University of Nevada, Reno

Retrospective Cue Benefits in Visual Working Memory are Limited to a Single Item at a Time

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Neuroscience

by

Ashley M. DiPuma

Dr. Edward Ester/Thesis Advisor

December, 2023

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

Ashley DiPuma

entitled

Retrospective Cue Benefits in Visual Working Memory are Limited to a Single Item at a Time

be accepted in partial fulfillment of the
requirements for the degree of

Master of Science

Edward Ester, Ph.D.
Advisor

Lars Strother, Ph.D.
Committee Member

Jennifer Hoy, Ph.D.
Graduate School Representative

Markus Kemmelmeier, Ph.D., Dean
Graduate School

December, 2023

ABSTRACT

Working memory (WM) performance can be improved by an informative cue presented during storage. This effect, termed a retrocue benefit, can be used to study limits on how human observers select and prioritize information stored in WM for behavioral output. There is disagreement about whether retrocue benefits extend to multiple WM items. One possibility is that relative to no- or neutral-cue trials multiple retrocues improve some aspects of memory performance (e.g., a reduction in random guessing) while worsening others (e.g., an increase in the probability of reporting a non-probed item). We tested this possibility in three experiments. Participants remembered arrays of four orientations or colors over a brief delay. One, two, or all four of these items were retrospectively cued, and at the end of the trial a single item was probed for recall. Participants' recall errors were lower during cue-one relative to cue-two and cue-four trials, and this benefit was driven primarily by a reduction in random guessing during cue-one trials. Moreover, recall precision, swap errors (i.e., reporting a non-probed item), and guessing rates were statistically indistinguishable across cue-two and cue-four trials. Thus, multiple simultaneously presented retrospective cues led to no performance improvement relative to an uninformative cue, providing further evidence that retrocue benefits in WM performance are limited to a single item at a time.

ACKNOWLEDGEMENTS

I want to express my sincere gratitude and appreciation to my mentor, Dr. Edward Ester. His support of my graduate education began at another university in a different state, and he has continued to be a source of guidance through the research and subsequent completion of my thesis.

I am deeply appreciative of Dr. Ester’s wisdom and grateful for his generous assistance and direction in contributing to my success.

Table of Contents

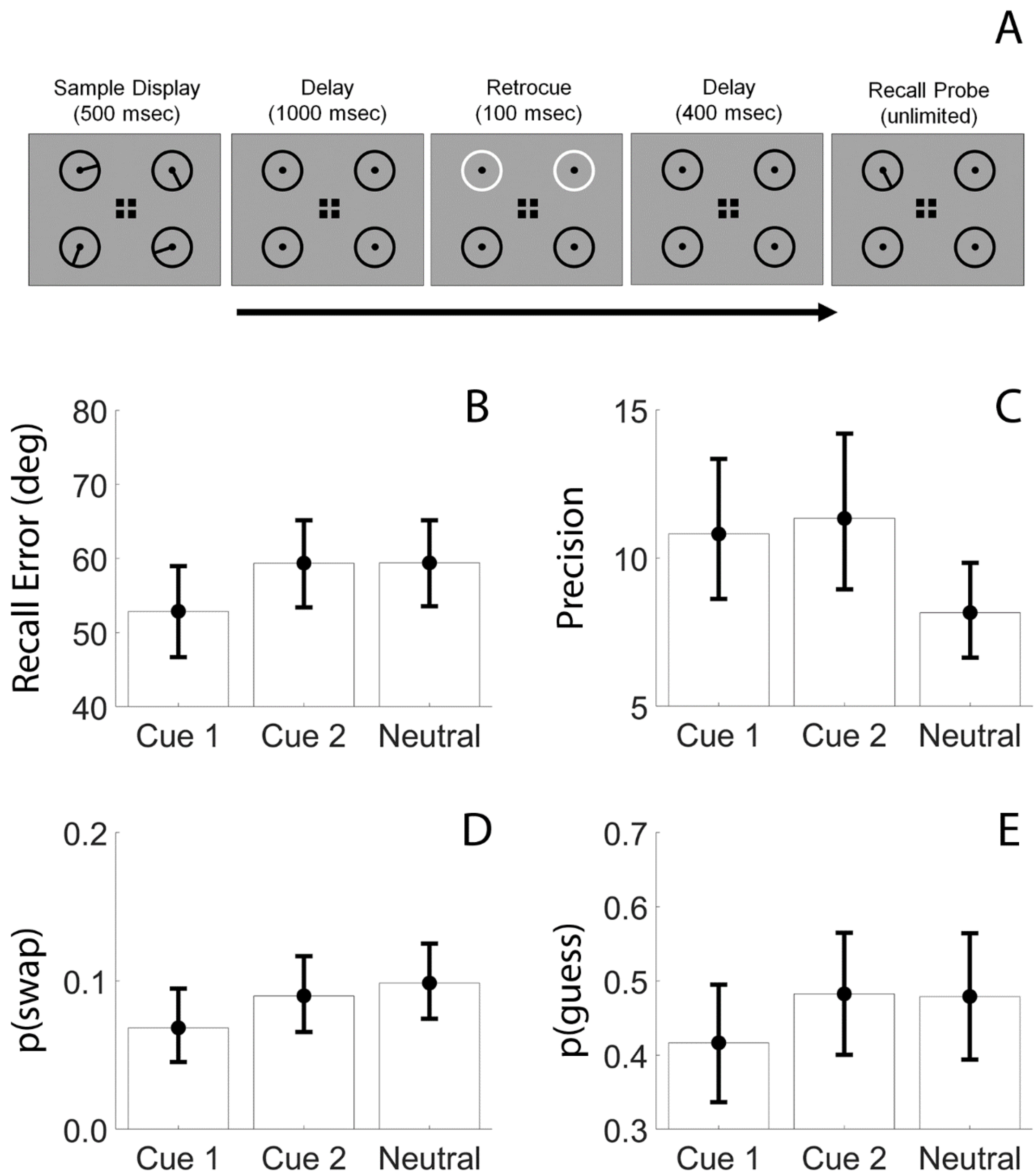

Figure 1. Design and Results of Experiment 1. (A) Task schematic showing a cue-two trial. Displays have been enlarged for exposition; see Methods for exact parameters. (B-E) Average absolute recall error (B), estimated recall precision (C), swap rates (D), and guess rates(E) as a function of cue condition. Error bars depict the 95% confidence interval of the mean.

