Competition between endogenous and exogenous attention delays the selection of task-relevant working memory content.

Edward F. Ester[1], Asal Nouri[2]

[1]Department of Psychology and Integrative Neuroscience Program, University of Nevada, Reno
[2]Center for Complex Systems & Brain Sciences, Florida Atlantic University

---

Funding: Supported by NSF 2050833 (E.F.E.)

Author Contributions: E.F.E. conceived and designed the experiment. E.F.E. and A.N. collected and analyzed the data. E.F.E. and A.N. co-wrote the manuscript.

Number of Figures: 8

Word Count (Intro + Results + Discussion, excepting figure captions): 4,851

Abstract

Working memory (WM) is a capacity- and duration-limited system that forms a temporal bridge between sensory phenomena and action. Capacity limits in WM necessitate the existence of external selection mechanisms that control access to this system (i.e., input gating), while changing behavioral demands necessitate the existence of internal selection mechanisms that prioritize existing WM representations for action (i.e., output gating). Whether similar mechanisms mediate WM input and output gating is unclear. We examined this possibility by exploiting a well-known distinction between top-down (endogenous) and bottom-up (exogenous) selection. Specifically, we examined how endogenous and exogenous factors compete to control the selection of WM content by examining their effects on EEG decoding performance. We found no evidence for a reflexive selection of task-irrelevant WM content when endogenous and exogenous factors conflicted. Instead, the selection of task-relevant WM content was delayed when these factors conflicted compared to when they were aligned. Thus, unlike external attention where the real-time balance of competition between endogenous and exogenous factors determines what information is selected, the selection of task-relevant WM content is delayed until competition between endogenous and exogenous factors is resolved.

Efficient behavioral selection requires rapid comparisons of sensory inputs with internal representations of motor affordances and goal states. Many of these comparisons take place in working memory (WM), a capacity- and duration-limited system that forms a temporal bridge between fleeting sensory phenomena and possible actions (D'Esposito & Postle, 2015; van Ede & Nobre, 2022). Capacity limits in WM necessitate the existence of external selection mechanisms that gate access to this system (i.e., "input gating"; Vogel et al., 2005; O'Reilly & Frank, 2006; Rac-Lubashevsky & Frank, 2021), while rapidly changing environmental circumstances necessitate the existence of internal selection mechanisms that prioritize behaviorally relevant subsets of information stored in WM for action (i.e., "output gating"; Chatham et al., 2014; Ester et al., 2018). Whether similar mechanisms mediate the selection of internal and external information is hotly debated (O'Reilly & Frank, 2006; Chun et al., 2011; Chatham & Badre, 2015; Souza & Oberauer, 2016; van Ede et al., 2020; Rac-Lubashevsky & Frank, 2021; van Ede & Nobre, 2022).

External sensory inputs can be selected based on top-down goals (i.e., endogenous selection) or bottom-up salience (i.e., exogenous selection), with selection ultimately determined by the balance of competition between these factors (Desimone & Duncan, 1995; Corbetta & Shulman, 2002; Buschman & Miller, 2007; Baluch & Itti, 2011; Wolfe, 2020). For example, exogenous factors can reflexively trigger the selection of salient, unexpected, or valuable stimuli (e.g., Theeuwes, 1992; Folk et al., 1992; Bacon & Egeth, 1994; Folk et al., 2009; Fukuda & Vogel, 2009; Anderson et al., 2011), disrupting top-down searches for less salient but task-relevant stimuli (e.g., Wolfe, 2020). These disruptions – termed attention capture – are frequently accompanied by

concurrent shifts in cortical and subcortical spatial priority maps that mediate eye movements and endogenous covert spatial attention (Corbetta & Shulman, 2002; Silver et al., 2005; Bisley & Goldberg, 2010; Sprague & Serences, 2013; Cosman et al., 2018).

Endogenous and exogenous factors also compete to control the selection of WM content (e.g., van Ede et al., 2020; van Ede & Nobre, 2022). However, less is known about how this competition is resolved. One possibility is that exogenous factors have the same effect on internal attention as they do on external attention: an involuntary and reflexive selection of salient but task-irrelevant information. For example, a salient external stimulus might trigger an exogenous selection of similar WM content. This would be the converse of WM-guided attention capture, where external attention is reflexively oriented to task-irrelevant stimuli that incidentally match attributes of stimuli stored in WM (Olivers et al., 2011). A second possibility is that exogenous factors trigger an automatic “refreshing” of matching but otherwise task-irrelevant WM content. This alternative is motivated by neurocomputational models of conjunctive coding (e.g., Schneegans & Bays, 2017; 2018) and short-term synaptic plasticity (Mongillo et al., 2008; Manohar et al., 2019; Li et al., 2021) where bottom-up inputs can reactivate spiking patterns and/or synaptic networks responsible for holding stimulus-specific information in WM. A third possibility is that internal selection is mediated by competition between endogenous and exogenous factors, with the latter delaying the selection of task-relevant information by the former. Although this possibility has been rejected in the external attention literature (e.g., Remington et al., 2001), it is motivated by a recent finding suggesting that orienting responses to task-relevant stimuli stored in WM are

delayed when experimental factors place endogenous and exogenous selection mechanisms in conflict (van Ede et al., 2020).

Here, we tested these models. We recorded EEG while participants performed a retrospectively cued WM task typically used to study internal attention (e.g., Griffin & Nobre, 2003; Landman et al., 2003; see Souza & Oberauer, 2016 and Myers et al., 2017 for recent comprehensive reviews). In different experimental blocks, a cue presented during WM storage indicated which of two memorized positions would be probed for recall (pro-cue trials) or which position would not be probed for recall (anti-cue trials). We reasoned that the anti-cue manipulation would create a state of conflict between endogenous and exogenous selection mechanisms, a point supported by studies documenting visual search costs when participants are cued to the identity of an upcoming distractor (e.g., Moher & Egeth, 2012; van Ede et al., 2020). Although this logic is familiar from prior research, we put it to novel use. Specifically, to discriminate between the internal selection models described in the previous paragraph, we quantified how anti-cues influenced the quality of cue-matching but task-irrelevant location-specific signals decoded from scalp EEG activity during WM storage. The capture and refreshing models predict that exogenous selection mechanisms engaged by anti-cues should prompt an inadvertent selection or refreshing of the cue-matching but task-irrelevant item on a subset of trials, yielding a transient increase in task-irrelevant position decoding performance. Conversely, the delayed selection model predicts that conflict generated by exogenous selection delays the endogenous selection of task-relevant WM content, yielding a delay in improved decoding performance during anti-cue relative to pro-cue trials.

Results

We recorded EEG while 40 human volunteers performed a retrospectively cued spatial recall task (Fig 1A). Participants remembered the positions of two discs over a brief delay, and a retrospective color cue presented 1.25 seconds later instructed participants to continue remembering the positions of both discs (i.e., uninformative trials) or to prioritize one of the discs for subsequent recall (i.e., informative trials). At the end of the trial, participants recalled the position of the task-relevant disc via mouse click. Behavioral performance was quantified via average response times and average absolute recall error (i.e., the average absolute difference between the correct and reported disc position). In separate experimental blocks, participants performed a pro-cue task or an anti-cue task. During the pro-cue task informative cues were assigned 100% validity; during the anti-cue task informative cues were assigned 0% validity (i.e., the cue color indicated which disc was task-irrelevant). This allowed us to disentangle the effects of endogenous and exogenous factors on the selection of WM content: during the pro-cue task the color cue indicates which of the two remembered stimuli are task relevant, and endogenous and exogenous selection mechanisms are aligned. During the anti-cue task, however, the color cue indicates which of the two stimuli are task-irrelevant, placing endogenous and exogenous selection mechanisms in competition (e.g., Moher & Egeth, 2012; van Ede et al., 2020). Task order (i.e., pro- followed by anti-cue or vice versa) was counterbalanced across participants, and participants were explicitly reminded about cue validity at the beginning of every block.

