Behavioral/Cognitive

Common Neural Mechanisms Control Attention and Working Memory

Ying Zhou (周颖),[1,2] Clayton E. Curtis,[2,3] Kartik K. Sreenivasan,[1,2,4*] and Daryl Fougnie[1,2*]

[1]Program in Psychology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates 129188, [2]Department of Psychology, New York University, New York, New York 10003, [3]Center for Neural Science, New York University, New York, New York 10003, and [4]Program in Biology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates 129188

Although previous studies point to qualitative similarities between working memory (WM) and attention, the degree to which these two constructs rely on shared neural mechanisms remains unknown. Focusing on one such potentially shared mechanism, we tested the hypothesis that selecting an item within WM utilizes similar neural mechanisms as selecting a visible item via a shift of attention. We used fMRI and machine learning to decode both the selection among items visually available and the selection among items stored in WM in human subjects (both sexes). Patterns of activity in visual, parietal, and to a lesser extent frontal cortex predicted the locations of the selected items. Critically, these patterns were strikingly interchangeable; classifiers trained on data during attentional selection predicted selection from WM, and classifiers trained on data during selection from memory predicted attentional selection. Using models of voxel receptive fields, we visualized topographic population activity that revealed gain enhancements at the locations of the externally and internally selected items. Our results suggest that selecting among perceived items and selecting among items in WM share a common mechanism. This common mechanism, analogous to a shift of spatial attention, controls the relative gains of neural populations that encode behaviorally relevant information.

Key words: attention; decoding; fMRI; selection; working memory

---

Significance Statement

How we allocate our attention to external stimuli that we see and to internal representations of stimuli stored in memory might rely on a common mechanism. Supporting this hypothesis, we demonstrated that not only could patterns of human brain activity predict which items were selected during perception and memory, but that these patterns were interchangeable during external and internal selection. Additionally, this generalized selection mechanism operates by changes in the gains of the neural populations both encoding attended sensory representations and storing relevant memory representations.

Introduction

Although theories of attention and working memory (WM) often emphasize their interrelatedness (Cowan, 1998a; Awh and Jonides, 2001; Oberauer, 2002; Chun, 2011; Gazzaley and Nobre, 2012), they are typically studied in isolation. Several lines of

empirical evidence highlight commonalities between attention and WM. For example, attention and WM share similar capacity and resource limitations (Cowan, 1998b; Marois and Ivanoff, 2005; but see Fougnie and Marois, 2006) and engage similar brain regions (LaBar et al., 1999; Pollmann and von Cramon, 2000; Awh and Jonides, 2001; Ranganath and D'Esposito, 2005; Ikkai and Curtis, 2011; Jerde et al., 2012). Moreover, the process of rehearsing items (Awh et al., 1999, 2000; Awh and Jonides, 2001; Jha, 2002; Postle et al., 2004; Theeuwes et al., 2005; Katus et al., 2014; Shen et al., 2015; Souza et al., 2020) and suppressing distracting information (Gazzaley et al., 2005; Sreenivasan and Jha, 2007) in WM may be attention-based. In turn, representations in WM can guide how we attend to sensory information (Olivers et al., 2006; Woodman and Luck, 2007; Gayet et al., 2013; Bahle et al., 2020; Sasin and Fougnie, 2020; Williams et al., 2022). However, since attention and WM are complex processes involving multiple cognitive operations, it remains unclear which underlying components may be shared between the two. Here, we focus on the process of selection, which refers to how task-

relevant information is prioritized over task-irrelevant information in both attention and WM. Critically, selection helps mitigate the strict resource or capacity limitations of attention and WM (Marois and Ivanoff, 2005).

The notion that a common process both selects among external sensory information and among internal WM representations is intriguing because it appeals to an intuitive common mechanism used to highlight relevant information. Nonetheless, there is little evidence linking attentional selection and WM selection. Researchers have intensively investigated attentional selection using pre-cueing paradigms, where a cue indicates which forthcoming stimulus to attend (Posner, 1980; Eriksen and Yeh, 1985; Murphy and Eriksen, 1987). Behaviorally, pre-cueing benefits the processing of selected information (Carrasco et al., 2000; Pestilli et al., 2009) at the cost of the processing of unselected information (Pestilli and Carrasco, 2005; Pestilli et al., 2007). Both neurons (Bushnell et al., 1981; Luck et al., 1997; Reynolds et al., 2000; Bisley and Goldberg, 2003) and voxels (Gandhi et al., 1999; Hopfinger et al., 2001; Liu et al., 2005; Serences and Boynton, 2007; Silver et al., 2007; Ikkai and Curtis, 2008) with receptive fields that match the locations of pre-cued items exhibit increased activity relative to those that match the locations of unattended items. Theories of selective attention, including computational models, posit that the benefits of attention stem from gain enhancements within the populations of neurons encoding selected task-relevant stimuli (Reynolds and Heeger, 2009; Carrasco, 2011). To study the process of selection in WM, researchers have used retro-cueing experimental paradigms, where a cue presented after WM items have been encoded signifies which memorandum will later be tested (Griffin and Nobre, 2003). Behaviorally, the quality of memory is better for the retroactively cued item compared with noncued items (Griffin and Nobre, 2003; Landman et al., 2003; Souza et al., 2016; Li et al., 2021). Neurally, WM representations are enhanced within the neural populations encoding the selected WM item (Lepsien et al., 2011; Sprague et al., 2016; Ester et al., 2018; Yoo et al., 2022).

We directly tested the hypothesis that a common neural mechanism underlies attentional and WM selection using fMRI and machine learning. Remarkably, classifiers trained on each type of selection were interchangeable in predicting the other, providing novel quantitative evidence for theories that posit a shared mechanism (Chun et al., 2011). In addition, using models of voxel receptive fields to visualize population activity, we observed elevated responses corresponding to the locations of the externally and internally selected items, suggesting that the shared selection mechanism involves differential gain.

Materials and Methods

Participants

Eleven neurologically healthy participants (ages 23–53; six females) with normal or corrected-to-normal vision participated in this experiment. The sample size was determined using previous fMRI studies comparing selection on perceptual and WM representations (Nobre et al., 2004; Tamber-Rosenau et al., 2011), and is equal to or larger than previous studies that compared within-condition and across-condition decoding performance of classifiers trained on fMRI data (Jerde et al., 2012; Rademaker et al., 2019; Kwak and Curtis, 2022), as well as those that used population receptive field (pRF)-weighted reconstruction analysis (Kwak and Curtis, 2022; Yoo et al., 2022). Participants provided written informed consent in accordance with procedures approved by the Institutional Review Board at New York University.

Experimental design

We generated stimuli and interfaced with the MRI scanner, button box, and eye tracker using MATLAB software (The MathWorks) and Psychophysics Toolbox 3 (Brainard, 1997). Stimuli were presented using a PROPixx DLP LED projector (VPixx) located outside the scanner room and projected through a waveguide and onto a translucent screen located at the head of the scanner bore. Participants viewed the screen at a total viewing distance of 63 cm through a mirror attached to the head coil. The display was a circular aperture with an ~32-degree of visual angle (dva) diameter. A trigger pulse from the scanner synchronized the onsets of stimulus presentation and image acquisition.

Participants performed a pre-cue task and a retro-cue task in the two scanning sessions. The task procedures are illustrated in Figure 1A. The fixation symbol in both tasks was a centrally-presented filled circle with a 0.3-dva radius. Subjects were required to maintain fixation in the center of the screen. Each pre-cue trial began with a 750-ms colored central fixation (0.4-dva radius) with three black placeholders. The color of the fixation indicated the target location in the upcoming stimulus screen. The distance from the screen center to the center of each placeholder was 6 dva, and the diameter of each placeholder was 8 dva. The pre-cue was followed by a 1500-ms ISI (Inter-stimulus interval), then by the stimulus for 1500 ms. Stimulus presentation consisted of three Gabor patches, one in each placeholder. The three placeholders were in three different colors, and subjects had to select the target Gabor in the placeholder with the pre-cued color. The three colors used in each trial were randomly selected from four colors (RGB = [255, 0, 0], [0, 200, 0], [0, 0, 255], [255, 165, 0]) and randomly distributed across the three locations (Left, Right, Bottom) so that the target location could not be predicted by the representations of the pre-cue. The stimulus presentation was followed by a 750-ms mask to diminish iconic memory (Sperling, 1960), then by a 3000-ms delay. This was followed by the presentation of a probe, which consisted of a circle and an oriented line. The length of the line and the diameter of the circle were both 6 dva. Subjects had to judge if the line was rotated clockwise or counterclockwise compared with the orientation of the selected Gabor. We adjusted the difference between the orientations of target and probe to titrate the behavioral performance to ~80%. Specifically, the difference between the orientations of the probe and target started at 20° and either increased by 1° after each error trial or decreased by 1° after four continuous correct trials (cf. Levitt, 1971). Subjects responded by pressing "1" for clockwise or "2" for counterclockwise. The probe screen lasted for 2250 ms regardless of subjects' responses. Subjects then received feedback consisting of the selected Gabor overlaid with the probe. The color of the probe indicated whether the response was correct (green: correct; red: incorrect). The intertrial interval lasted for 9750 ms. Each retro-cue trial began with a 1500-ms stimulus screen, which contained three Gabor patches in three black placeholders. The stimulus was followed by a 750-ms mask and a 1500-ms ISI. Following the ISI, subjects saw a retro-cue consisting of three colored placeholders surrounding a colored fixation. The color of the fixation point matched the color of one of the placeholders and indicated the location of the target Gabor. The colors on each trial were randomly selected from the four possible colors (see above) and randomly distributed across the three locations. The delay, probe, feedback, and intertrial interval were the same as those in the pre-cue task. Each subject completed two scanning sessions consisting of 10 runs each, with five pre-cue runs and five retro-cue runs presented in an interleaved order. Each run contained 18 trials, yielding 180 trials per condition (90 per session). Each run started with 13 dummy TRs (9750 ms) of a central fixation screen to allow for magnetic field stabilization.