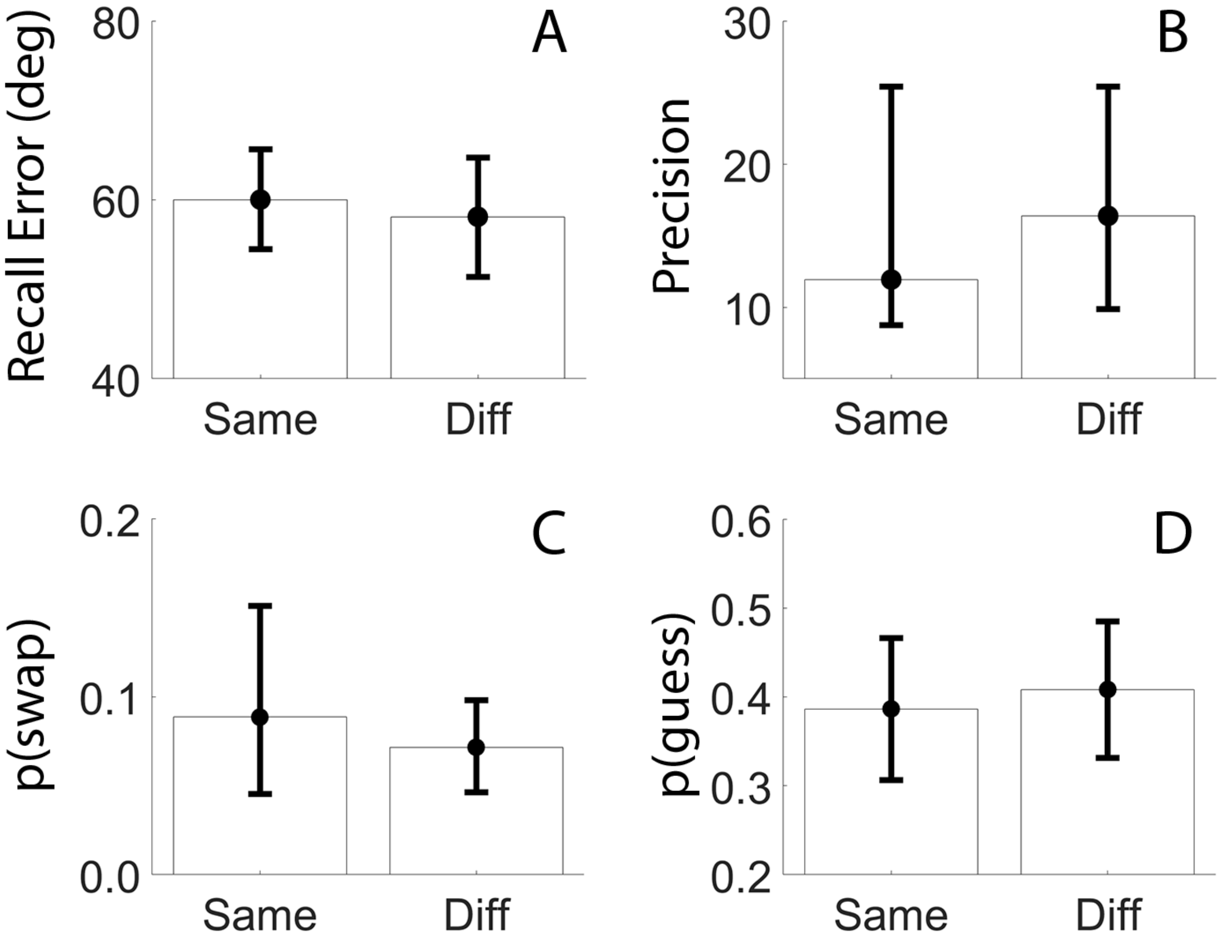

Figure 2. Hemifield Effects during Cue-two Trials in Experiment 1. We sorted participants average absolute recall errors (A), recall precision (B), swap rates (C), and guess rates (D) during cue-two trials according to whether the cued items appeared in the same visual hemifield or in different visual hemifields ("Diff"). Cue arrangement had no effect on any of these parameters. Error bars depict the 95% confidence interval of the mean.