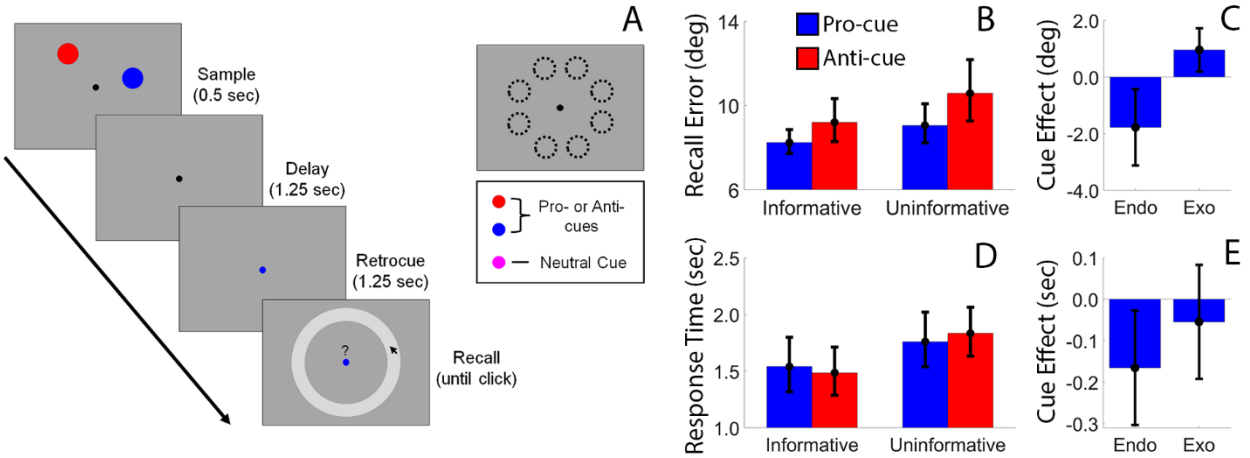

Figure 1. Memory Task and Performance*. (A) Participants remembered the locations of two discs over a blank delay. Each disc could appear at one of eight positions along the perimeter of an imaginary circle centered at fixation (upper right panel). (B) Effects of cue type (informative, uninformative) and task type (pro-cue, anti-cue) on average absolute recall errors. (C) We estimated the effects of exogenous factors on recall performance by computing the difference between informative pro-cue trials (i.e., where endogenous and exogenous factors are aligned) and informative anti-cue trials (i.e., where endogenous and exogenous factors are opposed). We estimated the effects of endogenous factors on recall performance by computing the difference between informative pro-cue trials and uninformative pro-cue trials minus the estimated effect of exogenous factors (see text for specifics). Identical analyses were also applied to participants response times (D, E). Error bars depict the 95% confidence interval of the mean. (N=150)*

Endogenous and Exogenous Factors Influence the Selection of Task-Relevant WM Content, but in Different Ways.

A two-factor repeated measures analysis of variance (ANOVA) applied to participants' average absolute recall errors (Fig 1B) revealed a main effect of cue type (i.e., informative vs. uninformative; [$F(1,39) = 15.854$, $p = 0.0003$, $\eta^2 = 0.289$]), a main effect of task (i.e., pro- vs. anti-cues; [$F(1,39) = 8.168$, $p = 0.0068$, $\eta^2 = 0.1732$]), and a significant interaction between these factors [$F(1,39) = 5.35$, $p = 0.0261$]. A complementary analysis of response times (Fig 1D) revealed a main effect of cue type [$F(1,39) = 483.046$, $p < 0.0001$, $\eta^2 = 0.9253$], no main effect of task [$F(1,39) = 0.022$, $p = 0.884$, $\eta^2 = 0.060$], and a significant interaction between these factors [$F(1,39) = 30.362$, $p < 0.0001$]. In planned comparisons, we sought further clarity on how endogenous and

exogenous factors influenced participants' memory performance. We reasoned that during informative pro-cue trials endogenous and exogenous factors are aligned while during informative anti-cue trials they conflict. Thus, to isolate the effects of exogenous factors on memory performance we compared participants' recall errors and response times across informative pro- and anti-cue trials (i.e., the simple effect of task for informative cues). Conversely, to quantify the effect of endogenous selection on memory performance, we first calculated differences in participants' recall errors and response times across informative and uninformative trials in the pro-cue task (i.e., the simple effect of cue type for the pro-cue task). Next, we subtracted the effects of exogenous factors estimated in the previous step from these differences to isolate the effects of endogenous factors on memory performance (i.e., while accounting for the fact that during the pro-cue task endogenous and exogenous cues are aligned while during the anti-cue task they are opposed). Endogenous factors had a facilitatory effect on task performance, lowering recall errors (M = 1.78°; 95% CI = 0.645°-3.112°; Fig 1C) and speeding response times (M = 0.165 sec; 95% CI = 0.014-0.305 sec; Fig 1E). In contrast, exogenous factors significantly worsened participants' recall errors (M = 0.961°; 95% CI = 0.191°-1.863°; Fig 1C) but had no effect on response times (M = -0.055, 95% CI = -0.073-0.189; Fig 1E). Thus, endogenous and exogenous factors had faciliatory and deleterious effects on participants' memory performance, respectively.

Manipulation Check: The Anti-cue Task Requires a Greater Degree of Cognitive Control than the Pro-cue task. A key assumption of our experimental approach holds that the anti-cue task produces conflict between endogenous and exogenous selection

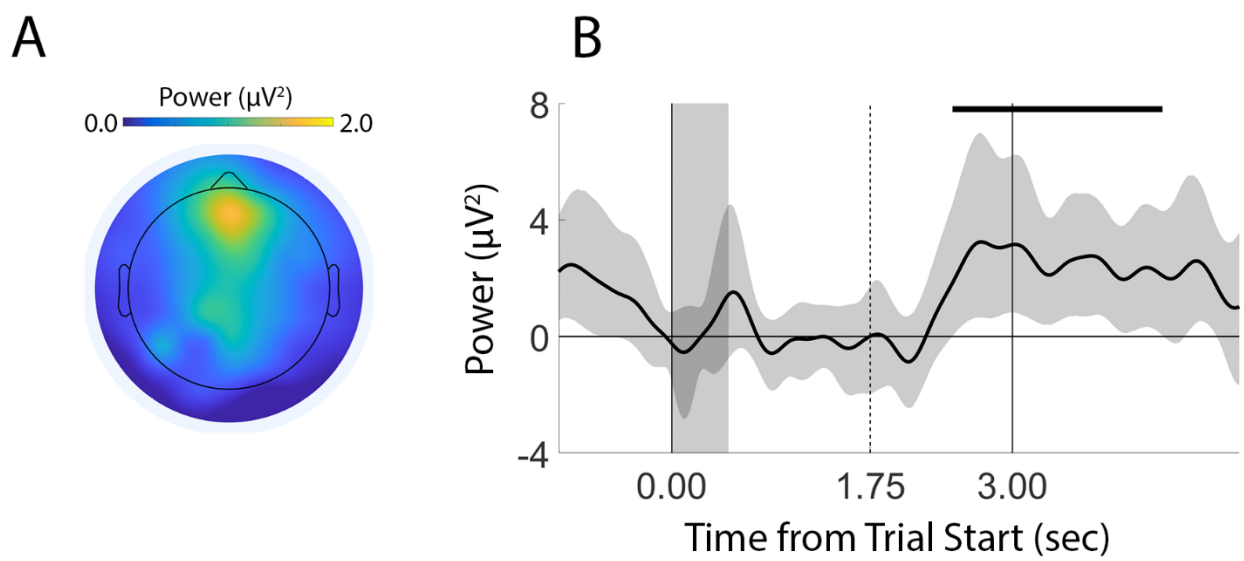

Figure 2. Frontal Theta Power is Greater During the Anti- vs. Pro-Cue Task. *(A) Difference in theta-power (4-7 Hz) scalp topography during the pro- and anti-cue tasks. Pro- and anti-cue theta power estimates were averaged over a period spanning 2.5-3.0 sec after trial start (i.e., 750-1250 ms after cue onset). Electrode-wise power estimates during the pro-cue task were subtracted from corresponding estimates during the anti-cue task, i.e., larger values indicate higher theta power during the anti- vs. pro-cue task. (B) Time-resolved differences in pro- and anti-cue theta power computed from frontal electrode sites. Theta power estimates were larger during the anti- vs. pro-cue task beginning approximately 600 ms after cue onset. Shaded regions depict the 95% confidence interval of the mean. Vertical solid lines at times 0.00 and 3.00 depict the onset of the sample and recall displays, respectively; the vertical dashed line at time 1.75 depicts the onset of an informative retrocue. The horizontal black bar at the top of the plot marks periods where the difference between anti- and pro-cue theta power was significantly greater than zero (cluster-based permutation tests; see Methods). See also Figure S1. (N = 188)*

mechanisms while the pro-cue task does not. Thus, we reasoned that a greater degree of cognitive control would be required during the anti-cue task compared to the pro-cue task. To verify this assumption, we computed estimates of theta power (4-7 Hz) over frontal electrode sites during the pro- and anti-cue tasks. Frontal theta power has robustly linked with the need for cognitive control (Cavanagh & Frank, 2014), scales with WM load (Jensen & Tesche, 2002), predicts successful working memory updating (Itthipurripat et al., 2013), and is inversely correlated with default mode network activity (Scheeringa et al., 2008). Thus, we expected larger frontal theta power estimates during

the anti-cue vs. the pro-cue task. We tested this prediction by analyzing data from informative trials during the pro- and anti-cue tasks. An analysis of theta power topography revealed significantly higher frontal theta power during the anti-cue relative to the pro-cue task (Fig 2A; Fig S1A, B). Time-resolved analyses of frontal theta power revealed that task-level differences began ~600 ms after retrocue onset and persisted well into the probe period (Fig 2B; S1C). Thus, the anti-cue task required a greater degree of cognitive control than the pro-cue task, supporting the general logic of our experimental approach.

Decoding Remembered Locations from Scalp EEG Activity

To understand how endogenous and exogenous factors influence the selection of WM content, we examined how pro- and anti-cues influenced our ability to decode stimulus positions from scalp EEG. Our approach builds on studies demonstrating that stimulus- and location-specific information can be decoded from scalp EEG signals (e.g., Foster et al., 2016) and that attending to an item or location stored in WM selectively boosts decoding for the attended information (e.g., Lewis-Peacock et al., 2012; LaRocque et al., 2013; Rose et al., 2016; Ester et al., 2018; Nouri & Ester, 2020).