Eye movements

Eye position coordinates (x,y) and pupil size were recorded at 500 Hz in the scanner using an EyeLink 2k (SR Research). Before each scanning session, eye position was calibrated using a 9-point calibration. Eye data were preprocessed in MATLAB using the freely available iEye toolbox (https://github.com/clayspacelab/iEye) using the following steps. (1) Data were transformed from raw pixel screen coordinates into dva. (2) Extreme values associated with loss of track and blinks were removed. (3) Data were smoothed with a Gaussian kernel (5-ms SD). (4) Each trial
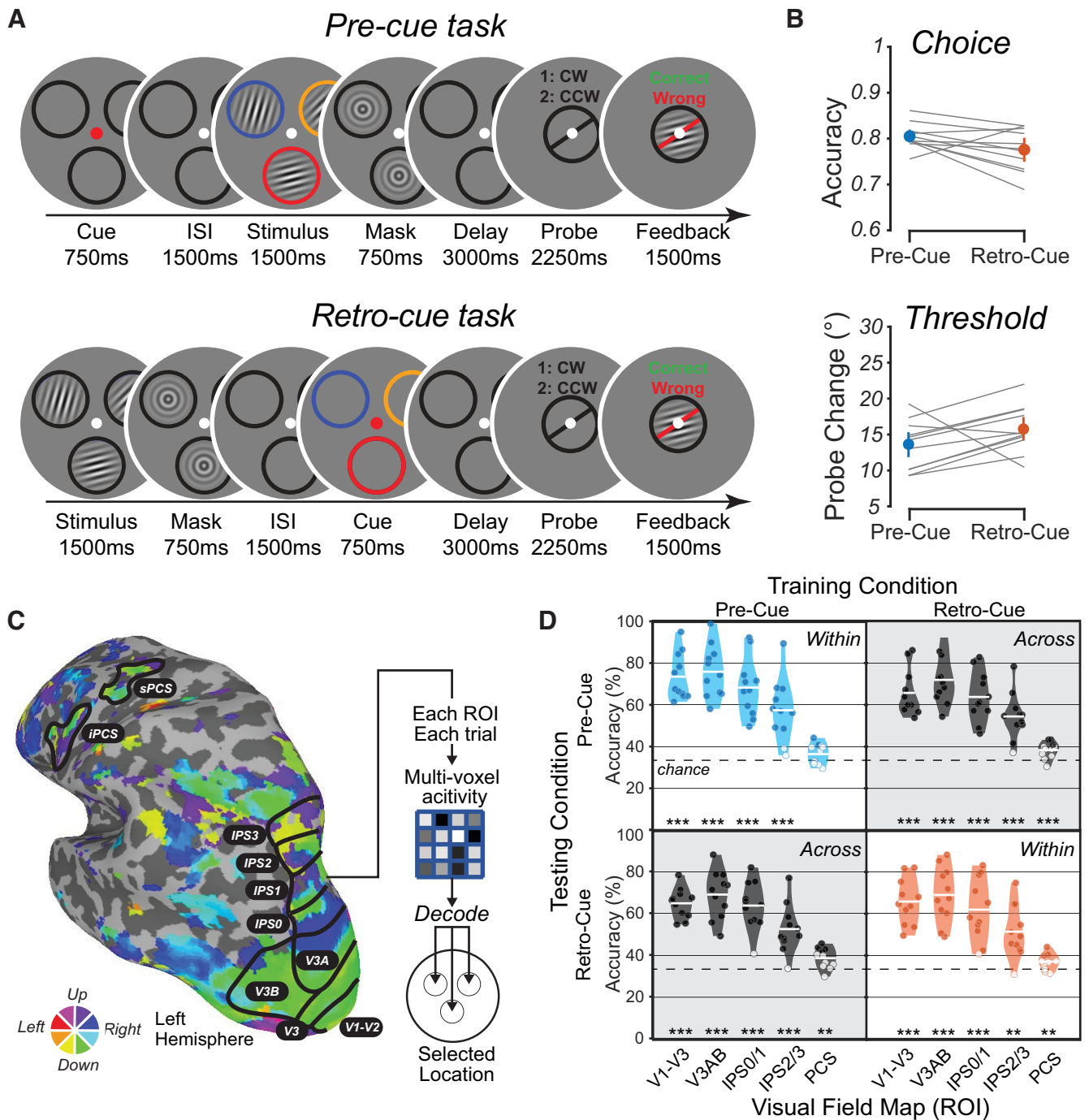
Figure 1. Methods, behavior, decoding analysis, and decoding results. ***A***, Pre-cue task: fixation cued the task-relevant color. Then, three oriented Gabors appeared within a colored placeholder ring. Participants encoded the Gabor whose ring color matched fixation. After a mask and delay, participants reported whether the orientation of the probe was tilted clockwise (CW; respond with button 1) or counterclockwise (CCW; respond with button 2) to the memorized target orientation. Retro-cue task: participants encoded the orientations of three Gabors that were followed by a mask and short ISI. Then a retro-cue indicated the task-relevant color. Participants were instructed to remember the orientation of the item at the location matching the fixation color and later compare to the probe orientation. ***B***, Probe task accuracy and probe change magnitude in pre-cue (blue) and retro-cue (red) conditions. Gray lines indicate individual subjects. The error bars denote 95% bootstrap confidence interval (CI). ***C***, Retinotopy for a representative subject. Color depicts voxels' preferred polar angle projected to the cortical surface of the left hemisphere. ***D***, Decoding performance across ROIs, for within-condition decoding of pre-cued location (upper left), retro-cued location (lower right), and cross-condition decoding of location (gray backgrounds). Gray dashed line indicates chance accuracy (33.3%; colored: individual performance was significantly higher than chance, $p < 0.05$; one-sided, FDR false detection rate (FDR) corrected across all ROIs for each decoding condition independently). The jittered dots indicate individual performance, filled dots indicate participants from whom decoding performance was higher than chance ($p < 0.05$, one-sided), and unfilled dots indicate participants with at-chance decoding performance. The horizontal line in the violin indicates the mean of accuracy. The stars in the bottom of each plot indicate the group-level performance is significantly higher than chance (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$; one-sided, FDR corrected across all ROIs for each decoding condition independently).

was drift-corrected by taking the mean over known epochs when the participant was fixating (the whole delay in pre-cue task, and the first delay before the cue in the retro-cue task) and subtracting that value from the entire trial.

acquired at the NYU Center for Brain Imaging on a 3T Siemens Prisma Scanner with a 64-channel head/neck coil. Functional scans for the selection experiment were acquired using an EPI pulse sequence with 44 slices and a voxel size of 2.53 mm (4× simultaneous-multi-slice acceleration; FoV = 200 × 200 mm, no in-plane acceleration, TE/TR = 30/750 ms, flip angle = 50°, bandwidth = 2290 Hz/pixel, 0.56-ms echo spacing, P → A phase encoding). Data for the retinotopic mapping session was acquired in a separate session at a higher resolution, with 56 slices and a voxel size of 2.03 mm (4× simultaneous multi-slice acceleration, FoV = 208 × 208 mm, no in-plane acceleration, TE/TR = 36/1200 ms, flip angle = 66°, bandwidth = 2604 Hz/pixel, 0.51 ms echo spacing, P → A phase encoding). To correct for local spatial distortions in the functional data offline, we estimated a field map of the field inhomogeneities by acquiring pairs of spin echo images with normal and reversed phase-encoding directions with an identical slice prescription to the functional data and no simultaneous-multi-slice acceleration (TE/TR = 45.6/3537 ms, three volumes per phase encoding direction). To enable precise localization of functional data, we collected T1-weighted whole-brain anatomic scans using a MP-RAGE sequence with 192 slices and a voxel size of 0.83 mm (FoV = 256 × 240 mm, TE/TR = 2.24/2400 ms, flip angle = 8°, bandwidth = 210 Hz/pixel) for each participant.