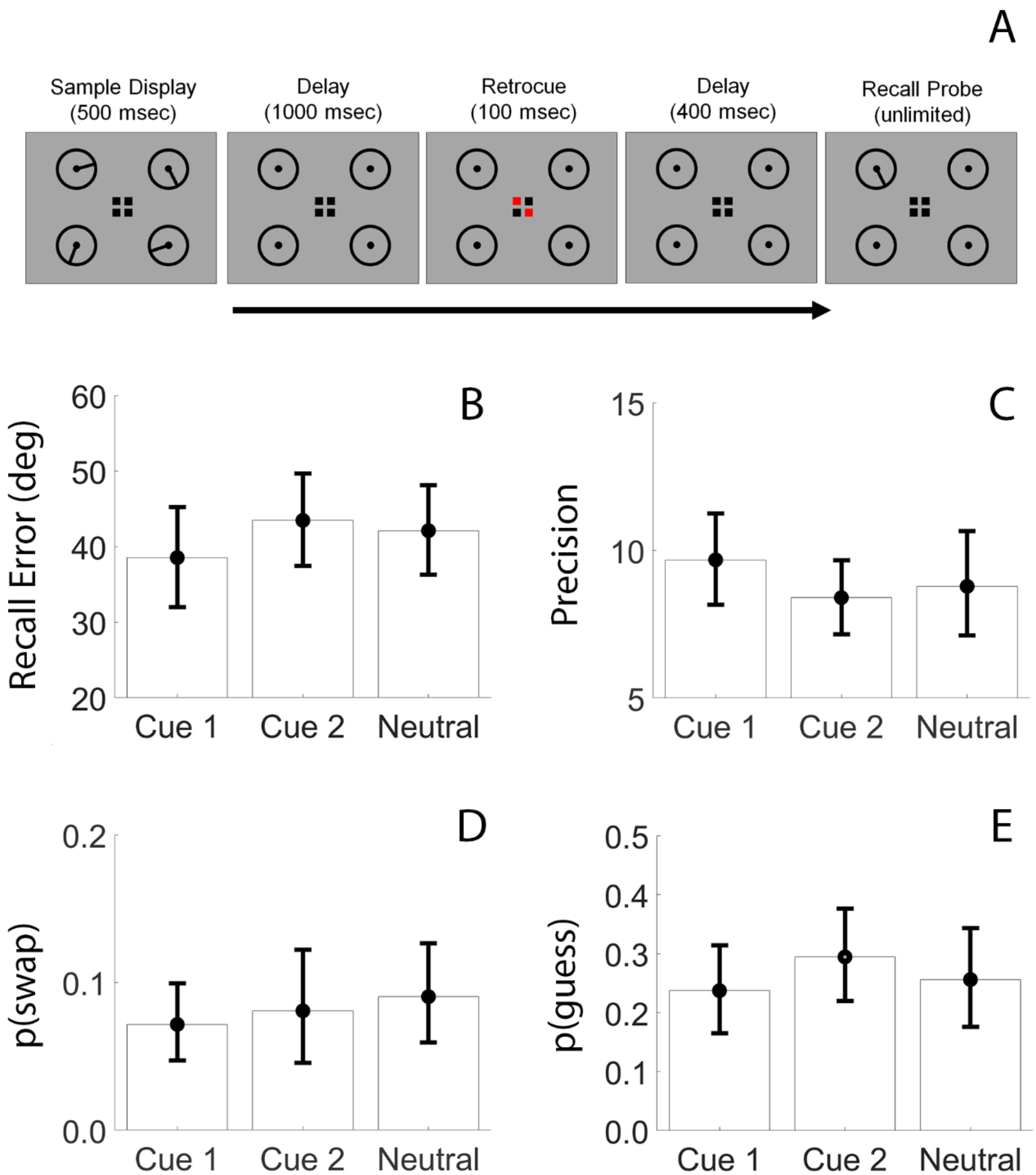

Figure 3. Design and Results of Experiment 2. (A) Task schematic showing a cue-two trial. Displays have been enlarged for exposition; see Methods for exact parameters. (B-E) Average absolute recall error (B), estimated recall precision (C), swap rates (D), and guess rates(E) as a function of cue condition. Error bars depict the 95% confidence interval of the mean.

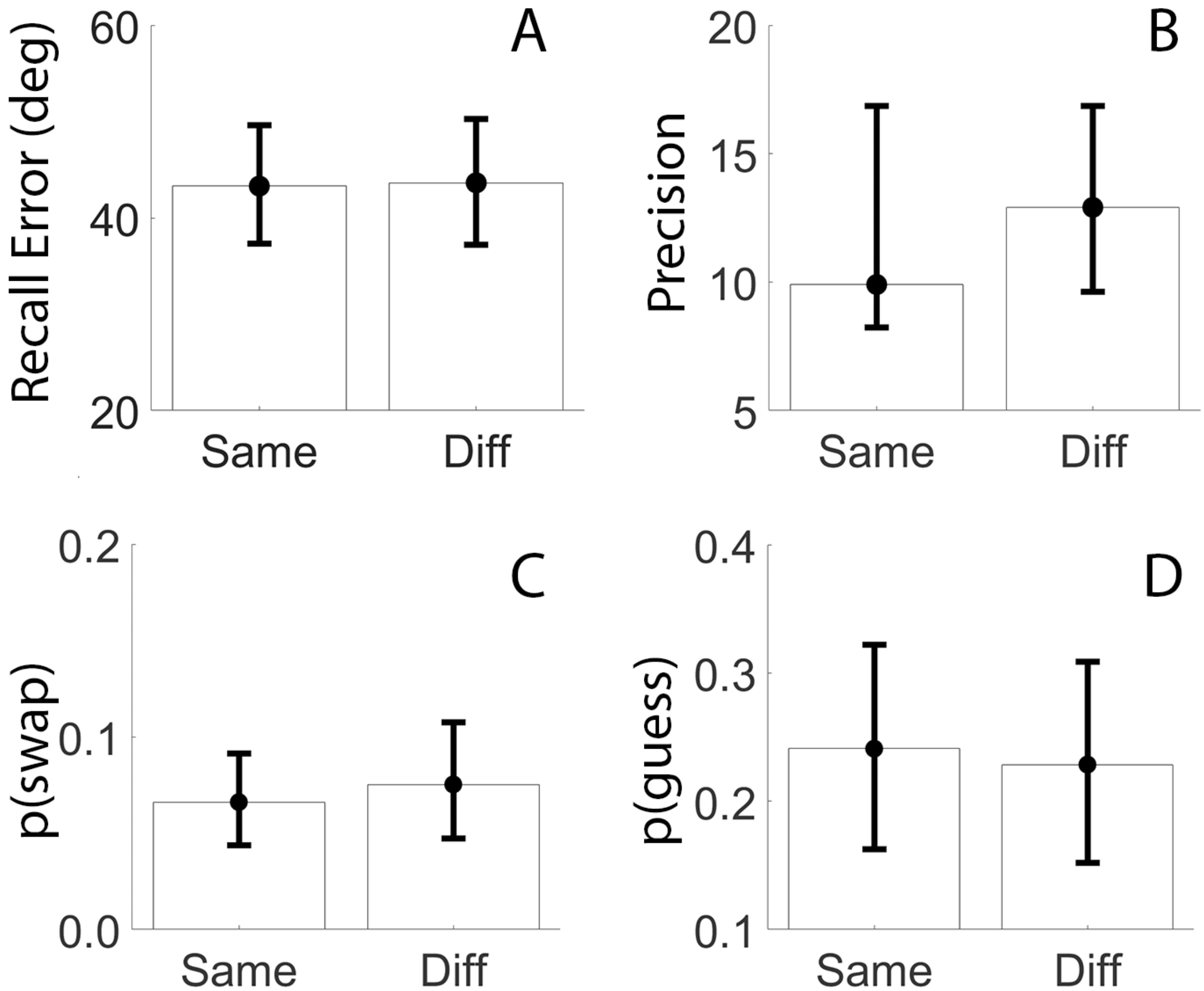

Figure 4. Hemifield Effects during Cue-two Trials in Experiment 2. We sorted participants average absolute recall errors (A), recall precision (B), swap rates (C), and guess rates (D) during cue-two trials according to whether the cued items appeared in the same visual hemifield or in different visual hemifields ("Diff"). Cue arrangement had no effect on any of these parameters. Error bars depict the 95% confidence interval of the mean.