We implemented a multivariate distance-based decoding analysis (e.g., Wolff et al., 2017) that was customized for our (parametric, circular) location space (e.g., Fig 1A; Fig 3). This approach is conceptually similar to image reconstruction techniques (i.e., “inverted encoding models”, Sprague et al., 2014; 2016; Ester et al., 2018) but does not require the experimenter to measure or specify a specific coding model or basis set. A

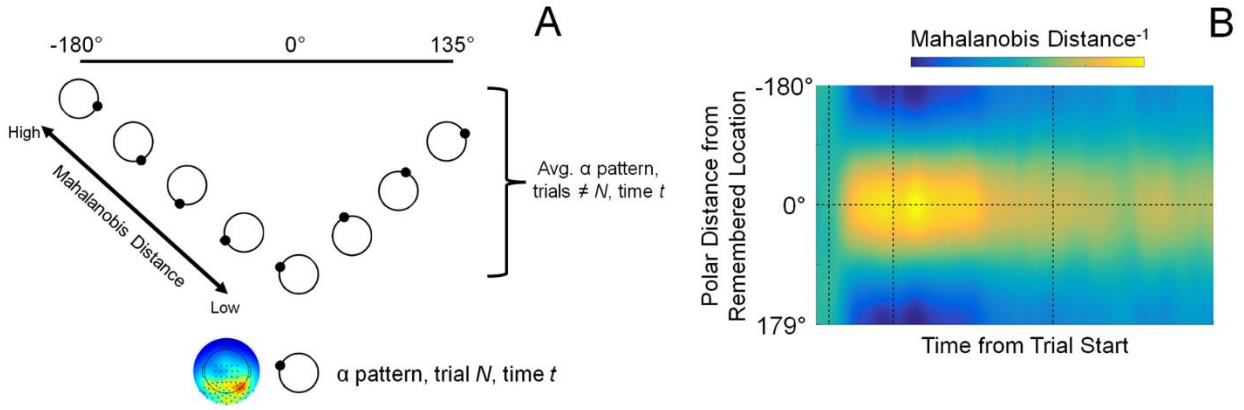

Figure 3. Decoding Analysis*. (A) We computed the multivariate (Mahalanobis) distance between patterns of alpha-band activity measured during each trial of a training data set with position-averaged patterns measured in an independent training data set. (B) If patterns of alpha-band activity contain robust position-specific signals, then multivariate distances should be smallest when comparing activation patterns associated with memory for a specific location in the test data set with activation patterns associated with memory for the same location in the training data set, and the multivariate distance between test and training data patterns should increase with polar distance. (N = 97)*

complete description of this analysis can be found in the accompanying STAR methods, but briefly we decoded the location of each disc by comparing trial-wise multielectrode patterns of occipitoparietal alpha-band (8-13 Hz) activity in a test data set with position-averaged patterns in an independent training data set. A key assumption of this approach is that patterns of activity associated with memory for neighboring positions in the training and test data sets should be more similar than patterns of activity associated with memory for distal positions. Thus, multivariate distances should be smallest when comparing activation patterns associated with memory for the same location in the training and test data sets, and multivariate distance should increase with the polar distance between the compared locations (e.g., Fig 3A). We therefore computed the multivariate (Mahalanobis) distance between activity patterns measured during each trial and time point of a test data set with averaged activity patterns associated with memory for all possible stimulus positions in an independent training

data set, yielding an *n* location x 1 vector of distance estimates. Trial-wise distance estimates were circularly shifted to a common center (0°, by convention) and averaged. If the measured EEG signal contains robust location information, this analysis should reveal an inverted “tuning-curve” like function over time (Fig 3B). This function was inverted and convolved with a cosine function, yielding a single decoding estimate for each test trial and timepoint. This entire analysis was performed separately for each disc (i.e., red vs. blue) for each cue condition (i.e., informative vs. uninformative) and task type (i.e., pro- vs. anti-cue), yielding a series of condition-specific decoding time series for each participant. To facilitate comparisons across cue conditions and tasks participant-level decoding time series were sorted by task relevance: during the pro-cue task decoding performance for the cue-matching disc was designated task-relevant and decoding performance for the cue-nonmatching disc was designated task-irrelevant; during the anti-cue task this mapping was reversed. Finally, we verified that our key findings (e.g., Figs 5-7) were not dependent on this specific decoding approach; an inverted encoding model analysis (STAR Methods) yielded qualitatively similar results (Fig S2).

Our pro-cue task is a direct replication of two earlier studies that reported rapid increases in task-relevant location decoding performance following the appearance of an informative color cue (Ester et al., 2018; Nouri & Ester, 2020). Thus, we expected to replicate these findings in the current study (Fig 4A). Next, we considered two models

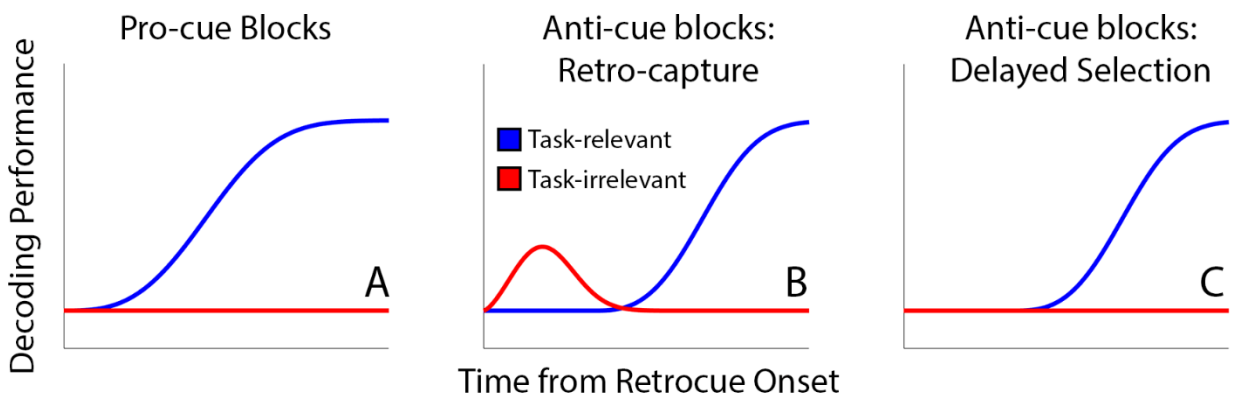

Figure 4. Selection Models. *Each panel depicts a hypothetical retrocue-locked decoding time course for task-relevant and task-irrelevant stimuli. Task relevance is defined by the relationship between the color of the retrospective cue and the position of the color-matching disc. For example, during pro-cue blocks the cue-matching disc is task-relevant while the cue-nonmatching disc is task-irrelevant; during anti-cue blocks the cue-nonmatching disc is task-relevant while the cue-matching disc is task-irrelevant. (N = 69)*

describing how exogenous mechanisms might influence the selection of task-relevant WM content during the anti-cue task (Fig 4). The first model was motivated by studies reporting exogenous capture but salient but task-irrelevant stimuli in the external attention literature (e.g., Theeuwes, 1992; Folk et al., 1992) and neurocomputational models of WM storage predicated on conjunctive coding (Schneegans & Bays, 2017; 2018) or pattern completion (e.g., Mongillo et al., 2008; Manohar et al., 2019). A capture-based account of internal selection holds that exogenous factors prompt reflexive selection of WM contents that incidentally match the anti-cue; conjunctive coding and pattern completion models predict that the anti-cue should prompt a “re-activation” of spiking or synaptic mechanisms responsible for storing color-location bindings in WM. Both accounts predict a transient increase in decoding performance for the task-irrelevant but cue-nonmatching item during the anti-cue task (Fig 4B). We used this prediction as a benchmark to evaluate an alternative model where exogenous mechanisms delay the selection of task-relevant WM content by endogenous

mechanisms (Fig 4C). This model predicts that competition between endogenous and exogenous factors delays the selection of task-relevant WM content, yielding an increase in task-relevant decoding performance that begins significantly later during the anti-cue task (where endogenous and exogenous factors compete to control selection) compared to the pro-cue task (where endogenous and exogenous factors are aligned).

Competition Between Endogenous and Exogenous Mechanisms Delays the Selection of Task-Relevant WM Content

Our experimental task (Fig 1A) was deliberately constructed so that the effects of endogenous and exogenous factors on the selection of WM contents could be measured during informative *and* uninformative trials. For example, during uninformative trials participants received an uninformative retrospective cue instructing them to remember the positions of both discs. Upon presentation of the probe display, this uninformative cue was replaced by a 100% valid (pro-) or 0% valid (anti-) cue instructing participants which disc to report. Conversely, during informative trials pro- and anti-cues were presented midway through the storage period. Since informative and uninformative trials had different response demands (i.e., pro- and anti-cues presented at the end of uninformative trials required an immediate response while pro- and anti-cues presented during the memory delay during informative trials did not), we analyzed data from these conditions separately.