MRI data preprocessing
T1-weighted anatomic images were segmented and cortical surfaces were constructed using the recon_all command in Freesurfer (version 6.0). Functional data were preprocessed with custom scripts using functions provided by AFNI. First, we applied the B_0 field map correction and reverse-polarity phase-encoding (reverse blip) correction. Next, we performed motion correction using a six-parameter affine transform, aligned the functional data to the anatomic images, and projected the data to the cortical surface. Spatial smoothing (5-mm FWHM on the cortical surface) was applied to the retinotopic mapping data, but no explicit smoothing was applied to the data from the selection experiment. Data from the selection experiment was re-projected into volume space, which incurs a small amount of smoothing along vectors perpendicular to the cortical surface. Finally, we removed linear trends from the time series data, and then normalized (z score) across all the time points within each run.

Estimating selection-related blood oxygen level-dependent (BOLD) activity
To identify activity related to selection in both our tasks, we used the 3dDeconvolve and 3dLSS commands in AFNI (https://afni.nimh.nih. gov/) to implement a least-squares-separate general linear model (LSS-GLM) to the preprocessed BOLD time series of the functional data. LSS-GLM has been shown to isolate single-trial activity in rapid event-related designs (Mumford et al., 2012; Arco et al., 2018). In the pre-cue condition, we modeled the selection, delay, and response events. The stimulus event was not modeled because it overlapped with the selection event. Based on previous estimates that symbolic central cues direct selection within 300 ms (Carrasco, 2011), we set the duration of the selection event equal to one TR (750 ms). The retro-cue model was similar to the pre-cue model, except the selection event was time-locked to the 750-ms retro-cue screen. Each event was modeled as a boxcar with the duration of the event convolved with a hemodynamic response function (HRF; GAM(10.9, 0.54) in 3dDeconvolve). The β coefficients for the selection event in each trial were estimated separately, resulting in 180 GLM iterations for each task condition. In each iteration, we modeled the selection event on the trial of interest with a single regressor and the selection events on all other trials with a separate regressor. The delay and response events were modeled with one regressor each. Thus, the GLM for each task condition included four regressors on each iteration (selection for the trial of interest, selection for all other trials, delay, and response). In addition, the model included six regressors for head motion and four regressors for data drift. The procedure was repeated with each trial in turn serving as the trial of interest, resulting in 180 selection betas per condition. Selection betas were normalized via z-scoring on a voxel-by-voxel and run-by-run basis before further analysis.

pRF mapping
pRF mapping was conducted using the procedures described previously (Mackey et al., 2017). Participants maintained central fixation while covertly tracking a bar aperture that swept across the screen in discrete steps in one of four orientations/directions: oriented vertically and traveling from left to right or right to left, or oriented horizontally and traveling from top to bottom or bottom to top. The bar aperture was divided into three rectangular segments (a central segment and two flanking segments) of equal sizes, each containing a random dot kinematogram (RDK). Participants' task was to identify which of the two flanking RDKs had the same direction of motion as the central RDK. The dot motions of all the three segments changed with each discrete step. Participants reported their answer with a button press. We adjusted the coherence of the RDK using a staircase procedure to maintain accuracy at ~75%. Each session contained eight to nine runs, each 5-min run consisted of 12 sweeps, and each sweep consisted of 12 discrete steps (one step every 2 s). The order of the four sweep directions was randomized within each run.

We fit the preprocessed BOLD time series for each voxel for each participant using the compressive spatial summation model (Kay et al., 2013):

$$\hat{r}(t) = \gamma \left[\iint S(x,y) N((x,y), \mathbf{I}\sigma) dxxy \right]^n,$$

where S is a binary stimulus image (1 in where the stimulus was presented and 0 in where the stimulus was not presented), $N((x,y), \mathbf{I}\sigma)$ is a normal distribution with mean (x, y) and variance $\mathbf{I}\sigma^2$, where $\mathbf{I}$ is a two-dimensional identity matrix describing a circular, symmetric Gaussian. The parameters of this model are the voxel's receptive field center (x, y) in dva, SD σ in dva, amplitude γ, and compressive spatial summation factor n (where $n \leq 1$). Parameters were fit with a GPU-accelerated course grid search over parameters, followed by a local optimization method.

Voxels' preferred phase angle ($\arctan(\frac{x}{y})$) and eccentricity ($\sqrt{x^2 + y^2}$) were visualized on the cortical surface. To define retinotopically organized regions of interest (ROIs), we restricted our analysis to voxels with >10% response variability explained by the pRF model. We then drew ROIs on each participant's cortical surface using reversals of the voxels' preferred phase angle as boundaries between neighboring visual regions (Mackey et al., 2017). We defined four ROIs in anterior and dorsal visual areas (V1–V3AB), four ROIs along the caudal–rostral intraparietal sulcus (IPS0–IPS3), and two ROIs along the dorsal-ventral precentral sulcus region (sPCS and iPCS), each with a full visual field representation (for an illustration of the ROIs for one subject, see Fig. 2A). Further fMRI analyses were conducted in ROIs that combined regions which share a foveal confluence (V1, V2, and V3; IPS0 and IPS1; IPS2 and IPS3; Wandell et al., 2005, 2007; Mackey et al., 2017). We also combined voxels in sPCS and iPCS into a single PCS ROI to roughly match the size of our other ROIs, although the results were comparable for the individual PCS ROIs.

Decoding analyses
We examined whether the machine learning algorithm trained by selection-related activity in the pre-cue task can decode the internally selected location in the retro-cue task, and vice versa. To decode the selected location, we trained sparse multinomial logistic regression (SMLR) classifiers in MATLAB using the Princeton MVPA toolbox (http://www.csbmb.princeton.edu/mvpa). The SMLR classifier is widely used to decode multiclass conditions (Krishnapuram et al., 2005; Pereira et al., 2009). The z-scored β coefficients for the selection event in each trial estimated using the LSS-GLM were used to train the classifiers. First, the classifiers were tested on the dataset in the same task condition as training (i.e., within-condition decoding) to examine whether the neural representations we extracted could predict the selected location (Fig. 1C). Since there were three possible locations, we used a leave-three-trials-out cross-validation scheme, in which the classifier was trained on the data
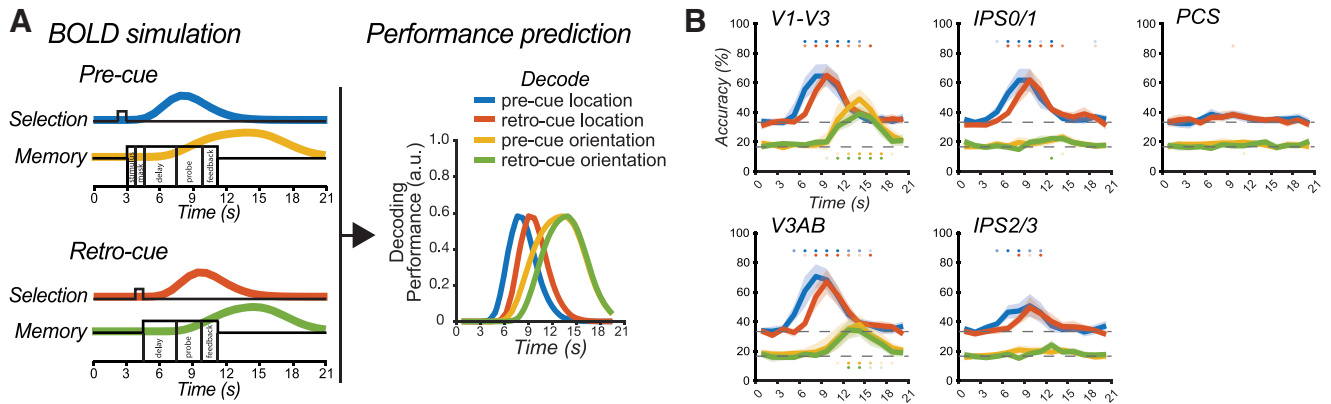
Figure 2. Temporal evolution of selection and memory decoding. *A*, Simulating selection and memory-related activity and predicting the performance of decoding target location and orientation. The neural activities for selection and memory were simulated by convolving the hypothetical duration with an HRF. The performance prediction was proportional to the amplitude of neural activities in arbitrary units. *B*, The time series of decoding performance for target location and orientation. The gray dashed lines key the chance for decoding location (33.3%) and orientation (16.7%). Blue and red lines stand for decoding location in pre-cue and retro-cue conditions, respectively. Yellow and green lines stand for decoding orientation in pre-cue and retro-cue conditions, respectively. The colored lines depict the mean of the performance for all subjects. The filled area around the line keys the 95% bootstrap CI. The dots on the top and bottom indicate the performance at the corresponding time points is significantly higher than chance (the darkest color means $p < 0.001$; the medium color means $p < 0.01$; the lightest color means $p < 0.05$; one-sided, FDR corrected across all time points for each ROI and decoding condition independently).

from all but three trials (one for each location) and then tested on the left-out trials in one task condition for each cross-validation fold, resulting in 60 cross-validation folds. Decoding accuracy was computed by comparing the true location labels with the classified labels across the 60 cross validations. Next, the classifier was trained on all of the data from one task condition and tested on each trial from the other task condition (i.e., across-condition decoding) to examine whether the neural representations were comparable in the two task conditions. To increase the reliability of our decoding estimates, we repeated the entire procedure 10 times, taking the mean of the decoding accuracy across iterations as the final within-condition and across-condition decoding accuracy.