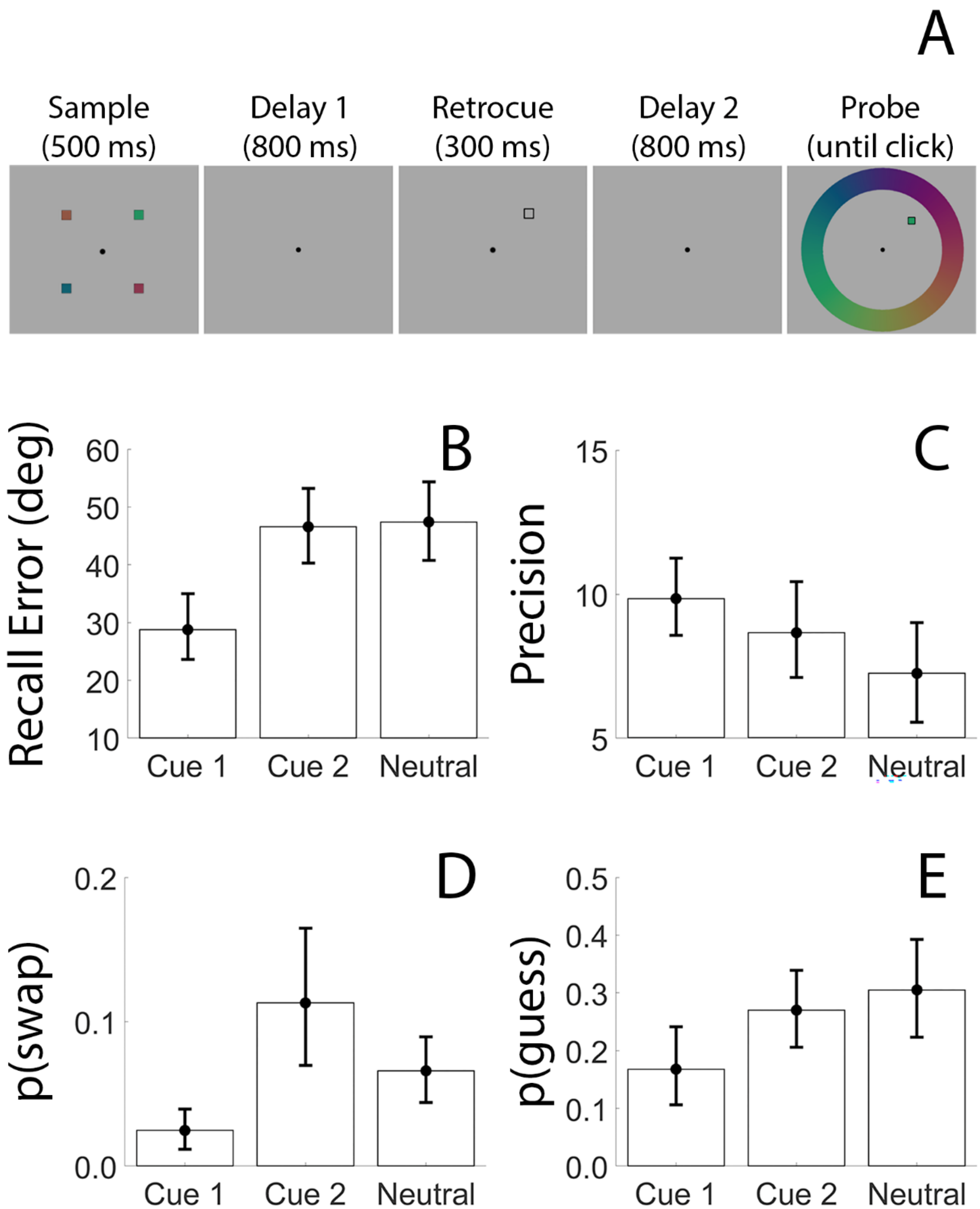

Figure 5. Design and Results of Experiment 3. (A) Task schematic showing a cue-one trial. Displays have been enlarged for exposition; see Methods for exact parameters. (B-E) Average absolute recall error (B), estimated recall precision (C), swap rates (D), and guess rates(E) as a function of cue condition. Error bars depict the 95% confidence interval of the mean.

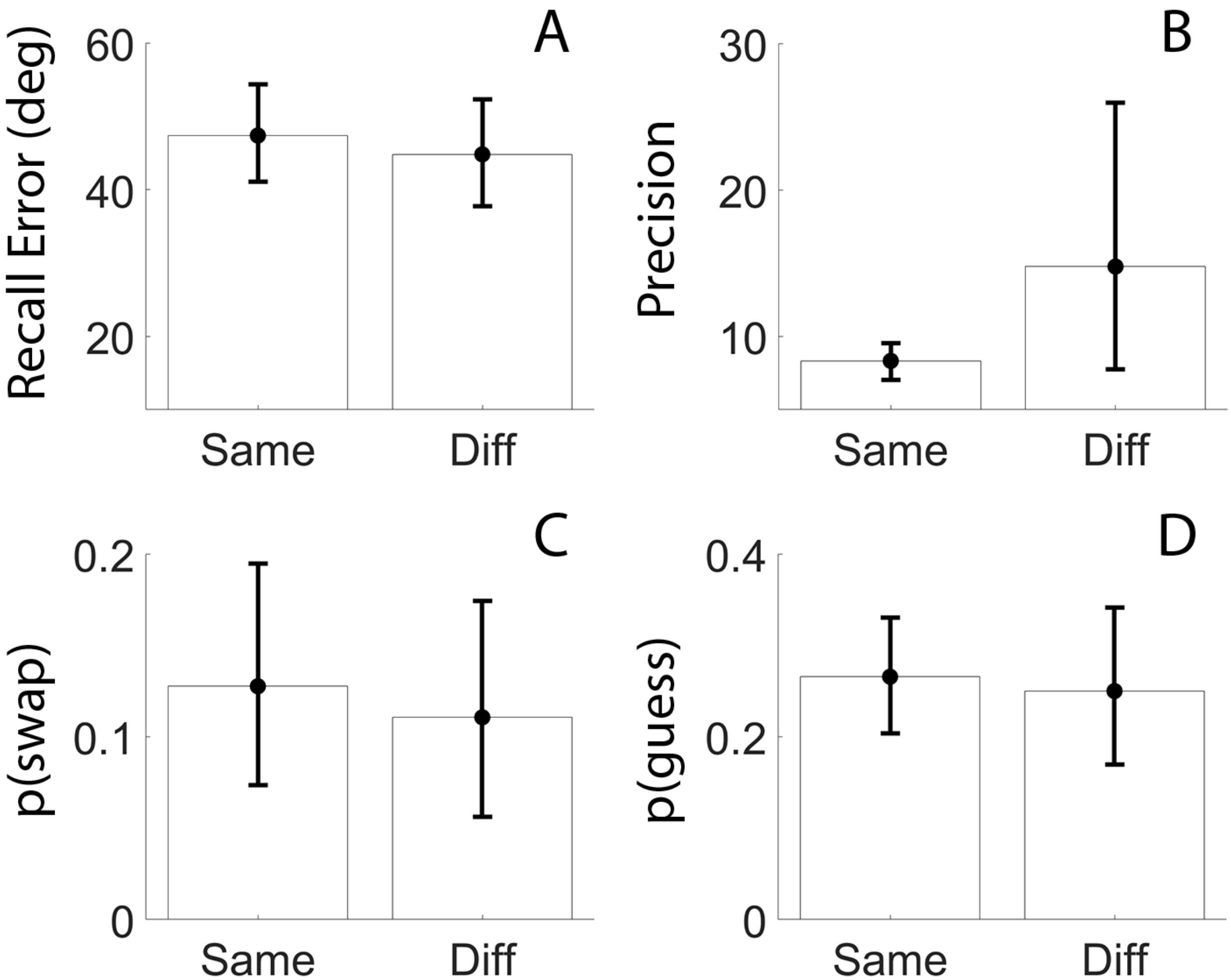

Figure 6. Hemifield Effects during Cue-two Trials in Experiment 3. We sorted participants average absolute recall errors (A), recall precision (B), swap rates (C), and guess rates (D) during cue-two trials according to whether the cued items appeared in the same visual hemifield or in different visual hemifields (“Diff”). Cue arrangement had no effect on any of these parameters. Error bars depict the 95% confidence interval of the mean.