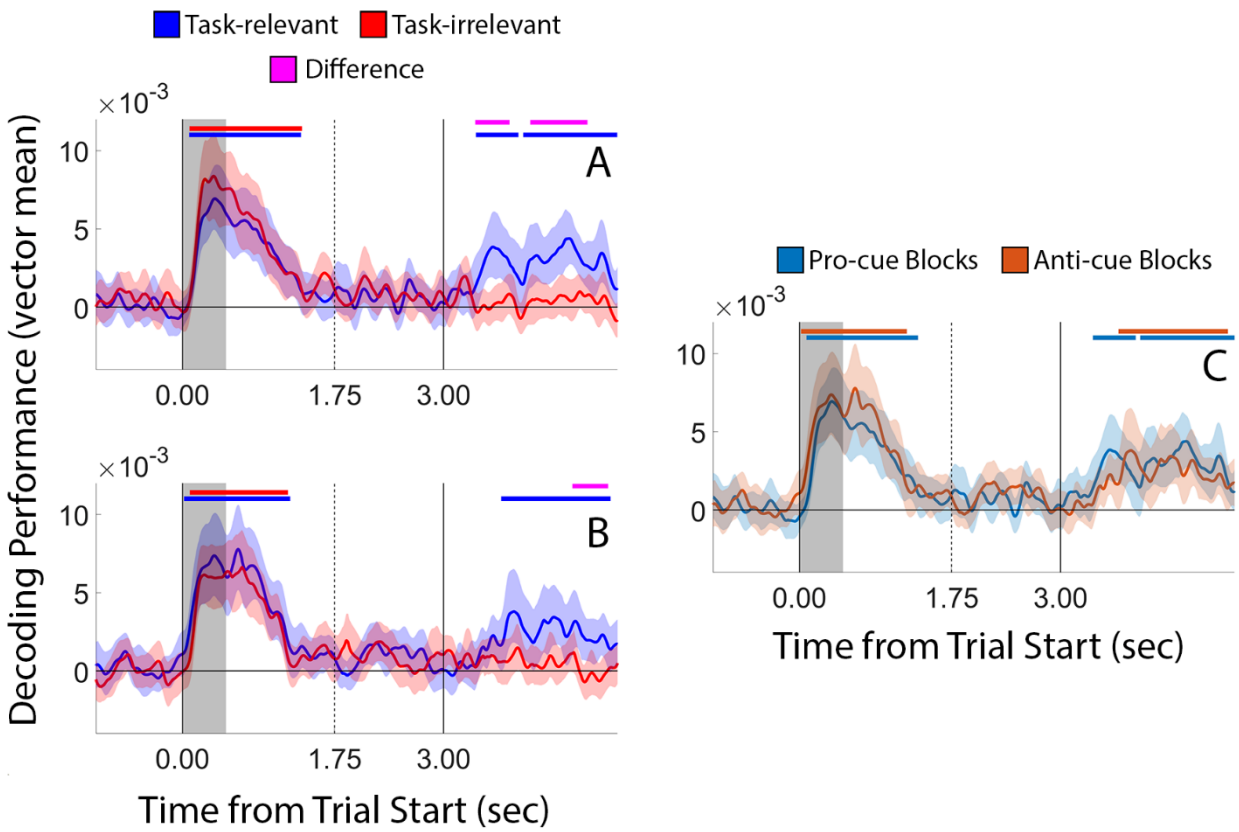

Figure 5. Location Decoding Performance During Neutral Trials. *(A, B) Decoding performance for task-relevant and task-irrelevant locations during pro-cue and anti-cue blocks, respectively. (C) Overlay of task-relevant location decoding performance for pro-cue and anti-cue blocks (i.e., the blue lines in panels A and B). Solid vertical lines at time 0.00 and 3.00 depict the onset of the sample and probe displays, respectively. The dashed vertical line at time 1.75 depicts the onset of the (uninformative) retrocue. Gray shaded region spanning 0.00-0.50 marks the duration of the sample display. Horizontal bars at the top of each plot mark intervals where decoding performance was significantly greater than zero (nonparametric cluster-based randomization test; see Methods) or intervals where decoding performance for one location was significantly greater than decoding performance for the other location. Shaded regions around each line depict bootstrapped confidence intervals of the mean. (N = 142)*

We first considered data from uninformative trials (Fig 5). Task-relevant and task-irrelevant location decoding performance in the pro-cue (Fig 5A) and anti-cue (Fig 5B) tasks increased rapidly during the sample display but returned to chance levels by the time the (uninformative) retrocue was presented 1.75 sec later. Task-irrelevant decoding performance remained at chance levels through the retrocue and probe displays while task-relevant decoding performance increased from chance- to above-

chance levels during the probe period. Visual comparisons of probe-locked task-relevant decoding performance suggested that above-chance decoding performance was reached earlier during the pro- relative to the anti-cue task (Fig 5C). To quantify this effect, we extracted and compared probe-locked task-relevant decoding time courses during the pro- and anti-cue tasks via cross-correlation. Specifically, we computed correlations between the timeseries of task-relevant decoding performance during the pro- and anti-cue tasks while temporally shifting the former by -1.0 to +1.0 sec in 4 msec intervals relative to the latter, yielding a correlation-by-lag function. Next, we compared the observed cross-correlation function to a null distribution computed by randomizing participant-level task labels (i.e., pro vs. anti-cue) 10,000 times. Thus, correlations larger than those expected by chance at negative lag values and correlations smaller than those expected by chance at positive lag values indicate that task-relevant decoding performance reaches above-chance levels earlier during the pro- vs. anti-cue task. Observed cross-correlation coefficients (Fig 7A) exceeded those expected by chance over lags spanning -0.33 to -0.22 sec and fell below those expected by chance over a period spanning +0.25 to +0.35 sec, confirming that task-relevant decoding performance reached above chance levels earlier during the pro- vs. anti-cue task.

A complementary analysis of cue-locked decoding performance during informative trials yielded similar conclusions (Fig 6; see also Fig S2). Here, task-relevant (but not task-irrelevant) decoding performance reached above-chance levels during the retrocue period during both the pro- and anti-cue tasks (Figs 6A and 6B,

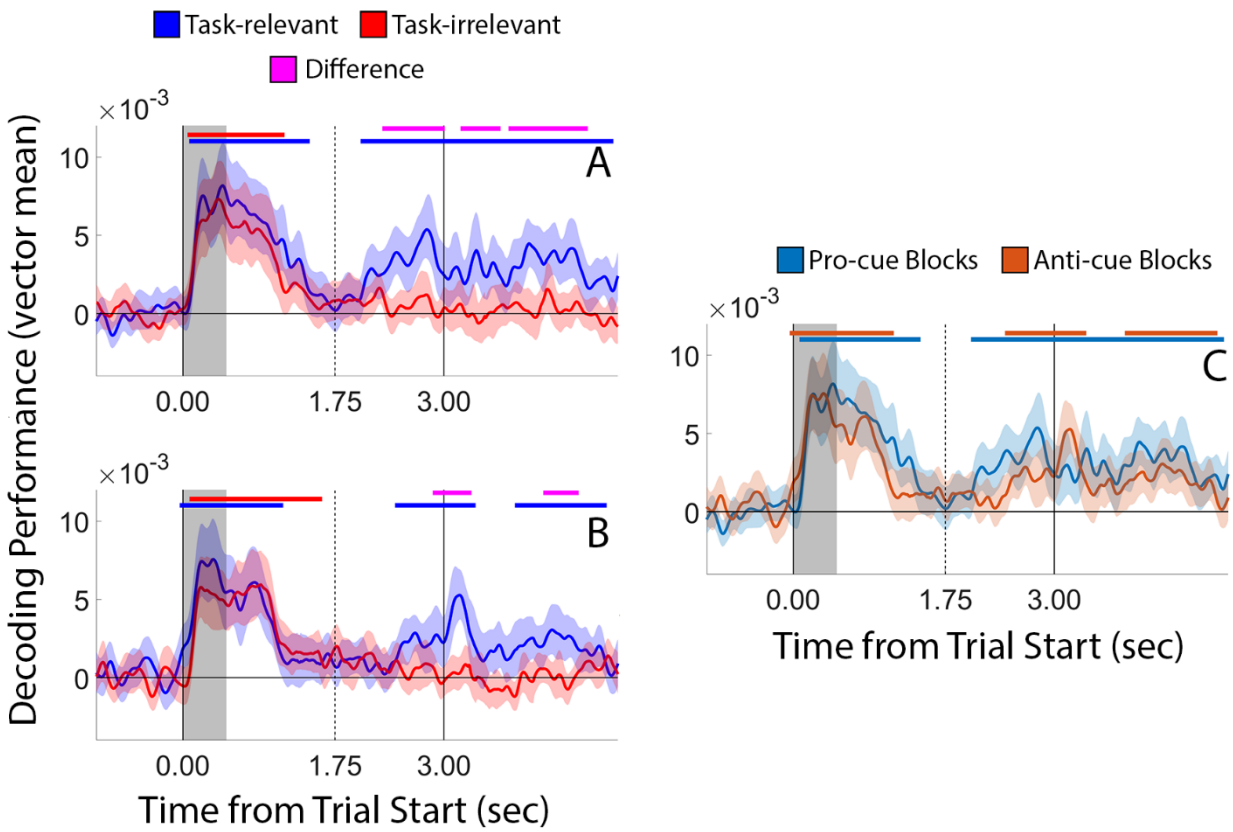

Figure 6. Location Decoding Performance During Informative Trials. *(A, B) Decoding performance for task-relevant and task-irrelevant locations during pro-cue and anti-cue blocks, respectively. (C) Overlay of task-relevant location decoding performance for pro-cue and anti-cue blocks (i.e., the blue lines in panels A and B). Conventions are identical to Fig 4. (N = 50)*

respectively). Visual comparisons of cue-locked decoding performance suggested that above-chance decoding performance was reached earlier during the pro- vs. anti-cue task (Fig 6C). This effect was confirmed by a cross-correlation analysis (Fig 7B): we observed cross-correlation coefficients exceeding those expected by chance over lags spanning -0.44 to -0.35 sec and falling below those expected by chance over lags spanning -0.14 to +0.13 and +0.28 to +0.70 sec.