To estimate the BOLD activity for selection and memory phases in the task, we computed the time series of decoding performance. We segmented the z-scored preprocessed BOLD time series from 0 to 21 s (i.e., the end of ITI (Inter-trial interval)) relative to the trial onset, and trained classifiers separately using the BOLD signal averaged across every 2 TRs. One set of classifiers was trained to decode target location, signifying the selection process, while the other set of classifiers was trained to decode target orientation, signifying the maintenance process. Decoding of the target location was conducted using the leave-three-trials-out cross-validation scheme described above, while decoding of target orientation was conducted using a leave-six-trials-out cross-validation scheme, corresponding to the six possible Gabor orientations.

Estimating population-level activity modulation
We tested whether external and internal selection both cause an activity increase in voxels whose receptive fields match the selected item's location relative to the unselected items' locations. To estimate the response of the neural populations at the selected location, we reconstructed a pRF-weighted map using the selection-related activity. This procedure essentially projects voxel activity in each ROI to visual space in screen coordinates (Kok and de Lange, 2014; Kwak and Curtis, 2022; Yoo et al., 2022). In each ROI, for every selected location (left, right, bottom) in each task condition (pre-cue, retro-cue), we created a reconstructed map using

$$\sum_i w_i((x,y), \mathbf{I}\sigma)\beta_i,$$

where $w_i((x,y), \mathbf{I}\sigma)$ is the weight associated with the *i*th voxel at location (x, y), and β_i is the averaged GLM-acquired selection period β at voxel *i* in trials with the same selected location. We defined w_i as the receptive field of voxel *i*, which was a two-dimensional Gaussian with

mean (μ_x, μ_y) and variance $\mathbf{I}\sigma^2$, where $\mathbf{I}$. is a two-dimensional identity matrix describing a circular, symmetric Gaussian (Fig. 3A). For each pRF-weighted selection activity map, we calculated the mean of selection-related activity in the selected and nonselected locations, and averaged across three target location conditions. The activity difference between the selected and nonselected locations was taken as the activity modulation because of selection.

**Relationship between decoding results and population-level activity modulation**
To investigate the relationship between our decoding results and population reconstructions, we conducted a median split of trials based on the magnitude of population-level modulation for the selected location and compared the relative classifier activation for the selected location for low-modulation and high-modulation trials. Classifier activation for each spatial location was calculated as the sum over voxels of the classifier training weights assigned to each voxel multiplied by the BOLD activity on a given trial, and was normalized to sum to 1 over the three classes. We used the relative classifier activation (where higher activation can be taken as greater classifier evidence for that class) for the selected location as a measure of the strength of classifier evidence for each of the three possible locations.

Distinguishing between possible mechanisms underlying enhanced population response
To explore whether observed increases in population response were consistent with relative gain modulation, we used a simulation to compare the expected influence on the spatiotopic population response under two plausible mechanisms, multiplicative gain (Reynolds and Heeger, 2009; Herrmann et al., 2010) and pRF shifts (Vo et al., 2017), with our observed population-level modulations. Each voxel's activity was simulated by the cumulative distribution function

$$\text{Activity} = N(r - \mu) = \int_{-\infty}^{r-\mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx,$$

where *r* denotes the radius of the selected Gabor stimulus, μ denotes voxel's estimated pRF center, and σ denotes voxel's estimated pRF size (given by the SD of the best-fit Gaussian; Fig. 4A). In the multiplicative gain model (Fig. 4B, left), selection-related modulation, *m*, of activity for voxels with pRF centers within the selected Gabor was given by
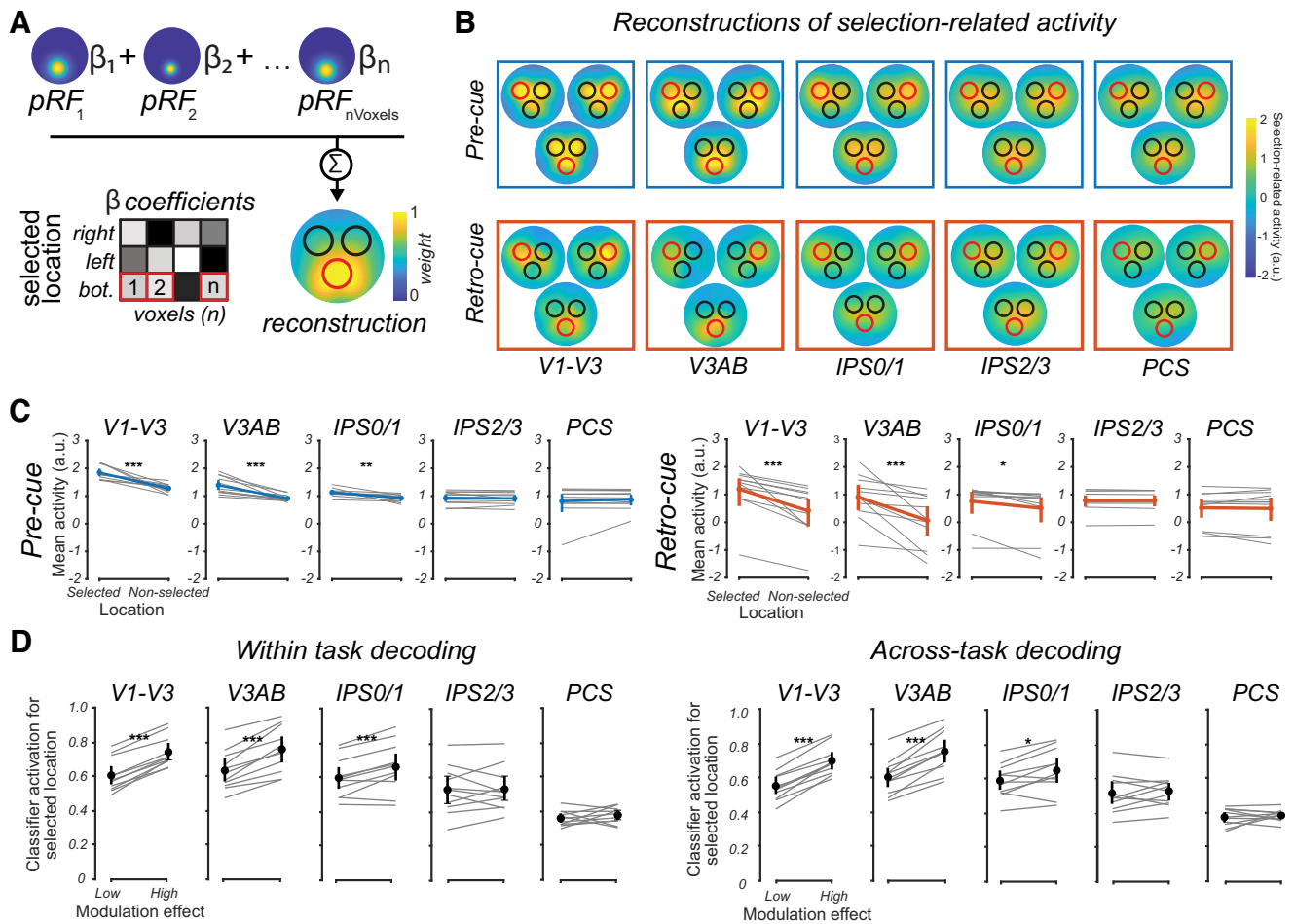
Figure 3. Reconstructing the topographic patterns of selection. ***A***, Procedure for reconstructing the pRF-weighted selection activity maps. For each condition and ROI, we multiplied the GLM-obtained selection-related β coefficients for each voxel, for each location (i.e., left, right, down) separately, by its pRF model parameters (i.e., position, size) and summed across all voxels. This procedure projects the relative population activity within the ROI in brain space into screen coordinates of visual space. ***B***, The pRF-weighted reconstruction maps for each condition and ROI. The circles in the map mark the three placeholder rings shown during the experiment, and the red ring keys the selected location. ***C***, The mean of pRF-weighted selection-related activity in the selected and nonselected locations in each task and ROI for individual subjects (gray) and across subjects (blue for pre-cue and red for retro-cue). These results indicate that changes in population gain underlie attentional and WM selection. ***D***, The relationship between the activity modulation observed in ***C*** and decoding results. Within-task (left) and across-task (right) decoding strength in the form of classifier activation is shown separately for trials with low and high modulation. Error bars are the 95% bootstrap CI. Stars in the top indicate the significance of the difference between the two conditions indicated in the x-axis (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$; one-sided, FDR corrected across all ROIs for each task condition independently).

$$m = 1 + 0.6\frac{p(d)}{p(0)} = 1 + 0.6e^{-\frac{d^2-3.5^2}{2\sigma^2}},$$

where d denotes the distance between voxel's pRF center and the center of the selected Gabor and σ denotes voxel's estimated pRF size; (cf. Reynolds and Heeger, 2009; see their Fig. 1; Herrmann et al., 2010; see their Fig. 5). In the pRF shift model (Fig. 4B, right), we assumed that voxels' pRF centers would shift a distance of s toward the center of the selected location according to

$$s = 0.2 \times (1 - e^{-1.5d}),$$

where d denotes the distance between voxel's pRF center and the center of the selected Gabor (cf. Vo et al., 2017; see their Fig. 2e). To assess which model best described our data, we calculated the sum of squared error between simulated and observed activity modulation.