INTRODUCTION

Working memory (WM) enables the temporary storage and manipulation of information no longer in the sensorium. This system has a limited capacity (e.g., Ma et al., 2014; Luck & Vogel, 2013), and mechanisms of selective attention are needed to control what information gains access to WM and to prioritize existing WM representations for behavioral output. It is well-established that WM performance can be facilitated by an informative cue presented during storage (e.g., Griffin & Nobre, 2003; Landman et al., 2003), and this effect - termed a retrocue benefit - can be used to explore the behavioral and neural consequences of prioritizing information stored in memory (e.g., Souza & Oberauer, 2016; Sprague et al., 2016; Myers et al., 2017; Ester et al., 2018; Nouri & Ester, 2020).

Several studies have demonstrated that human observers use retrocues to flexibly shift attention between different items stored in WM. For example, Landman et al. (2003) presented participants with multiple sequentially-presented retrocues and found a retrocue benefit for the last-cued item in the sequence (see also Li & Saiki, 2014; Maxcey et al., 2015). Using a similar procedure, Souza et al. (2016) found that retrocue benefits can extend to multiple sequentially-cued items, provided each cued item is equally likely to be probed at the end of a trial. Thus, participants can sequentially prioritize different items stored in WM for later report. Whether participants can *simultaneously* prioritize different items is less clear. Some studies have found that multiple simultaneous retrospective cues encouraging participants to prioritize a subset of items stored in WM confer no performance benefit above a no-cue or neutral cue condition (Makovski & Jiang, 2007; Oberauer & Bialkova, 2009). In one example, Makovski and Jiang (2007) showed participants displays containing six colored discs. After a blank delay, participants were required to report whether a single probed disc matched the color of the disc it replaced. During the delay period, participants were shown displays containing 0, 1, 2, 3, or 6 spatial cues. Participants were informed that during cue-present trials (i.e., 1, 2, 3, or 6 spatial cues) the probed item would always be drawn from the subset of cued colors. These cues informed the participant which memory item(s) were most likely to be probed at the end of the trial. Relative to cue-zero baseline, change detection performance was enhanced when participants were cued to one position, but not when they were cued to two, three, or six positions. However, other studies have reported WM performance benefits for multiple simultaneous retrospective cues under specific circumstances (e.g., Delvinne & Holt, 2012; Matsukura & Vecera, 2015; Heuer & Schubö, 2016). For example, Delvenne and Holt (2012) found a retrocue benefit in change detection performance when participants were cued to prioritize two items that appeared in different visual hemifields, but not in the same visual hemifield. Likewise, Heuer and Schubö (2016) presented participants with spatial and feature retrocues and found that while feature-based cues yielded benefits for multiple cued items presented at both contiguous and non-contiguous locations, spatial cues only yielded benefits when the cued items appeared at contiguous locations.

To our knowledge, all studies examining multiple *simultaneous* retrocue benefits on WM performance have relied on aggregate measures of memory performance such as average change detection accuracy or absolute recall error. These measures have the advantage of simplicity but make it difficult to determine how cues influence memory performance. For example, retrocue benefits in change detection performance could reflect (a) increased recall precision for cued relative to uncued memory items, (b) a decrease in the likelihood of forgetting the cued memory item, (c) a decrease in the likelihood of confusing the cued memory item with an uncued memory item (i.e., a swap error), or (d) some combination of the above. Indeed, several studies have reported that compared to a neutral- or no-cue condition, a single informative retrocue improves recall precision, lowers swap rates, and lowers the probability of random guessing (e.g., Murray et al., 2013; Pertzov et al., 2013; Souza et al., 2014; Gunseli et al., 2015; Makovki & Pertzov, 2015; Souza et al., 2016). What happens when participants are cued to prioritize multiple items? It is possible that directing attention to multiple stimuli held in memory confers benefits on some aspects of memory performance (e.g., reducing the likelihood of random guessing) while harming other aspects of memory performance (e.g., increasing the likelihood of swap errors). However, this pattern would be opaque to discretized or aggregate measures of memory performance. To illustrate, consider a hypothetical experiment where participants encode four items into memory and are subsequently probed to recall the identity of one item. During storage participants receive a cue indicating that the to-be-probed item will be drawn from a subset of two items stored in memory (i.e., informative cue trials) or a cue indicating that all four items are equally likely to be probed (i.e., uninformative cue trials). Suppose that the results of this experiment show that during uninformative cue trials participants correctly recall the identity of the probed item with probability 0.7, incorrectly recall the identity of a non-probed item (i.e., a task-irrelevant stimulus) with probability 0.1, and randomly guess with probability 0.2, but during informative cue trials participants recall the identity of the probed item with same probability of 0.7, but recall the identity of a non-probed item with greater probability 0.25, and randomly guess with lower probability 0.05. Clearly, these patterns would indicate that participants are processing stored information differently during informative and uninformative cue trials, yet they would yield (nearly) identical average absolute recall error estimates.

We investigated the effects of multiple simultaneous retrospective cues on WM performance in three experiments. Participants were retrospectively cued to prioritize zero, one, or two orientations (Experiments 1 or 2) or colors (Experiment 3) stored in WM for subsequent recall. Analyses of participants' average absolute recall errors revealed a significant performance benefit during cue-one relative to cue-two and cue-zero trials, replicating several earlier findings (e.g., Makovski & Jiang, 2007; Oberauer & Bialkova, 2009). Analyses of participants' recall precision, swap rates, and guess rates revealed that superior recall performance during cue-one relative to cue-two and cue-zero trials was driven primarily by lower guessing rates, consistent with other prior findings (e.g., Pertzov et al., 2013; Murray et al., 2013). However, precision estimates, swap rates, and guess rates were identical during cue-two and cue-zero trials. Thus, we found no evidence to support the hypothesis that multiple simultaneous retrocues influence memory performance compared to an uninformative cue display. This, in turn, lends further support to the hypothesis that retrocue-based access to the contents of WM is limited to one item at a time (Makvoski & Jiang, 2007; Oberauer & Bialkova, 2009; Souza & Oberauer, 2016).