Alternative Explanations

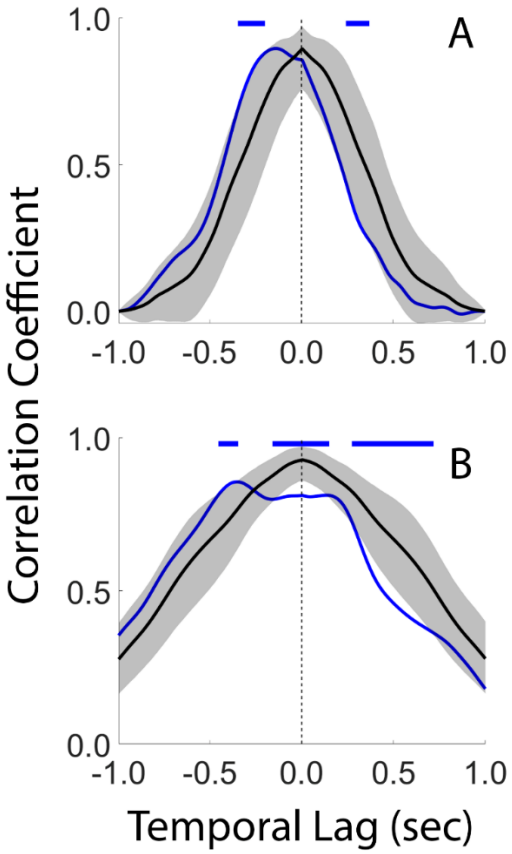

Figure 7. Cross-correlation Results. *(A) Cross-correlation between the time series of task-relevant location decoding during the response period of neutral pro-cue and neutral anti-cue trials. Peaks over negative values depict epochs where the time series of task-relevant decoding performance during pro-cue blocks leads the time series of decoding performance during anti-cue blocks. (B) An identical analysis was applied to decoding time series measured during the cue-to-probe interval of informative pro-cue and anti-cue trials. The black line and shaded area depict the mean and 95% confidence interval of cross-correlation coefficients computed by randomizing time series condition labels (i.e., pro- vs. anti-cue block) 10,000 times. Blue horizontal bars at the top of the plot depict intervals where observed cross-correlations exceeded the 95th or 2.5th percentile of the null distribution. (N = 127)*

The first model of internal exogenous attention we considered (Fig 4B) predicts that exogenous factors prompt either a reflexive selection of the cue-matching but task-irrelevant disc or an automatic refreshing of position information during the anti-cue task. This prediction is difficult to reconcile with the findings summarized in Figs 5-7: during the anti-cue task, we observed no evidence for above-chance task-irrelevant decoding performance during uninformative (Fig 5B) or informative trials (Fig 6B). However, evidence for retrospective capture by the cue-matching but task-irrelevant disc may have been obscured by trial averaging. For example, perhaps the retrocapture effect is small, short, lived, or intermittent (i.e., occurring on only a subset of trials). To examine these possibilities in greater detail, we recomputed task-irrelevant decoding performance during the anti-cue task after sorting participants' data by median recall error (i.e., "high" vs. "low"), reasoning that since exogenous factors have a deleterious effect on participants' recall errors (Fig 1C),

exogenous selection of the task-irrelevant disc – as indexed by higher task-irrelevant decoding performance – should be more evident during high recall error trials. However, this was not the case: we observed no evidence for above-chance task-irrelevant decoding performance during low- or high-error informative (Fig S3A) or uninformative trials (Fig S3B). Thus, it is unlikely that the pattern of exogenous-then-endogenous selection predicted by the retro-capture model was obscured by trial-averaging.

The delayed selection model (Fig 4C) predicts that endogenous and exogenous selection compete to control internal selection during anti-cue trials, and that selection proceeds only once this competition has been resolved. The data reported here are nominally consistent with this prediction: task-relevant decoding performance reached above-chance levels during the pro-cue task (i.e., when endogenous and exogenous factors are aligned) earlier than during the anti-cue task (i.e., when endogenous and exogenous factors were opposed; Figs 5C, 6C; 7). We considered several alternative explanations for this pattern. For example, one trivial possibility is that that it simply takes participants longer to interpret the anti-cues. However, this explanation is difficult to reconcile with the fact that neither the main effect of task (i.e., pro-cue vs. anti-cue; Fig 1D) nor the simple effect of task (Fig 1E) on response times during informative cue trials reached significance. A second possibility is that delayed above-chance decoding performance during the anti-cue task was caused by carryover effects. For example, although task order was counterbalanced across observers, perhaps participants who completed the pro-cue task followed by the anti-cue task had extra difficulty interpreting anti-cues compared to participants who performed the anti-cue task followed by the pro-cue task. To test this possibility, we compared the time-courses of task-relevant

decoding performance during informative anti-cue trials in participants who performed the pro-cue task followed by the anti-cue task (N = 17) or vice versa (N = 23). For both groups, task-relevant decoding performance reached above chance levels shortly before or immediately after the onset of the probe display (Fig S4). If anything, the onset of above-chance decoding performance occurred later for participants who performed the anti-cue task second vs. those who performed the anti-cue task first, though this difference was not significant (p = 0.146; randomization test, see STAR Methods). Thus, order effects cannot account for delays in task-relevant decoding performance during the anti-cue vs. pro-cue blocks.

Finally, we considered the hypothesis that the selection of cue-matching but task-irrelevant positions during the anti-cue task were obscured by successful cognitive control. Specifically, we reasoned that selection of the task-irrelevant position might be more likely during trials where participants "spaced out" compared to trials when they were actively engaged in the task. To test this hypothesis, we re-computed cue-matching but task-irrelevant location decoding performance after sorting participants' alpha-band EEG data by frontal theta power (Fig 2), reasoning that inadvertent selection of cue-matching but task-irrelevant stimuli would be more likely during trials where frontal theta power was low vs. high. However, we observed no evidence for above-chance

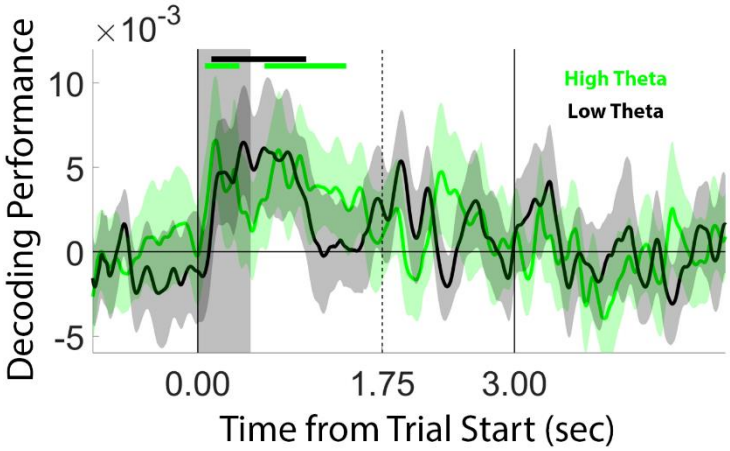

Figure 8. Task-Irrelevant Decoding Performance During the Anti-cue Task as a Function of Frontal Theta Power. Conventions are identical to those in Figs 5-6.

decoding of the cue-matching but task-irrelevant position during high- or low-theta power trials (Fig 8). This analysis provides converging evidence suggesting that exogenous factors do not lead to a selection or re-activation of cue-matching but task-irrelevant WM content, but instead delay the endogenous selection of task-relevant WM content.

Discussion

Selective attention can be allocated to sensory inputs and internal representations based on voluntary, endogenous factors or involuntary, exogenous factors. An enormous literature suggests that external selection is mediated by competition between endogenous and exogenous factors, with the focus of selection determined by the balance of competition between these factors (e.g., Desimone & Duncan, 1995; Corbetta & Shulman, 2002; Connor et al., 2004). Here, we show that exogenous factors influence internal selection by a different mechanism: delayed selection. Capitalizing on recent findings (e.g., Ester et al., 2018; Nouri & Ester, 2020) we used EEG to decode the contents of spatial WM after a 100% informative or 0% informative retrospective cue (i.e., pro- and anti-cues, respectively). We found no evidence for exogenous selection or refreshing of cue-matching but task-irrelevant WM content during the anti-cue task. Instead, the selection of task-relevant WM content was simply delayed by ~250 ms during the anti-cue relative to the pro-cue task. Thus, the selection of task-relevant WM content is delayed until competition between endogenous and exogenous factors is resolved.