Statistical analysis
For our behavioral analyses, we used paired-sample t tests to examine differences between performance across tasks. For our fMRI analyses, all statistics were calculated using a nonparametric permutation analysis (Rademaker et al., 2019). This method is appropriate here because there

was no a priori reason to believe that the data would be normally distributed. Specifically, we repeated each analysis with shuffled trial labels (i.e., selected location: left, right, bottom) 1000 times to build an empirical null distribution of the test statistic of interest (e.g., decoding accuracy). For individual-level analyses, the percentage of the empirical null distribution that was equal to or larger than the real data were taken as the p-value. For group-level analyses, our test statistic was the t-value derived from a paired t test of the real data (vs zero or chance), and the empirical null distribution was the corresponding t value for each of the 1000 iterations. The p-value was the percentage of the null distribution equal to or larger than the true t score.

Data availability
Behavioral, eye-tracking, and MRI data, and analysis code are available at https://osf.io/jqu95/.

Results

Behavioral performance
The accuracy of memory judgments in the pre-cue and retro-cue tasks indicated that participants performed each task well (Fig. 1B). Each condition used a separate staircase with a target accuracy of 80%. As expected, we found no significant difference in
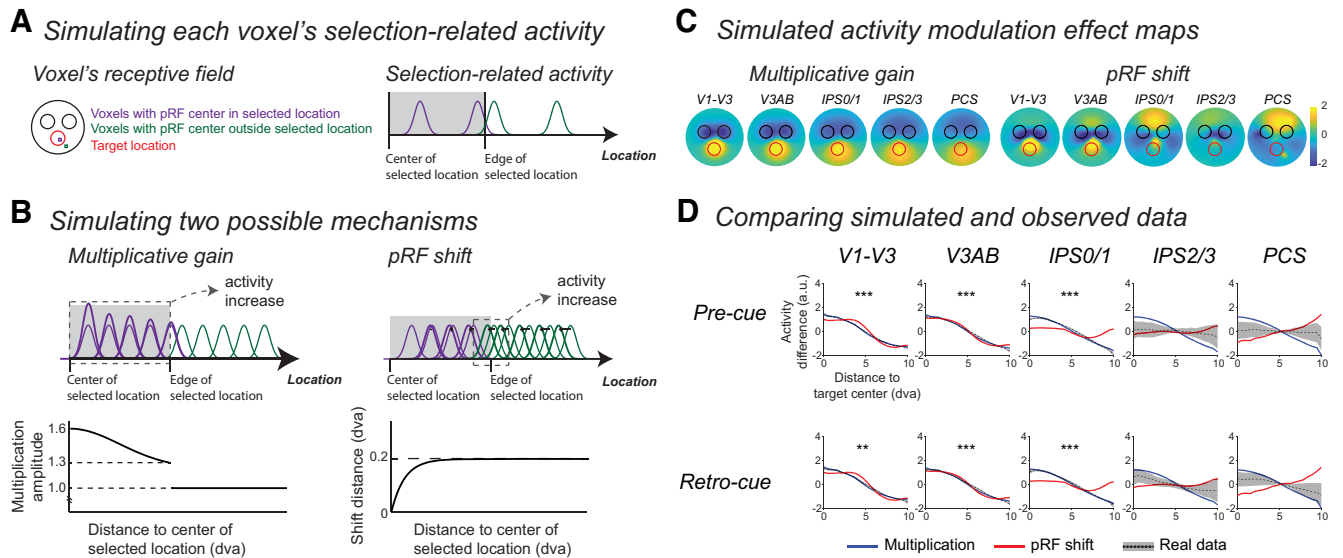
Figure 4. Simulating multiplicative gain and pRF shift mechanisms. *A*, Schematic illustrating the simulation of selection-related activity. *B*, Schematic illustrating the simulation of the two possible mechanisms. In the multiplicative gain model, activity enhancement was modulated by multiplying the selection-related activity of voxels with pRF centers in the selected location. In the pRF shift model, voxels' pRF centers were shifted toward the center of the selected location. *C*, The simulated activity modulation maps for the two possible mechanisms. We reconstructed the pRF-weighted map using the simulated selection-related activity using the same procedure as that used to generate Figure 3*B*, isolated activity modulation effect by subtracting the mean of the pRF-weighted map for the two nonselected locations from that for the selected location [e.g., middle − (left + right)/2], and collapsed the modulation maps across three selected locations. The circles in the map mark the three placeholder rings shown during the experiment, and the red ring keys the selected location. *D*, Comparison between simulated and observed data. The modulation effect decays as a voxel's pRF center increases in distance from the selected location. The blue and red lines are for multiplicative gain and pRF shift mechanism, respectively, and the black line is for the observed data. The filled area around the line indicates the 95% bootstrap CI. The stars in the top of each plot indicate that the sum of squared error between the multiplication simulation data and real data were significantly lower than that between the pRF shift simulation data and real data (***p < 0.001, **p < 0.01, *p < 0.05; one-sided, FDR corrected across all ROIs). The observed data were significantly more consistent with the multiplicative gain model in most ROIs.

accuracy between the pre-cue (81%) and retro-cue (78%) conditions, $t_{(10)} = 2.084$, $p = 0.064$, Cohen's $d = 0.628$, $BF_{10} = 1.421$. At threshold, the angular differences between the sample and probe orientations in the pre-cue (13.64°) and retro-cue (15.76°) conditions did not significantly differ, $t_{(10)} = 1.856$, $p = 0.093$, Cohen's $d = 0.560$, $BF_{10} = 1.074$, suggesting that performance was well-matched across the two tasks.

Decoding target location from isolated selection-related BOLD activity
First, we estimated selection-related activity on each trial using a GLM that modeled each of the trial components for each cueing condition (Materials and Methods). Second, for each condition separately, we trained classifiers to predict the cued locations (i.e., left, right, down; Fig. 1*C*). For both pre-cue and retro-cue trials, most of the ROIs could successfully decode the selected location (Fig. 1*D*), confirming that our decoding procedures were robust. Third, to test our main hypothesis regarding a shared neural selection mechanism, we performed across-condition decoding. We trained classifiers using pre-cue data and attempted to decode the cued location using retro-cue data. Similarly, we trained classifiers using retro-cue data and decoded the cued locations using pre-cue data. Critically, we found that most of the ROIs could successfully decode the selected location across cueing conditions (Fig. 1*D*). The performance of across-condition decoding was ~90% of within-condition decoding, suggesting that the neural activity patterns were nearly interchangeable across the two tasks. We ruled out that these results were because of gaze shifts during selection (Theeuwes et al., 2005, 2009; Hedge and Leonards, 2013); classifiers trained on eye position were unable to predict the selected location (all FDR corrected $ps \geq 0.05$).

Temporal evolution of selection and memory
To confirm that our decoding results indeed represented selection as opposed to incidental maintenance of the selected location, we examined the time course of decoding accuracy. This analysis was motivated by the fact that location information was essential during the selection process itself, but ceased to be task-relevant once selection was complete (as the memory probe was presented at fixation). Thus, we predicted that the time course of decoding accuracy for location would rise following the onset of the selection cue and return to chance once selection was complete, while the time course of orientation decoding (signifying the maintenance, retrieval, and response processes) would peak after selection and remain above chance for the duration of the trial. To formalize this prediction, we separately convolved the selection regressor and the combined delay + response + feedback regressors from our GLM with the HRF to simulate the expected time courses of the selection and memory processes, respectively (Fig. 2*A*). We then compared these simulated time courses with the actual decoding time course for location and orientation. Critically, the decoding time course analysis was completely independent from the GLM used to derive the simulated time courses (see Materials and Methods). The decoding accuracy time series for location and orientation (Fig. 2*B*) closely matched our simulated time courses. Specifically, the rapid rise and fall of location decoding in visual and parietal cortex matched the simulated time course of selection, whereas orientation decoding in visual cortex mirrored the simulated memory time course. Moreover, location decoding accuracy peaked earlier in the pre-cue relative to the retro-cue condition, consistent with the relative timing of selection in the two tasks. Together, these observations support our claim that location decoding reflected the selection process while orientation decoding reflected memory processes.