METHODS

Participants. A total of 98 volunteers participated in this study. 45 volunteers from the Florida Atlantic University community participated in Experiment 1, 28 volunteers from the University of Nevada, Reno community participated in Experiment 2, and 25 volunteers from the Brock University community participated in Experiment 3. All participants were aged 18-40 and self-reported normal or corrected-to-normal visual acuity. All experimental procedures were approved by local institutional review boards, and all volunteers gave both written and oral informed consent before enrolling in the study. Data from 8 participants in Experiment 1 were discarded due to chance-level task performance (i.e., average absolute recall errors ≥ 85° in the cue-one condition). Thus, the data reported here reflect the remaining 90 participants (37 in Experiment 1, 28 in Experiment 2, and 25 in Experiment 3).

Stimulus Displays and Testing Environment. Participants in each experiment were seated in a dimly lit and quiet room for the duration of testing. Stimuli for Experiments 1 and 2 were generated in MATLAB and rendered using Psychtoolbox 3 software extensions. Stimuli in Experiment 1 were rendered on a 19-inch Dell CRT monitor cycling at 75 Hz; stimuli in Experiment 2 were rendered on a 27-inch LCD monitor cycling at 240 Hz. Participants were seated approximately 65 cm from the display (head position was unconstrained). Stimuli for Experiment 3 were generated in Python and rendered on a 20’’ LCD display using PsychoPy software (Peirce et al., 2019). Participants were seated approximately 57 cm from the display (head position was unconstrained).

Experiment 1 – Exogenously Cued Orientation Recall. A trial schematic is shown in Figure 1A. Participants were instructed to maintain fixation on a small dot (subtending 0.25° visual angle from a viewing distance of 60 cm) for the duration of each trial. Each trial began with a sample display containing four "clock-face" stimuli at 45°, 135°, 225°, and 315° polar angle along the perimeter of an imaginary circle (radius 5°) centered at the fixation point. Each stimulus subtended 2.5° (diameter) and contained a bar (1.25° length, 8-pixel stroke width) whose orientation was randomly and independently chosen from a uniform circular distribution on the interval (0°, 359°]. The sample display was presented for 500 ms and followed by a 1000 ms blank delay. A cue display was presented for 100 ms, followed by a 400 ms blank delay and a probe display containing a single clock-face stimulus. The probe stimulus was assigned a random orientation value, and participants were instructed to adjust it to match the sample stimulus it replaced using the left and right arrow keys. Participants entered their final response by pressing the spacebar. Participants were instructed to prioritize accuracy and no response deadline was imposed.

We retrospectively cued participants to prioritize zero, one, two, or all four stimuli. Cues were rendered by flashing the circular outline of the relevant stimuli white for 100 ms (see Figure 1A). During cue-one trials we randomly cued one of the four stimuli, subject to the constraint that each location was cued equally often within a single block of 60 trials. During cue-two trials we randomly cued two of the four stimuli. We did not explicitly control the spatial relationship between the cued stimuli, i.e., whether they appeared in the same vs. different hemifields, but we did investigate possible effects of cue location in post-hoc analyses (e.g., Figure 2). The cue-zero and cue-four conditions were included as neutral baselines. Both conditions yielded equivalent performance, so data from these trials were pooled to create a single neutral cue condition (specifically, we analyzed the cue-zero and cue-four trials separately for each observer and then averaged the data across conditions). When present, cues were 100% valid in the sense that the probe always appeared at a cued location. Each participant completed 7 (N = 1), 8 (N = 2), 10 (N = 36) or 11 (N = 6) blocks of 60 trials as time permitted (participants were given a maximum of 1.5 hours to complete the experiment). Performance feedback in the form of average absolute report error was given at the end of each block.

Experiment 2 – Endogenous Orientation Recall. Experiment 1 used sudden onset cues that are known to trigger reflexive shifts of attention (Jonides & Yantis, 1988), thus, Experiment 2 was conducted to examine whether the findings of Experiment 1 would generalize to a scenario where participants were encouraged to endogenously select cued items. Experiment 2 was identical to Experiment 1, with the exception that (a) we eliminated the cue-zero condition and (b) we replaced the peripheral, exogenous cues used in Experiment 1 with central, endogenous cues. Specifically, we replaced the central fixation point used in Experiment 1 with a four-spoke fixation grid, here each spoke pointed towards one of the four stimulus locations (see Figure 3A). Participants were retrospectively cued to remembered stimuli by changing individual spokes on the fixation grid from black to red for 100 ms (see Figure 3A, which depicts an example cue-two trial). During cue-one trials we randomly cued one of the four stimuli, subject to the constraint that each location was cued equally often within a single block of 60 trials. During cue-two trials we randomly cued two of the four stimuli. Again, we made no attempt to control the spatial relationship between the cued stimuli (e.g., same vs.

different hemifields), but examined whether this factor influenced performance in post-hoc analyses (Figure 4). Each participant completed 5 (N = 2), 7 (N = 2), or 8 (N = 24) blocks of 60 trials as time permitted (participants were given a maximum of 1.5 hours to complete the experiment). Performance feedback (average absolute report error relative to the probed item) was given at the end of each block.

Experiment 3 – Color Recall. To examine whether the findings of Experiments 1 and 2 generalized to a new feature space, we conducted a third experiment where participants were retrospectively cued to 1, 2, or all four of the remembered colors (Figure 5A). Cues were rendered by displaying an outline of the relevant stimuli location for 300 ms (see Figure 5A). Stimulus colors were randomly selected from a 360° isoluminant CIE L*a*b color space with a minimum spacing of 30°. A sample display containing four colored squares (subtending 1° at a radial distance of 6° from fixation from a viewing distance of 57 cm) was presented for 500 ms followed by an 800 ms blank delay. A cue display was presented for 300 ms, followed by another 800 ms blank delay. Finally, a probe display containing one outline square was presented along with a color wheel; participants indicated their memory for the color that appeared at the outline location by clicking on the color wheel. Participants were instructed to prioritize accuracy, and no response deadline was imposed. During cue-one trials we randomly cued one of the four stimuli, there was no formal constraint on the number of times a location could be cued in a given block of trials. During cue-two trials we randomly cued two of the four stimuli. We made no attempt to control the spatial relationship between the cued stimuli (e.g., same vs. different hemifields). Participants completed 100 trials in the cue-four, cue-one, and cue-two conditions, breaks were given every 25 trials. Performance feedback was not given.

