Influence

Component	Attentional Capture	Representational Shift
Cognitive Function	Attentional guidance, priority signal	Perceptual bias, evidence integration
Behavioral Signature	Transient trajectory curvature ($AUC_{capture}$)	Sustained trajectory curvature (AUC_{shift}), endpoint bias
Temporal Dynamics	Fast, peaks early in movement and then decays rapidly	Slow, emerges early and remains sustained throughout movement
Key Correlate	Movement onset latency	Magnitude of final perceptual bias
Task Condition	Present in WM → Perception (prospective)	Present in WM ↔ Perception (bidirectional)
Mechanisms	Biased competition, top-down priority signaling	Bayesian integration, predictive coding (WM as Prior)