The current findings have major implications for models of external and internal attentional selection. External attention is directed to sensory inputs, and competition between endogenous and exogenous factors sometimes result in attention “capture” by unexpected, salient, and valuable but otherwise task-irrelevant stimuli (Theeuwes, 1992; Folk et al., 1992; Bacon & Egeth, 1994; Folk et al., 2009; Fukuda & Vogel, 2009; Anderson et al., 2011). Endogenous and exogenous factors also compete to control the selection of internal representations (e.g., van Ede et al., 2020; van Ede & Nobre,

2022), but this competition is resolved in a fundamentally different way. We observed no evidence for the selection of salient internal representations akin to capture effects seen in external attention. Instead, we found that the selection of task-relevant internal representations was delayed when endogenous and exogenous selection mechanisms were placed in conflict compared to when they were aligned. This finding supports claims that the selection of external sensory inputs and internal memory representations are mediated by different mechanisms (Chun et al., 2011; van Ede & Nobe, 2022).

The current findings also have implications neurocomputational models of WM based on conjunctive coding (e.g., Schneegans & Bays, 2017; 2018) or synaptic signaling (e.g., Manohar at al. 2019; Li et al., 2021). On the one hand, conjunctive coding models predict that WM representations are maintained by spiking activity in feature- and/or location-specific neural populations. While the exact mechanisms vary by implementation (see Scheegans & Bays, 2017 and Oberauer & Lin, 2017 for representative examples), these models generally predict that a feature probe in one dimension (e.g., orientation) activates spiking patterns in neural populations that code this feature and those that code other features of the same object (e.g., color) and/or its location. This, in turn, enables robust read-out of the probed and non-probed stimulus dimensions by downstream neural populations. While these models were not developed to describe the anti-cue task contemplated here, one could reasonably predict an increase in task-irrelevant decoding performance after presentation of an anti-cue based on their general architecture. We observed no evidence for such an effect, and it remains to be seen whether these models can be modified to predict behavioral and neural data during pro- and anti-cue tasks.

On the other hand, pattern completion models predict that the contents of WM reside in different neural states – an “active” state mediated by sustained spiking activity and a “latent” state mediated by short-term synaptic plasticity. Presentation of a feature probe that matches a stimulus stored in a latent format reinstates activity patterns evoked when that stimulus was encoding, prompting and/or “refreshing” of the neural representation of the probe-matching item through pattern completion (Mongillo et al., 2008; Manohar et al., 2019). This prediction enjoys some support: a representation stored in WM item can be “re-activated” (as indexed by above-chance EEG decoding performance) by a task-irrelevant sensory input (Wolff et al., 2017) or a transcranial magnetic stimulation (TMS) pulse applied over sensory cortex (Rose et al., 2016). Conversely, in the current study we found no evidence for a reactivation of cue-matching but task-irrelevant WM content following presentation of an anti-cue. However, one salient difference between the current study and prior work is that in the latter, an informative retrospective cue instructed presented prior to the “impulse” stimulus instructed participants which of two remembered stimuli should be prioritized for report. Thus, re-activation of information stored in synaptic traces may be contingent on the network responsible for storing information to be selected or otherwise primed for decision making and action.

A recent study reported that the onset of small-but-robust biases in gaze position (referred to hereafter as microsaccades) towards the memorized location of a task-relevant stimulus were delayed following the appearance of an anti- vs. pro-cue (van Ede et al., 2020). The authors of this study speculated that delays in selecting task-relevant information during an anti-cue task were caused by an automatic “refreshing” of

cue-matching but task-irrelevant WM content. However, as the authors of this study acknowledged, this hypothesis cannot be substantiated based on microsaccades alone. First, although attention-related modulations of cortical and subcortical processing are larger during trials containing microsaccades towards the location of a (covertly) attended stimulus, clear attention-related modulations are also observed in the absence of microsaccades (Yu et al., 2022; Liu et al., 2022; though see Lowet et al. 2018 for a contradictory result). Second, any of the three models examined in the current study (i.e., retro-capture, automatic refreshing, or delayed selection) could have produced the delays in microsaccade biases reported by van Ede et al. (2020). Although we did not (and could not, lacking a functional eyetracker at the time this study was conducted) record eye position with sufficient spatiotemporal resolution needed to replicate the findings reported by van Ede et al., we observed no evidence for an enhancement of cue-matching but task-irrelevant position information during the anti-cue task that would be predicted by the account endorsed by van Ede and colleagues (2020). Thus, we speculate that the delays in microsaccade biases reported by these authors were generated by competition between endogenous and exogenous factors. Further research could compare the time-courses of oculomotor and neural signals associated with internal selection to test this possibility.

To summarize, the current findings support recent suggestions that endogenous and exogenous selection mechanisms compete to control access to internal WM representations. However, this competition is resolved in a fundamentally different way than seen during external attention. Specifically, endogenous and exogenous competition does not produce an errant selection or refreshing of salient but task-

irrelevant WM content. Instead, this competition delays the selection of task-relevant memory content by endogenous mechanisms. This result reveals a fundamental difference between the operation of external and internal attentional selection, and suggests that they are mediated by fundamentally different mechanisms.

Methods

Participants. 42 human volunteers (both sexes) participated in a single 2.5-hour testing session. Participants were recruited from the Florida Atlantic University community via campus advertisements and remunerated at $15/h in Amazon.com gift cards. All participants gave both written and oral informed consent in compliance with procedures established by the local institutional review board, and all participants self-reported normal or corrected-to-normal visual acuity. Two participants voluntarily withdrew from the study prior to completing both cue conditions (i.e., pro-cue vs. anti-cue); data from these participants were excluded from final analyses. Thus, the data reported here reflect the remaining 40 participants.

Testing Environment. Participants were seated in a dimly-lit and acoustically shielded recording chamber for the duration of the experiment. Stimuli were generated in MATLAB and rendered on a 17'' Dell CRT monitor cycling at 85 Hz (1024 x 768 pixel resolution) via PsychToolbox3 software extensions (Kleiner et al., 2007). Participants were seated approximately 65 cm from the display (head position was unconstrained). To combat fatigue and/or boredom, participants were offered short breaks at the end of each experimental block.

Spatial Retrocue Task. A task schematic is shown in Fig 1A. Each trial began with the presentation of an encoding display lasting 500 ms. The encoding display contained two colored circles (blue and red; subtending 1.75° visual angle from a viewing distance of 65 cm) rendered at of eight polar locations (22.5° to 337.5° in 45° increments) along the perimeter of an imaginary circle (radius 7.5° visual angle) centered on a circular fixation point (subtending 0.25°) rendered in the middle of the display. The locations of the two

discs were counterbalanced across each task (i.e., pro-cue vs. anti-cue), though not necessarily within an experimental block. Participants were instructed to maintain fixation and refrain from blinking for the duration of each trial.

The sample display was followed by a 1.25 sec blank display and a 1.25 sec retrocue display. Retrocues were defined by a change in the color of the fixation point; during informative trials the fixation point changed colors from black to either blue or red (i.e., matching the color of a remembered disc). During neutral trials, the fixation point changed colors from black to purple (the “average” of blue and red). At the end of the trial, a response display containing a mouse cursor, “?” symbol, and a probe circle appeared. Participants were required to report the location of a remembered disc by clicking along the perimeter of the probe circle (during neutral trials, the fixation point changed colors from purple to either blue or red during the response display, indicating which item should be reported). Participants were instructed to prioritize accuracy over speed, and no response deadline was imposed. The trial terminated as soon as the participant clicked on a location. Sequential trials were followed by a 1.5-2.5 sec blank period (randomly and independently selected from a uniform distribution after each trial).

During the first half of the experiment (e.g., experimental blocks 1-8), each participant was assigned to the pro-cue or anti-cue task. In the pro-cue task, participants were informed that during informative would be required to click on the location of the disc matching the retrocue color. Conversely, in the anti-cue task participants were informed that during informative trials they would be required to click on the location of the disc that did not match the retrocue color. Participants completed

eight blocks of 56 trials in the both the pro-cue and anti-cue tasks. Task order (i.e., eight blocks of the pro-cue task followed by eight blocks of the anti-cue task or vice versa) was counterbalanced across participants.

EEG Acquisition and Preprocessing. Continuous EEG was recorded from 63 uniformly distributed scalp electrodes using a BrainProducts “actiCHamp” system. The horizontal and vertical electrooculogram (EOG) were recorded from bipolar electrode montages placed over the left and right canthi and above and below the right eye, respectively. Live EEG and EOG recordings were referenced to a 64th electrode placed over the right mastoid and digitized at 1 kHz. All data were later re-referenced to the algebraic mean of the left- and right mastoids, with 10-20 site TP9 serving as the left mastoid reference.

Data preprocessing was carried out via EEGLAB software extensions (Delorme & Makeig, 2004) and custom software. Data preprocessing steps included the following, in order: (1) resampling (from 1 kHz to 250 Hz), (2) bandpass filtering (1 to 50 Hz; zero-phase forward- and reverse finite impulse response filters as implemented by EEGLAB), (3) epoching from -1.0 to +5.0 relative to the start of each trial, (4) identification, removal, and interpolation of noisy electrodes via EEGLAB software extensions, and (5) identification and removal of oculomotor artifacts via independent components analysis as implemented by EEGLAB. After preprocessing, location decoding analyses focused exclusively on the following 10-20 occipitoparietal electrodes: P7, P5, P3, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO2, PO4, PO8, O1, O2, Oz.