Reconstructing maps of selection-related activity

Although these results derived from machine learning approaches provide strong evidence for interchangeable patterns of activity during attentional and WM selection, it is notoriously difficult to make direct inferences about neural mechanisms based on significant classification (Freeman et al., 2011; Naselaris et al., 2011; Serences and Saproo, 2012). We used each voxel's pRF parameters (i.e., position and size) to project selection-related activity within each ROI to the screen coordinates of visual space (Fig. 3A). We used these reconstructions to visualize how selection impacted the distribution of activity within the retinotopically organized maps. Notably, we found increased activity in the portions of the maps containing the selected target relative to the distractors for both the pre-cue and retro-cue conditions (Fig. 3B). To quantify these results, we compared the mean activity within the selected and nonselected locations. For both pre-cue and retro-cue conditions, the mean activity at the selected location was higher in V1–V3, V3AB, and IPS0/1 (Fig. 3C). To investigate whether this relative difference in activity explained the significant cross-decoding between pre-cued and retro-cued selection, we conducted a median split of trials for each subject based on the magnitude of population modulation for the selected location and compared classifier evidence (quantified as the classifier activation for the selected location relative to the activation for the nonselected locations) for low-modulation and high-modulation trials. Both within-task and across-task classifier evidence were significantly greater for high-modulation relative to low-modulation trials in V1–V3, V3AB, and IPS0/1 (Fig. 3D). Such relative difference in spatiotopic activity modulation explains why we found significant cross-decoding between selection during pre-cue and retro-cue tasks. Most provocatively, these findings are highly consistent with the effects of multiplicative gain enhancement that have been observed at the population (Corbetta et al., 1990; Sprague and Serences, 2013) and single-neuron (Connor et al., 1997; McAdams and Maunsell, 1999; Treue and Martínez Trujillo, 1999) level; as such, our results point to a plausible neural mechanism underlying a shared mechanism of external and internal selection.

To provide further support for the notion that gain enhancement is the mechanism underlying our observed changes in population response, we used simulations to directly compare two mechanisms of selection. We simulated fMRI data expected under multiplicative gain (Fig. 4B, left) and under pRF shifts (Fig. 4B, right) and compared these data to our observed data in Figure 3C. Critically, while both multiplicative gain (Reynolds and Heeger, 2009; Herrmann et al., 2010) and RF shifts (Vo et al., 2017) have been observed with fMRI during spatial selection, these mechanisms make different predictions about the spatial pattern of modulation in voxels that represent the selected location. Specifically, the gain model predicts that the largest enhancement in response would occur in voxels whose RFs cover the center of the selected location, while the RF shift model predicts a larger enhancement in voxels whose RFs cover the periphery of the selected location (i.e., voxels whose RF centers shift from outside the selected location to inside the selected location). The simulated topographic patterns of activity shown in Figure 4C are consistent with the prediction of the multiplicative gain mechanism. The sum of squared error between observed and simulated activity enhancement modulation suggests that our observed data were significantly more consistent with the multiplicative gain model in V1–V3, V3AB, and IPS1/0 (all FDR corrected $ps \leq 0.004$), but not in IPS2/3 or PCS where the two models did not significantly differ (all FDR corrected $ps \geq 0.080$;

Fig. 4D). These findings bolster the claim that gain modulation supports a common selection operation.

Discussion

While both attention and WM involve the preferential processing of task relevant information, here we addressed the extent to which they draw on shared mechanisms. Importantly, the intuitive and appealing notion that attentional and WM selection reflect a single underlying process (Chun et al., 2011; Kiyonaga and Egner, 2013) lacks direct evidence. Behavioral studies using dual task paradigms report inconsistent effects of a concurrent attention-demanding task on WM selection (Janczyk and Berryhill, 2014; Lin et al., 2021; but see Hollingworth and Maxcey-Richard, 2013; Rerko et al., 2014; Makovski and Pertzov, 2015). Furthermore, while neural studies consistently observe that attentional and WM selection evoke activity in overlapping brain regions (Griffin and Nobre, 2003; Nobre et al., 2004; Kuo et al., 2009; Gazzaley and Nobre, 2012), this is a qualitative and not a quantitative conclusion, and crucially ignores the possibility that selection processes may draw on distinct neural mechanisms that coexist in the same brain regions. As an illustrative example, perception of first-order and second-order motion have been found to activate nearly identical regions in V1 (Nishida et al., 2003), but a quantitative comparison demonstrated that different populations were responsible for each (Ashida et al., 2007). Our findings establish a stronger case for the overlap of selection operations for perceptual and mnemonic information than prior work by identifying interchangeable patterns of activity in human visual, parietal, and to a lesser extent frontal cortices during attentional and WM selection. An important distinction between our work and existing MVPA studies comparing attention, perception, and WM is that these other studies have focused on overlap in perceptual and mnemonic representations (Serences et al., 2009; Jerde et al., 2012), as opposed to the operations that facilitate the use of these representations to guide behavior. In addition, our results represent an advance over previous studies that found task-generalized decoding only in prefrontal cortex (Panichello and Buschman, 2021) or not at all (Tamber-Rosenau et al., 2011) by providing evidence for a generalized selection mechanism that spans multiple cortical regions. An important area for future studies will be to identify the specific contexts that engage this shared mechanism.

Another key theoretical advance of our work lies in identifying a putative mechanism, activity enhancement of spatiotopic population-level responses in the selected location (i.e., gain enhancement), that underlies selection in attention and WM. Enhanced sensory responses during attentional selection is a well-established finding at the population level (Corbetta et al., 1990; Mangun et al., 1993; Sprague and Serences, 2013), most likely because of a multiplicative scaling of neuronal responses within the attended receptive field (Connor et al., 1997; McAdams and Maunsell, 1999; Treue and Martínez Trujillo, 1999; Williford and Maunsell, 2006) that drives preferential processing of perceptual information. In contrast, there is comparatively little evidence for the role of gain enhancement in WM selection besides qualitative similarities between WM-induced modulations of visual processing and the enhancement of sensory processing during attentional selection (Awh et al., 2000; Sreenivasan and Jha, 2007; Sreenivasan et al., 2007; Merrikhi et al., 2017). Our findings suggest that the vast literature on multiplicative

gain can be leveraged to better understand how we select information from within WM. According to theory, increases in neural gain enhance the signal-to-noise ratio, and therefore precision, of neural representations (Zemel et al., 1998; Ma et al., 2006). Thus, the gain increases we observed associated with selection might control which items are prioritized in WM. Notably, while gain enhancement is the most plausible mechanism based on our simulation results (Fig. 4), we do not claim to have conclusively ruled out all mechanisms of selection. Importantly, our results constrain theory by demonstrating that any plausible mechanism would need to produce equivalent relative activity modulation at the populational level.

Do our decoding results reflect the selection process itself or the outcome of the selection process (i.e., the consequence of having selected a particular location; Myers et al., 2017)? We considered this question in two ways – first by examining the time course of location decoding to distinguish between selection and memory for the selected information. We found that the strongest decoding of location was time-locked to the selection events in both tasks, with decoding accuracy quickly falling to chance once the location information was no longer relevant. In contrast, orientation decoding peaked later after selection and remained above chance for the entire memory delay. This pattern of findings indicates a transient process by which the task-relevant location was selected followed by prolonged maintenance of the target's orientation. Thus, location decoding in our data likely represents the selection process itself. The time-limited representation of selection that we observed may help explain important discrepancies in the behavioral literature – some studies find dual-task costs between attention and selection in WM (Janczyk and Berryhill, 2014; Lin et al., 2021) and others do not (Hollingworth and Maxcey-Richard, 2013; Rerko et al., 2014; Makovski and Pertzov, 2015). We argue that studies that have failed to observe interference generally assume that attention is continuously applied to maintain selection in WM (Hollingworth and Maxcey-Richard, 2013), while those that find interference generally put the secondary task temporally near the selection cue (Janczyk and Berryhill, 2014). Further elucidation of the temporal profiles of selection in WM should be an important area for future investigation.

Second, despite the fact that our study was not specifically designed to distinguish the sources controlling selection from the effects of selection, our findings intriguingly point to potentially distinct roles of visual and association cortex. In visual cortex, not only could we decode the selected location, but in later time points we could decode the target orientation held in memory, suggesting a role in WM storage (Curtis and D'Esposito, 2003; Serences et al., 2009; Sreenivasan et al., 2014), potentially as a consequence of receiving top-down selection signals (Sprague et al., 2016; Rahmati et al., 2018; Yoo et al., 2022). On the other hand, in parietal and frontal cortex, we had robust decoding of selected location, while decoding of memorized orientation was inconsistent across the delay and across ROIs (Fig. 2B), consistent with the idea that gain enhancement in topographically organized regions of parietal and frontal cortex reflects the sources of top-down signals controlling which locations are selected. While we cannot completely rule out the possibility that unsuccessful orientation decoding was because of larger RF sizes or increased spatial heterogeneity in the representation of features in these regions (for a successful demonstration of orientation decoding in frontoparietal cortices, see Ester et al., 2015), our findings are reminiscent of the role of frontoparietal cortex in orienting attention and prioritization (Corbetta and Shulman, 2002; Serences and Yantis,

2006; Jerde et al., 2012) and gating (Frank et al., 2001; Chatham et al., 2014), and spatial cognition more broadly (Heilman et al., 1985; Mesulam, 1999; Corbetta et al., 2000; Vandenberghe et al., 2001; Yantis et al., 2002; Srimal and Curtis, 2008; Szczepanski et al., 2010; Mackey et al., 2016). It is worth noting that a previous study that failed to find a common multivoxel pattern across attentional and WM selection used different features (location and object) for selection across tasks (Tamber-Rosenau et al., 2011). Given that neural substrates of selection are sensitive to differences in the medium of selection (Giesbrecht et al., 2003), conclusions about a common selection mechanism ought to be drawn from comparisons which rigorously match the attention and WM tasks. Here, we compare across the same selection medium (location), equating relevant features, as well as behavioral performance. Consistent with the idea that WM selection relies on internally-directed shifts of attention that highlight task-relevant information, our results suggest that a common mechanism underlies selection during attention and WM.