20 of 25 of the participants additionally completed 200 trials of an unreliable cue-one condition in which the cue was informative but not always valid (the cued item was probed 50% of the time); the results are not analyzed here.

Data Analysis and Statistics. Data from each experiment were analyzed using two complementary methods. To get an overall view of participants' task performance, we computed estimates of average absolute recall error (i.e., the average absolute angular difference between the orientation or color reported by the participant and the actual probed orientation or color). We also fit participants' recall data using a parametric model of the form:

$$p(\hat{\theta}) = (1 - \gamma - \beta)\Phi_\sigma(\hat{\theta} - \theta) + \gamma\frac{1}{2\pi} + \beta\frac{1}{m}\sum_i^m \Phi_\sigma(\hat{\theta} - \theta_i^*)$$

where θ is the target feature value, $\hat{\theta}$ is the reported feature value, γ is the proportion of trials where the subject guesses, β is the probability of misremembering the target location, $\{\theta_1^*,\ \theta_2^*, \ldots \theta_m^*\}$ are the values of the *m* nontarget items, and Φ_σ is a von Mises distribution with mean 0 and standard deviation σ (Bays et al., 2009). The effects of cue number (i.e., cue-one, cue-two, etc.) on these parameters were estimated via one-way repeated-measures analysis of variance (ANOVA) with cue number as the sole model factor. Where appropriate, false-discovery-rate-corrected post-hoc comparisons were performed via repeated measures t-tests. Throughout the manuscript, we report condition averages, 95% confidence intervals, and effect sizes (η^2 and Cohen's *d*). Non-significant effects were probed with Bayesian pairwise t-tests with uniform priors to quantify evidence for the null hypothesis using custom MATLAB software (available for download at https://github.com/klabhub/bayesFactor). The result of a Bayesian t-test is a

Bayes Factor, typically denoted BF_{10}. For example, a Bayes Factor of 3.0 provides 3-to-1 odds favoring the alternative over the null hypothesis. Since Bayesian analyses were restricted to null effects (estimated using frequentist statistics), we computed an inverse Bayes Factor BF_{01} describing the strength of evidence favoring the null over the alternative hypothesis, i.e., $BF_{01} = \frac{1}{BF_{10}}$.

RESULTS

Experiment 1. The results of Experiment 1 are summarized in Figure 1B-E. A one-way ANOVA applied to participants' recall errors (Figure 1B) revealed a main effect of cue number (i.e., 0, 1, 2, or 4), $F(2, 72) = 22.81$, $p < 1e^5$, $\eta^2 = 0.388$. Post-hoc comparisons revealed that this effect was driven by superior performance during the cue-one relative to cue-two trials (M = 52.88° vs. 59.37°, respectively; $t(36) = 5.456$, $p < 1e^5$, $d = 0.34$; 95% CI of the difference = 4.25-8.90°) and during cue-one relative to neutral trials (M = 52.88° vs. 59.40°; $t(36) = 5.19$, $p < 1e^5$, $d = 0.35$; 95% CI of the difference = 4.25-9.08°). Recall performance during cue-two and neutral trials was statistically indistinguishable, $t(36) = 0.04$, $p = 0.963$; $BF_{01} = 5.65$ (for reference, a BF_{01} of 5.0 indicates 5-to-1 odds favoring the null hypothesis; see *Data Analysis and Statistics,* Methods). These findings are consistent with earlier studies failing to find a multiple retrocue benefit in WM (e.g., Makvoski & Jiang, 2007). Complementary analyses of participants' recall precision, swap rates, and guess rates revealed a more detailed picture. Cue number had no effect on estimates of recall precision (Figure 1C; $F(2, 72) = 2.64$, $p = 0.078$) nor swap rates (Figure 1D; $F(2, 72) = 2.90$, $p = 0.06$). However, cue number had a significant effect on guess rates (Figure 1E; $F(2, 72) = 3.151$, $p = 0.048$, $\eta^2 = 0.08$). Visual inspection of Figure 1E suggests that this effect was driven by lower guessing rates during cue-one relative to cue-two trials (M = 0.417 and 0.482, respectively; 95% CI of the difference = -0.023-0.136) and/or lower guessing rates during cue-one-relative to neutral trials (M = 0.417 vs. 0.479; 95% CI of the difference = 0.015-0.122). However, post-hoc comparisons between these conditions did not survive correction for multiple comparisons (t(36) = 1.86, 0.15, and 2.26; p = 0.107, 0.885, and 0.09; and BF_{01} = 1.20, 5.60, and 0.60 for the comparisons of cue-one vs. cue-two, cue two vs. neutral, and cue-one vs. neutral trials, respectively).

We also investigated whether the spatial positions of the cued stimuli during cue-two trials influenced memory performance. For example, Delvenne & Holt (2012) found multiple simultaneous retrocue benefits when cued stimuli were arranged in different visual hemifields, but not in the same visual hemifield. We tested this possibility by sorting cue-two trials according to the spatial arrangement of retrospectively cued items (i.e., same vs. different hemifields) and recomputing average absolute recall error, recall precision, swap rates, and guess rates within each group. Cue arrangement had no impact on any of the parameters we examined, with equivalent performance during same- and different-hemifield trials for average absolute recall error (Figure 2A; t(36) = 1.272, p = 0.212), recall precision (Figure 2B; t(36) = 1.098, p = 0.279), swap rates (Figure 2C; t(36) = 0.555, p = 0.582) or guess rates (Figure 2D; t(36) = 0.671, p = 0.507). These conclusions were supported by Bayesian t-tests (BF_{01} = 2.69, 3.24, 4.89 and 4.58 for recall errors, recall precision, swap rates, and guess rates, respectively.

Experiment 2. The results of Experiment 1 demonstrated that participants' absolute recall performance was worse during cue-two and neutral trials compared to cue-one trials. Furthermore, precision estimates, swap rates, and guess rates were statistically indistinguishable across cue-two and neutral trials. Next, we sought to replicate and extend these findings by examining whether a similar pattern was observed when participants were endogenously (as opposed to exogenously) simultaneously cued to prioritize multiple items.