Data Cleanup. Prior to analyzing participants’ behavioral or EEG data, we excluded all trials where the participant responded with a latency of < 0.4 sec (we attributed these trials to accidental mouse clicks following the onset of the probe display rather than a

deliberative recall of a specific stimulus position) and more than 3 standard deviations above the average response time across all experimental conditions. This resulted in an average (±1 S.E.M.) loss of 14.43 ±0.93 trials (or 1.67% ± 0.11% of trials) across participants but had no qualitative effect on any of the findings reported here.

Decoding Spatial Positions from Posterior Alpha-Band EEG Signals. Location decoding was based on the multivariate distance between EEG activity patterns associated with memory for specific positions. This approach is an extension of earlier parametric decoding methods (Wolff et al., 2017) designed for use in circular feature spaces (e.g., Fig 3A). Following earlier work (e.g., Ester et al., 2018), we extracted spatiotemporal patterns of alpha-band activity (8-13 Hz) from 17 occipitoparietal electrode sites (see *EEG Acquisition and Preprocessing* above). The raw timeseries at each electrode was bandpass filtered from 8-13 Hz (zero-phase forward-and-reverse filters as implemented by EEGLAB software), yielding a real-valued signal f(t). The analytic representation of f(t) was obtained via Hilbert transformation:

$$z(t) = f(t) + if(t)$$

where *i* is the imaginary operator and *if(t)* = $A(t)e^{i\varphi(t)}$. Alpha power was computed by extracting and squaring the instantaneous amplitude A(t) of the analytic signal z(t).

Location decoding performance was computed separately for each disc (i.e., blue vs. red), trial type (i.e., informative vs. neutral) and each task (i.e., pro-cue vs. anti-cue) on a timepoint-by-timepoint basis. In the first phase of the analysis, we sorted data from each condition into 10 unique training and test data sets using stratified sampling while

ensuring that each training set was balanced across remembered positions (i.e., we ensured that each training data set contained an equal number of observations where the location of the remembered stimulus was at 22.5°, 67.5°, etc.). We circularly shifted the data in each training and test data set to a common center (0°, by convention) and computed trial-averaged patterns of responses associated with memory for each disc position in each training data set. Next, we computed the Mahalanobis distance between trial-wise activation patterns in each test data set with position-specific activation patterns in the corresponding test data set, yielding a location-wise set of distance estimates. If scalp activation patterns contain information about remembered positions then distance estimates should be smallest when comparing patterns associated with memory for identical positions in the training and test data set and largest when comparing opposite positions (i.e., those ±180° apart), yielding an inverted gaussian-shaped function (e.g., Fig 3B). Trial-wise distance functions were averaged and sign-reversed for interpretability. Decoding performance was estimated by convolving timepoint-wise distance functions with a cosine function, yielding a metric where chance decoding performance is equal to 0. Decoding results from each training- and test-data set pair were averaged (thus ensuring the internal reliability of our approach), yielding a single decoding estimate per participant, timepoint, and task condition.

Cross-correlation Analyses. Temporal differences in task-relevant location decoding performance were estimated via cross-correlation analyses. For neutral trials, we extracted task-relevant decoding performance during pro- and anti-cue blocks over a period spanning 0.0 to 1.0 seconds following the onset of the probe display (i.e., when

an informative cue instructed participants which disc to recall). We computed correlation coefficients between pro- and anti-cue decoding time series while systematically shifting the pro-cue time series from -1.0 to +1.0 sec relative to the anti-cue decoding time series (Fig 7A, blue line). We compared these correlation coefficients to a distribution of correlation coefficients computed under the null hypothesis (i.e., no systematic difference in pro- and anti-cue decoding time series) by repeating the same analysis 10,000 times while randomizing the decoding condition labels (i.e., pro- vs. anti-cue) for each participant (Fig 7A, black line and shaded area). An identical analysis was performed on task-relevant pro- and anti-cue decoding task performance from 0.0 to 1.75 during informative trials (Fig 7B). We deliberately selected a longer temporal interval for analysis during informative trials as we expected increases in pro- and anti-cue decoding performance to begin during the retrocue period and persist into the ensuing response period.

Quantifying Frontal Theta Power. Analyses of frontal theta power focused on informative trials from the pro- and anti-cue tasks. The raw timeseries at each scalp electrode was bandpass filtered from 4-7 Hz (zero-phase forward-and-reverse filters as implemented by EEGLAB software), yielding a real-valued signal f(t). The analytic representation of this signal was obtained via Hilbert transformation, and theta power was computed by extracting and squaring the instantaneous amplitude A(t) of the analytic signal z(t). Topographic maps of theta power during the pro- and anti-cue tasks (Figs 2A, S1A, S1B) were obtained by averaging power estimates over trials and a temporal window spanning 2.5 to 3.0 sec following the start of each informative trial (i.e., 750 to 1250 ms after cue onset). Based on these maps, we limited further analyses

to power estimates measured at four frontal electrode sites: AFz, Fz, F1, and F2. Data from these electrodes were used to compute time-resolved estimates of theta power during the pro- and anti-cue tasks (Fig S1C) and task differences in theta power (Fig 2B). In a final analysis, we extracted and computed trial-wise estimates of theta power during the anti-cue task (using the same electrodes and temporal window described in the previous paragraph). We sorted participants' anti-cue EEG data into low- and high-theta power groups after applying a media split to power estimates, then decoded the location of the cue-matching but task-irrelevant item within each group (Fig 8). This allowed us to test whether evidence for exogenous capture and/or "refreshing" of cue-matching memory traces was more likely to occur on trials with low- vs. high theta power.

Inverted Encoding Model. To verify the generality of our findings across analytic approaches, we reconstructed position-specific WM representations from spatiotemporal patterns of alpha-band activity using an inverted encoding model (Fig S2). Our approach was conceptually and quantitatively identical to that used in earlier studies (Foster et al., 2016; Ester et al., 2018). We modeled alpha power at each scalp electrode as a weighted sum of eight location-selective channels, each with an idealized tuning curve (a half-wave rectified cosine function raised to the 8th power, with the maximum response of each function normalized to 1). The predicted responses of each channel during each trial were arranged in a *k* channels by *n* trials design matrix C. The relationship between the EEG data and the predicted responses in *C* is given by a general linear model of the form:

$$B = WC + N$$

where B is an m electrode by n trial training data matrix, W is an m electrode by k channel weight matrix, and N is a matrix of residuals (i.e., noise).

To estimate W, we constructed a training data set containing an equal number of trials for each stimulus position (i.e., 22.5-337.5° in 45° increments). We identified the location φ with the fewest r repetitions and constructed a training dataset B_{trn} (m electrodes by n trials) and weight matrix C_{trn} (n trials by k channels) by randomly selecting (without replacement) 1 to r trials for each of the eight possible stimulus positions. The training dataset was used to compute a weight for each channel C_i using ordinary least-squares estimation:

$$V_i = \frac{\sum_i^{-1} W_i}{W_i^T \sum_i^{-1} W_i}$$

where T and -1 are the matrix transpose and inversion operations, respectively. Σ_i is the regularized noise covariance matrix for each channel i, estimated as:

$$\sum_i \frac{1}{n-1} \varepsilon_i \varepsilon_i^T$$

where n is the number of training trials and ε_i is a matrix of residuals:

$$\varepsilon_i = B_{trn} - W_i C_{trn,i}$$

Estimates of ε_i were obtained by regularization-based shrinkage using an analytically determined shrinkage observation. In this way, an optimal spatial filter v_i was estimated for each channel C_i, yielding an m electrode by k filter matrix V.

Next, we constructed a test dataset B_{tst} (m electrodes by n trials) containing data from all trials not included in the training data set and estimated trial-by-trial channel responses C_{tst} (k channels by n trials):

$$C_{tst} = V^T B_{tst}$$

Trial-wise channel response estimates were interpolated to 360°, circularly shifted to a common center (0°, by convention), and averaged, yielding a single reconstruction per participant, time point, cue condition (i.e., informative vs. uninformative) and task (i.e., pro vs. anti-cue). Condition-wise channel response functions were averaged, converted to polar form, and projected onto a vector with angle 0°:

$$r = |z|\cos(\arg(z)), \qquad z = ce^{2i\rho}$$

Where *c* is a vector of estimate channel responses and *ρ* is a vector of angles at which the channels peak. To ensure internal reliability, this entire analysis was repeated 100 times, and unique (randomly chosen) subsets of trials were used to define the training and test data sets during each iteration. The results were then averaged across permutations to obtain a single reconstruction strength estimate for each participant, task condition, and timepoint.

Statistical Comparisons – Behavioral Data. Participants' behavioral data (i.e., absolute average recall error and average response time; Fig 1B-E) were analyzed using standard repeated-measures parametric statistics (e.g., t-test, ANOVA); for these comparisons we report test statistics, p-values, and effect size estimates.

Statistical Comparisons – Decoding Performance & Inverted Encoding Model. The decoding analysis and inverted encoding model we used assume chance-level performance of 0. Likewise, direct comparisons of decoding performance or reconstruction strength across conditions (e.g., pro-cue vs. anti-cue) assume null statistics of 0. Thus, we evaluated task-relevant and task-irrelevant decoding performance (Figs 5-6, 8, S1-S4) by generating null distributions of decoding performance (or differences in decoding performance across conditions) by randomly inverting the sign of each participant's data with 50% probability and averaging the data

across participants. This procedure was repeated 10,000 times, yielding a 10,000-element null distribution for each time point. Finally, we implemented a cluster-based permutation test (Maris & Oostenveld, 2007; Wolff et al. 2017) with cluster-forming and cluster-size thresholds of $p < 0.05$ to correct for multiple comparisons across time points.