References

Arco JE, González-García C, Díaz-Gutiérrez P, Ramírez J, Ruz M (2018) Influence of activation pattern estimates and statistical significance tests in fMRI decoding analysis. J Neurosci Methods 308:248–260.

Ashida H, Lingnau A, Wall MB, Smith AT (2007) FMRI adaptation reveals separate mechanisms for first-order and second-order motion. J Neurophysiol 97:1319–1325.

Awh E, Jonides J (2001) Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci 5:119–126.

Awh E, Jonides J, Smith EE, Buxton RB, Frank LR, Love T, Wong EC, Gmeindl L (1999) Rehearsal in spatial working memory: evidence from neuroimaging. Psychol Sci 10:433–437.

Awh E, Anllo-Vento L, Hillyard SA (2000) The role of spatial selective attention in working memory for locations: evidence from event-related potentials. J Cogn Neurosci 12:840–847.

Bahle B, Thayer DD, Mordkoff JT, Hollingworth A (2020) The architecture of working memory: features from multiple remembered objects produce parallel, coactive guidance of attention in visual search. J Exp Psychol Gen 149:967–983.

Bisley JW, Goldberg ME (2003) Neuronal activity in the lateral intraparietal area and spatial attention. Science 299:81–86.

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:433–436.

Bushnell MC, Goldberg ME, Robinson DL (1981) Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. J Neurophysiol 46:755–772.

Carrasco M (2011) Visual attention: the past 25 years. Vision Res 51:1484–1525.

Carrasco M, Penpeci-Talgar C, Eckstein M (2000) Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. Vision Res 40:1203–1215.

Chatham CH, Frank MJ, Badre D (2014) Corticostriatal output gating during selection from working memory. Neuron 81:930–942.

Chun MM (2011) Visual working memory as visual attention sustained internally over time. Neuropsychologia 49:1407–1409.

Chun MM, Golomb JD, Turk-Browne NB (2011) A taxonomy of external and internal attention. Annu Rev Psychol 62:73–101.

Connor CE, Preddie DC, Gallant JL, Van Essen DC (1997) Spatial attention effects in macaque area V4. J Neurosci 17:3201–3214.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215.

Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1990) Attentional modulation of neural processing of shape, color, and velocity in humans. Science 248:1556–1559.

Corbetta M, Kincade JM, Ollinger JM, McAvoy MP, Shulman GL (2000) Voluntary orienting is dissociated from target detection in human posterior parietal cortex. Nat Neurosci 3:292–297.

Cowan N (1998a) Attention and memory: an integrated framework. Oxford: Oxford University Press.

Cowan N (1998b) Visual and auditory working memory capacity. Trends Cogn Sci 2:77.

Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7:415–423.

Eriksen CW, Yeh YY (1985) Allocation of attention in the visual field. J Exp Psychol Hum Percept Perform 11:583–597.

Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. Neuron 87:893–905.

Ester EF, Nouri A, Rodriguez L (2018) Retrospective cues mitigate information loss in human cortex during working memory storage. J Neurosci 38:8538–8548.

Fougnie D, Marois R (2006) Distinct capacity limits for attention and working memory: evidence from attentive tracking and visual working memory paradigms. Psychol Sci 17:526–534.

Frank MJ, Loughry B, O'Reilly RC (2001) Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cogn Affect Behav Neurosci 1:137–160.

Freeman J, Brouwer GJ, Heeger DJ, Merriam EP (2011) Orientation decoding depends on maps, not columns. J Neurosci 31:4792–4804.

Gandhi SP, Heeger DJ, Boynton GM (1999) Spatial attention affects brain activity in human primary visual cortex. Proc Natl Acad Sci U S A 96:3314–3319.

Gayet S, Paffen CLE, Van der Stigchel S (2013) Information matching the content of visual working memory is prioritized for conscious access. Psychol Sci 24:2472–2480.

Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. Trends Cogn Sci 16:129–135.

Gazzaley A, Cooney JW, Rissman J, D'Esposito M (2005) Top-down suppression deficit underlies working memory impairment in normal aging. Nat Neurosci 8:1298–1300.

Giesbrecht B, Woldorff MG, Song AW, Mangun GR (2003) Neural mechanisms of top-down control during spatial and feature attention. Neuroimage 19:496–512.

Griffin IC, Nobre AC (2003) Orienting attention to locations in internal representations. J Cogn Neurosci 15:1176–1194.

Hedge C, Leonards U (2013) Using eye movements to explore switch costs in working memory. J Vis 13:18.

Heilman KM, Bowers D, Coslett HB, Whelan H, Watson RT (1985) Directional hypokinesia: prolonged reaction times for leftward movements in patients with right hemisphere lesions and neglect. Neurology 35:855–859.

Herrmann K, Montaser-Kouhsari L, Carrasco M, Heeger DJ (2010) When size matters: attention affects performance by contrast or response gain. Nat Neurosci 13:1554–1559.

Hollingworth A, Maxcey-Richard AM (2013) Selective maintenance in visual working memory does not require sustained visual attention. J Exp Psychol Hum Percept Perform 39:1047–1058.

Hopfinger JB, Woldorff MG, Fletcher EM, Mangun GR (2001) Dissociating top-down attentional control from selective perception and action. Neuropsychologia 39:1277–1291.

Ikkai A, Curtis CE (2008) Cortical activity time locked to the shift and maintenance of spatial attention. Cereb Cortex 18:1384–1394.

Ikkai A, Curtis CE (2011) Common neural mechanisms supporting spatial working memory, attention and motor intention. Neuropsychologia 49:1428–1434.

Janczyk M, Berryhill ME (2014) Orienting attention in visual working memory requires central capacity: decreased retro-cue effects under dual-task conditions. Atten Percept Psychophys 76:715–724.

Jerde TA, Merriam EP, Riggall AC, Hedges JH, Curtis CE (2012) Prioritized maps of space in human frontoparietal cortex. J Neurosci 32:17382–17390.

Jha AP (2002) Tracking the time-course of attentional involvement in spatial working memory: an event-related potential investigation. Brain Res Cogn Brain Res 15:61–69.

Katus T, Andersen SK, Müller MM (2014) Common mechanisms of spatial attention in memory and perception: a tactile dual-task study. Cereb Cortex 24:707–718.

Kay KN, Winawer J, Mezer A, Wandell BA (2013) Compressive spatial summation in human visual cortex. J Neurophysiol 110:481–494.

Kiyonaga A, Egner T (2013) Working memory as internal attention: toward an integrative account of internal and external selection processes. Psychon Bull Rev 20:228–242.

Kok P, de Lange FP (2014) Shape perception simultaneously up- and down-regulates neural activity in the primary visual cortex. Curr Biol 24:1531–1535.

Krishnapuram B, Carin L, Figueiredo MAT, Hartemink AJ (2005) Sparse multinomial logistic regression: fast algorithms and generalization bounds. IEEE Trans Pattern Anal Mach Intell 27:957–968.

Kuo BC, Rao A, Lepsien J, Nobre AC (2009) Searching for targets within the spatial layout of visual short-term memory. J Neurosci 29:8032–8038.

Kwak Y, Curtis CE (2022) Unveiling the abstract format of mnemonic representations. Neuron 110:1822–1828.e5.

LaBar KS, Gitelman DR, Parrish TB, Mesulam M (1999) Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. Neuroimage 10:695–704.

Landman R, Spekreijse H, Lamme VAF (2003) Large capacity storage of integrated objects before change blindness. Vision Res 43:149–164.

Lepsien J, Thornton I, Nobre AC (2011) Modulation of working-memory maintenance by directed attention. Neuropsychologia 49:1569–1577.

Levitt H (1971) Transformed up-down methods in psychoacoustics. J Acoust Soc Am 49:467–477.

Li Z, Tong M, Chen S, Qian J (2021) Effect of attentional selection on working memory for depth in a retro-cueing paradigm. Mem Cognit 49:747–757.

Lin YT, Sasin E, Fougnie D (2021) Selection in working memory is resource-demanding: concurrent task effects on the retro-cue effect. Atten Percept Psychophys 83:1600–1612.

Liu T, Pestilli F, Carrasco M (2005) Transient attention enhances perceptual performance and FMRI response in human visual cortex. Neuron 45:469–477.

Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J Neurophysiol 77:24–42.

Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. Nat Neurosci 9:1432–1438.

Mackey WE, Devinsky O, Doyle WK, Golfinos JG, Curtis CE (2016) Human parietal cortex lesions impact the precision of spatial working memory. J Neurophysiol 116:1049–1054.

Mackey WE, Winawer J, Curtis CE (2017) Visual field map clusters in human frontoparietal cortex. Elife 6:e22974.

Makovski T, Pertzov Y (2015) Attention and memory protection: interactions between retrospective attention cueing and interference. Q J Exp Psychol (Hove) 68:1735–1743.

Mangun GR, Hillyard SA, Luck SJ (1993) Electrocortical substrates of visual selective attention. In: Attention and performance XIV: synergies in experimental psychology, artificial intelligence, and cognitive neuroscience, pp 219–243. Cambridge: The MIT Press.

Marois R, Ivanoff J (2005) Capacity limits of information processing in the brain. Trends Cogn Sci 9:296–305.

McAdams CJ, Maunsell JH (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. J Neurosci 19:431–441.

Merrikhi Y, Clark K, Albarran E, Parsa M, Zirnsak M, Moore T, Noudoost B (2017) Spatial working memory alters the efficacy of input to visual cortex. Nat Commun 8:15041.

Mesulam MM (1999) Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. Philos Trans R Soc Lond B Biol Sci 354:1325–1346.

Mumford JA, Turner BO, Ashby FG, Poldrack RA (2012) Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 59:2636–2643.

Murphy TD, Eriksen CW (1987) Temporal changes in the distribution of attention in the visual field in response to precues. Percept Psychophys 42:576–586.

Myers NE, Stokes MG, Nobre AC (2017) Prioritizing information during working memory: beyond sustained internal attention. Trends Cogn Sci 21:449–461.

Naselaris T, Kay KN, Nishimoto S, Gallant JL (2011) Encoding and decoding in fMRI. Neuroimage 56:400–410.

Nishida S, Sasaki Y, Murakami I, Watanabe T, Tootell RB (2003) Neuroimaging of direction-selective mechanisms for second-order motion. J Neurophysiol 90:3242–3254.

Nobre AC, Coull JT, Maquet P, Frith CD, Vandenberghe R, Mesulam MM (2004) Orienting attention to locations in perceptual versus mental representations. J Cogn Neurosci 16:363–373.

Oberauer K (2002) Access to information in working memory: exploring the focus of attention. J Exp Psychol Learn Mem Cogn 28:411–421.

Olivers CNL, Meijer F, Theeuwes J (2006) Feature-based memory-driven attentional capture: visual working memory content affects visual attention. J Exp Psychol Hum Percept Perform 32:1243–1265.

Panichello MF, Buschman TJ (2021) Shared mechanisms underlie the control of working memory and attention. Nature 592:601–605.

Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: a tutorial overview. Neuroimage 45:S199–S209.

Pestilli F, Carrasco M (2005) Attention enhances contrast sensitivity at cued and impairs it at uncued locations. Vision Res 45:1867–1875.

Pestilli F, Viera G, Carrasco M (2007) How do attention and adaptation affect contrast sensitivity? J Vis 7:9.1–12.

Pestilli F, Ling S, Carrasco M (2009) A population-coding model of attention's influence on contrast response: estimating neural effects from psychophysical data. Vision Res 49:1144–1153.

Pollmann S, von Cramon DY (2000) Object working memory and visuospatial processing: functional neuroanatomy analyzed by event-related fMRI. Exp Brain Res 133:12–22.

Posner MI (1980) Orienting of attention. Q J Exp Psychol 32:3–25.

Postle BR, Awh E, Jonides J, Smith EE, D'Esposito M (2004) The where and how of attention-based rehearsal in spatial working memory. Brain Res Cogn Brain Res 20:194–205.

Rademaker RL, Chunharas C, Serences JT (2019) Coexisting representations of sensory and mnemonic information in human visual cortex. Nat Neurosci 22:1336–1344.

Rahmati M, Saber GT, Curtis CE (2018) Population dynamics of early visual cortex during working memory. J Cogn Neurosci 30:219–233.

Ranganath C, D'Esposito M (2005) Directing the mind's eye: prefrontal, inferior and medial temporal mechanisms for visual working memory. Curr Opin Neurobiol 15:175–182.

Rerko L, Souza AS, Oberauer K (2014) Retro-cue benefits in working memory without sustained focal attention. Mem Cognit 42:712–728.

Reynolds JH, Heeger DJ (2009) The normalization model of attention. Neuron 61:168–185.

Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. Neuron 26:703–714.

Sasin E, Fougnie D (2020) Memory-driven capture occurs for individual features of an object. Sci Rep 10:19499.

Serences JT, Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. Neuron 55:301–312.

Serences JT, Saproo S (2012) Computational advances towards linking BOLD and behavior. Neuropsychologia 50:435–446.

Serences JT, Yantis S (2006) Selective visual attention and perceptual coherence. Trends Cogn Sci 10:38–45.

Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. Psychol Sci 20:207–214.

Shen M, Huang X, Gao Z (2015) Object-based attention underlies the rehearsal of feature binding in visual working memory. J Exp Psychol Hum Percept Perform 41:479–493.

Silver MA, Ress D, Heeger DJ (2007) Neural correlates of sustained spatial attention in human early visual cortex. J Neurophysiol 97:229–237.

Souza AS, Rerko L, Oberauer K (2016) Getting more from visual working memory: retro-cues enhance retrieval and protect from visual interference. J Exp Psychol Hum Percept Perform 42:890–910.

Souza AS, Czoschke S, Lange EB (2020) Gaze-based and attention-based rehearsal in spatial working memory. J Exp Psychol Learn Mem Cogn 46:980–1003.

Sperling G (1960) The information available in brief visual presentations. Psychol Monogr 74:1–29.

Sprague TC, Serences JT (2013) Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. Nat Neurosci 16:1879–1887.

Sprague TC, Ester EF, Serences JT (2016) Restoring latent visual working memory representations in human cortex. Neuron 91:694–707.

Sreenivasan KK, Jha AP (2007) Selective attention supports working memory maintenance by modulating perceptual processing of distractors. J Cogn Neurosci 19:32–41.

Sreenivasan KK, Katz J, Jha AP (2007) Temporal characteristics of top-down modulations during working memory maintenance: an event-related potential study of the N170 component. J Cogn Neurosci 19:1836–1844.

Sreenivasan KK, Curtis CE, D'Esposito M (2014) Revisiting the role of persistent neural activity during working memory. Trends Cogn Sci 18:82–89.

Srimal R, Curtis CE (2008) Persistent neural activity during the maintenance of spatial position in working memory. Neuroimage 39:455–468.

Szczepanski SM, Konen CS, Kastner S (2010) Mechanisms of spatial attention control in frontal and parietal cortex. J Neurosci 30:148–160.

Tamber-Rosenau BJ, Esterman M, Chiu Y-C, Yantis S (2011) Cortical mechanisms of cognitive control for shifting attention in vision and working memory. J Cogn Neurosci 23:2905–2919.

Theeuwes J, Olivers CNL, Chizk CL (2005) Remembering a location makes the eyes curve away. Psychol Sci 16:196–199.

Theeuwes J, Belopolsky A, Olivers CNL (2009) Interactions between working memory, attention and eye movements. Acta Psychol (Amst) 132:106–114.

Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399:575–579.

Vandenberghe R, Gitelman DR, Parrish TB, Mesulam MM (2001) Functional specificity of superior parietal mediation of spatial shifting. Neuroimage 14:661–673.

Vo VA, Sprague TC, Serences JT (2017) Spatial tuning shifts increase the discriminability and fidelity of population codes in visual cortex. J Neurosci 37:3386–3401.

Wandell BA, Brewer AA, Dougherty RF (2005) Visual field map clusters in human cortex. Philos Trans R Soc Lond B Biol Sci 360:693–707.

Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. Neuron 56:366–383.

Williams JR, Brady TF, Störmer VS (2022) Guidance of attention by working memory is a matter of representational fidelity. J Exp Psychol Hum Percept Perform 48:202–231.

Williford T, Maunsell JHR (2006) Effects of spatial attention on contrast response functions in macaque area V4. J Neurophysiol 96:40–54.

Woodman GF, Luck SJ (2007) Do the contents of visual working memory automatically influence attentional selection during visual search? J Exp Psychol Hum Percept Perform 33:363–377.

Yantis S, Schwarzbach J, Serences JT, Carlson RL, Steinmetz MA, Pekar JJ, Courtney SM (2002) Transient neural activity in human parietal cortex during spatial attention shifts. Nat Neurosci 5:995–1002.

Yoo AH, Bolaños A, Hallenbeck GE, Rahmati M, Sprague TC, Curtis CE (2022) Behavioral prioritization enhances working memory precision and neural population gain. J Cogn Neurosci 34:365–379.

Zemel RS, Dayan P, Pouget A (1998) Probabilistic interpretation of population codes. Neural Comput 10:403–430.