References

Anderson BA, Laurent PA, Yantis S. Value-driven attentional capture. Proceedings of the National Academy of Sciences USA, 108:10367-10371 (2011)

Bacon WF, Egeth HE. Overriding stimulus-driven attentional capture. Perception & Psychophysics, 55:485-496 (1994)

Baluch F, Ittit L. Mechanisms of top-down attention. Trends in Neurosciences, 34:210-224 (2011)

Beck VM, Luck SJ, Hollingworth A. Whatever you do, don't look at the …: Evaluating guidance by an exclusionary attentional template. Journal of Experimental Psychology: Human Perception & Performance, 44:645-662 (2018)

Berger H. Über das elektroenkephalogramm des menschen. Archiv Für Psychiatrie und Nervenkrankheiten (1929)

Bisley JW, Goldberg ME. Attention, intention, and priority in the parietal lobe. Annual Review of Neuroscience 33:1-21 (2013)

Buschman TJ, Miller EK. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315:1860-1862 (2007)

Carlisle NB, Nitka AW. Location-based explanations do not account for active attentional suppression. Visual Cognition 27:305-316 (2019)

Cavanagh JF, Frank MJ. Frontal theta as a mechanism for cognitive control. Trends in Cognitive Sciences 18:414-421 (2014)

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

Chatham CH, Badre D Multiple gates on working memory. Current Opinion in Behavioral Sciences 1:23-31 (2015)

Chun MM, Golomb JD, Turk-Browne NB. A taxonomy of external and internal attention. Annual Review of Psychology, 62:73:101 (2011)

Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience, 3:201-215 (2002)

Cosman JD, Lowe KA, Zinke W, Woodman GF, Schall JD. Prefrontal control of visual distraction. Current Biology 28:414-420 (2018)

D'Esposito M, Postle BR. The cognitive neuroscience of working memory. Annual Review of Psychology, 66:115-142 (2015)

Delorme A, Makeig S. EEGLAB: An open-source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134:9-21 (2004)

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

Ester EF, Anderson DE, Serences JT, Awh E. A neural measure of precision in visual working memory. Journal of Cognitive Neuroscience 25:754-761 (2013)

Ester EF, Nouri A, Rodriguez L. Retrospective cues mitigate information loss in human cortex during working memory storage. Journal of Neuroscience 38:8538-8548 (2018)

Folk CL, Remington RW, Johnston JC. Involuntary covert orienting is contingent on attentional control settings. Journal of Experimental Psychology: Human Perception & Performance 18:1030-1044 (1992)

Folk CL, Ester EF, Troemel K. How to keep attention from straying: Get engaged! Psychonomic Bulletin & Review 16:127-132 (2009)

Foster JJ, Sutterer DW, Serences JT, Vogel EK, Awh E. The topography of alpha-band activity tracks the content of spatial working memory. Journal of Neurophysiology 115:168-177 (2016)

Fukuda K, Vogel EK. Human variation in overriding attentional capture. Journal of Neuroscience 29:8726-8733 (2009)

Griffin IC, Nobre AC. Orienting attention to locations in internal representations. Journal of Cognitive Neuroscience 15:1176-1194 (2003)

Itthipuripat S, Wessel JR, Aron AR. Frontal theta is a signature of successful working memory manipulation. Experimental Brain Research 224:255-262 (2013)

Jensen O, Tesche CD. Frontal theta activity in humans increases with memory load in a working memory task. European Journal of Neuroscience 15:1395-1399

Johnson MR, Mitchell KJ, Raye CL, D'Esposito M, Johnson MK. A brief thought can modulate activity in extrastriate visual areas: Top-down effects of refreshing just-see visual stimuli. NeuroImage 37:290-299 (2007)

Kiyonaga A, Egner T. Working memory as internal attention: Toward an integrative account of internal and external selection processes. Psychonomic Bulletin & Review 20:228-242 (2013)

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

LaRocque JJ, Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. Decoding attended information in short-term memory: an EEG study. Journal of Cognitive Neuroscience 25:127-142. (2013)

Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. Neural evidence for a distinction between short-term memory and the focus of attention. Journal of Cognitive Neuroscience 24:61-79 (2012)

Lowet E, Gomes B, Srinivasan K, Zhou H, Schafer RJ, Desimone R. Enhanced neural processing by covert attention only during microsaccades directed towards the attended stimulus. Neuron 99:207-214 (2018)

Li J, Huang Q, Han Q, Mi Y, Luo H. Temporally coherent perturbation of neuroal dynamics during retention alters human multi-item working memory. Progress in Neurobiology 201:102023 (2021)

Liu B, Nobre AC, van Ede F. Functional but not obligatory link between microsaccades and neural modulation by covert spatial attention. bioRxiv (2022)

Luck SJ, Gaspelin N, Folk CL, Remington RW, Theeuwes J (2020) Progress toward resolving the attention capture debate. Visual Cognition 1-21 (2020)

Manohar SG, Zokaei N, Fallon SJ, Vogel TP, Husain M. Neural mechanisms of attention to items in working memory. Neuroscience & Biobehavioral Reviews, 101:1-12 (2019)

Maris E, Oostenveld R. Nonparametric statistical testing of EEG and MEG data. Journal of Neuroscience Methods 164:177-190 (2007)

Moher J, Egeth HE. The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. Attention, Perception, & Psychophysics, 74:1590-1605 (2012)

Mongillo G, Barak O, Tsodyks M. Synaptic theory of working memory. Science 319:1543-1546 (2008)

Myers NE, Stokes MG, Nobre AC. Prioritizing information during working memory: Beyond sustained internal attention. Trends in Cognitive Sciences 21:449-461 (2017)

Nouri A, Ester EF. Recovery of information from latent memory stores decreases over time. Cognitive Neuroscience 11:101-110 (2020)

Oberauer K, Lin H-Y. An interference model of visual working memory. Psychological Review, 124:21-59 (2017)

O'Reilly RC, Frank MJ. Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. Neural Computation 18:283-328 (2006)

Olivers CNL, Peters J, Houtkamp R, Roelfsema PR. Different states in visual working memory: When it guides attention and when it does not. Trends in Cognitive Sciences 15:327-334 (2011)

Pan Y, Popov T, Frisson S, Jensen O. Saccades are locked to the phase of alpha oscillations during natural reading. bioRxiv (2022)

Rac-Lubashevsky R, Frank MJ. Analogous computations in working memory input, output, and motor gating: Electrophysiological and computational modeling evidence. PLoS Computational Biology 17(6) (2021)

Remington RW, Folk CL, Mclean JP. Contingent attentional capture or delayed allocation of attention? Perception & Psychophysics 63:298-307 (2001)

Rose NS, LaRocque JJ, Riggall AC, Gosseries O, Starrett MJ, Meyering EE, Postle BR. Reactivation of latent working memories with transcranial magnetic stimulation. Science 354:1136-1139 (2016)

Scheeringa R, Bastiaansen MCM, Petersson KM, Oostenveld R, Norris DG, Hagoort P. Frontal theta EEG activity correlates negatively with the default mode network in resting state. International Journal of Psychophysiology 67:242-251 (2008)

Schneegans S, Bays PM (2017) Neural architecture for feature binding in visual working memory. Journal of Neuroscience 37:3913-3925 (2017)

Schneegans S, Bays PM (2018) New perspectives on binding in visual working memory. British Journal of Psychology, 110:207-244 (2018)

Silver MA, Ress D, Heeger DJ. Topographic maps of visual spatial attention in human parietal cortex. Journal of Neurophysiology 94:1358-1371

Souza AS, Oberauer K. In search of the focus of attention in working memory: 13 years of the retrocue effect. Attention Perception & Psychophysics 78:1839-1860 (2016)

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

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

Sprague TC, Ester EF, Serences JT. Reconstructions of information in visual spatuial working memory degrade with memory load. Current Biology 24:2174-2180 (2014)

Theeuwes J. Perceptual selectivity for color and form. Perception & Psychophysics 51:599-606 (1992)

van Ede F, Board AG, Nobre AC. Goal-directed and stimulus-driven selection of internal representations. Proceedings of the National Academy of Sciences 17:24590-24598 (2020)

van Ede F, Nobre AC. Turning attention inside-out: How working memory serves behavior. Annual Review of Psychology (2022)

Vogel EK, McCollough AW, Machizawa MG. Neural measures reveal individual differences in controlling access to working memory. Nature 438:500-503 (2005)

Wolfe JM. Visual search: How do we find what we are looking for? Annual Review of Vision Science 6:1-24 (2020)

Wolff MJ, Jochim J, Akyurek EG, Stokes MG. Dynamic hidden states underlying working-memory-guided behavior. Nature Neuroscience 20:864-871 (2017)

Wolff MJ, Kandemir G, Stokes MG, Akyurek EG. Unimodal and bimodal access to sensory working memories by auditory and visual impulses. Journal of Neuroscience 40:671-681 (2020)

Yu G, Herman JP, Katz LN, Krauzlis RJ. Microsaccades as a marker not a cause for attention-related modulation. Elife (2022)