The results of Experiment 2 are summarized in Figure 3B-E. A one-way ANOVA applied to participants' recall errors (Figure 3B) revealed a main effect of cue number, $F(2, 54) = 13.81$, $p < p < 1e^5$, $\eta^2 = 0.388$. Post-hoc comparisons revealed that this effect was driven by superior performance during the cue-one relative to cue-two trials (M = 38.52° and 43.45°, respectively; $t(27) = 4.89$, $p < 1e^4$, $d = 0.279$; 95% CI of the difference = 2.93-6.82°) and superior performance during cue-one relative to neutral trials (M = 38.52° and 42.09°, respectively; $t(27) = 3.49$, $p = 0.002$, $d = 0.19$; 95% CI of the difference = 1.54-5.50°), replicating the findings of Experiment 1 and prior research (e.g., Makovski & Jiang, 2007; Oberauer & Bialkova, 2009). Cue number had no effect on recall precision (Figure 3C; $F(2, 54) = 1.205$, $p = 0.308$) or swap rates (Figure 3D; $F(2, 54) = 0.595$, $p = 0.553$), but did have a significant effect on guess rates (Figure 3E; $F(2, 54) = 3.580$, $p = 0.035$, $\eta^2 = 0.12$). Visual inspection of Figure 3E suggests that this effect was driven by lower guessing rates during cue-one vs. cue-two and cue-one vs. neutral trials (M = 0.237, 0.294, and 0.256, for cue-one, cue-two, and neutral trials, respectively). Indeed, false-discovery-rate-corrected post-hoc comparisons revealed significantly higher guess rates during cue-one vs. cue-two trials [$t(27) = 2.872$, $p = 0.0235$, $d = 0.276$; 95% CI of the difference = 0.018-0.095]; but no difference between cue-one and neutral trials ($t(27) = -0.80$, $p = 0.431$; $BF_{01} = 3.72$) or between cue-two and neutral trials ($t(27) = 1.76$, $p = 0.135$; $BF_{01} = 1.28$). The spatial positions of cued stimuli during cue-two trials (i.e., same vs. different hemifield) had no impact on participants' recall errors (Figure 4A; $t(27) = 0.232$, $p = 0.812$), recall precision (Figure 4B; $t(27) = 1.511$, $p = 0.142$), swap rates (Figure 4C; t(27) = 0.550, p = 0.587), or guess rates (Figure 4D; t(27) = 0.377, p = 0.709). These conclusions were supported by Bayesian t-tests (BF_{01} = 4.86, 1.80, 4.34 and 4.67 for recall errors, recall precision, swap rates, and guess rates, respectively). Thus, the results of Experiment 2 are consistent with those of Experiment 1: first, participants' absolute recall performance was worse during cue-two and neutral trials compared to cue-one trials, and this effect was driven by a reduction in guess rates during cue-one relative to cue-two trials (Figure 3E). Second, estimates of recall precision, swap rates, and guess rates were statistically indistinguishable during cue-two relative to neutral trials.

Experiment 3. The results of Experiment 3 are summarized in Figure 5B-E. A one-way ANOVA applied to participants' recall errors (Figure 5B) revealed a main effect of cue number [F(2, 48) = 40.91, p < p < $1e^5$, η^2 = 0.630]. Post-hoc comparisons revealed that this effect was driven by superior performance during cue-one relative to cue-two trials (M = 28.78° and 46.57°, respectively; t(24) = 8.07, p < $1e^5$, d = 1.107; 95% CI of the difference = 13.56-22.13°) and during cue-one relative to neutral trials (M = 28.78° and 47.41°, respectively; t(24) = 6.98, p < $1e^5$, d = 1.127; 95% CI of the difference = 13.73-24.17°). Cue number also had a significant effect on swap rates (Figure 5D; F(2, 48) = 8.921, p = 0.0005, η^2 = 0.271) and guess rates (Figure 5E; F(2, 48) = 5.273, p = 0.009, η^2 = 0.180). Post-hoc analyses revealed significantly greater swap rates during cue-two vs. cue-one trials [M = 0.113 vs. 0.025, respectively; t(24) = 3.99, p = 0.0016, d = 0.797; 95% CI of the difference = 0.047-0.133] and during neutral vs. cue-one trials [M = 0.066 vs. 0.025, respectively; t(24) = 3.026, p = 0.0087, d = 0.838; 95% CI of the difference =

0.014-0.067], but no difference in swap rates during cue-two and neutral trials [t(24) = 1.86, p = 0.075; BF_{01} = 1.07]. Complementary analyses revealed significantly greater guess rates during neutral relative to cue-one trials [M = 0.305 vs. 0.168, respectively; t(24) = 3.27, p = 0.009, d = 0.679; 95% CI of the difference = 0.061-0.221] but no difference in guess rates during neutral relative to cue-two trials [M = 0.305 vs. 0.270, respectively; t(24) = 0.847, p = 0.405; 95% CI of the difference = -0.044-0.114; BF_{01} = 3.43]. Differences in guess rates during cue-one and cue-two trials were not statistically significant, [t(24) = 2.12, p = 0.066; 95% CI of the difference = -0.002-0.188; BF_{01} = 0.704]. Finally, we found no effects of hemifield (i.e., same vs. different) on task performance during cue-two trials (Figure 6; t(24) = 1.262, 1.274, 0.392, and 0.276 for recall errors, recall precision, swap rates, and guess rates, respectively; all $p > 0.215$). These conclusions were supported by Bayesian t-tests (BF_{01} = 2.33, 2.30, 4.42 and 4.58 for recall errors, recall precision, swap rates, and guess rates, respectively). Thus, the results of Experiment 3 are consistent with those of Experiments 1 and 2, with the exception that swap errors were more common during cue-two and cue-four relative to cue-one trials (Fig 5D); However, once again there were no observed differences in recall precision or error types (i.e., swaps vs. guesses) during cue-two relative to neutral trials.

DISCUSSION

Retrospective cue paradigms can be used to study the consequences of allocating attention to items stored in working memory (Griffin & Nobre, 2003; Landman et al., 2003). Retrospectively cueing a single item stored in memory leads to improvements in absolute recall error, recall precision, the probability of reporting a non-target item from memory (i.e., a swap error) and the probability of randomly guessing (Murray et al., 2013; Pertzov et al., 2013; Souza et al., 2014; Gunseli et al., 2015; Makovki & Pertzov, 2015; Souza et al., 2016). Multiple sequentially presented retrospective cues also improve WM performance, either for the last cued item in a sequence (Landman et al., 2003; Li & Saiki, 2014; Maxcey et al., 2015) or for all cued item in a sequence (Souza et al., 2016), depending on the specific task structure. Considerably less is known about whether human observers can use multiple *simultaneously* presented retrocues to prioritize two or more of several items stored in memory, with some findings arguing against this possibility (e.g., Makovski & Jiang, 2007; Oberauer & Bialkova, 2009) and other suggesting that it is possible only under certain circumstances, such as when simultaneously cued items appear in different visual hemifields (Delvenne & Holt, 2012), or when the cued items appear in neighboring spatial positions (Matsukura & Vecera, 2015; Heuer & Schubö, 2016). It is possible that multiple simultaneously presented retrospective cues improve some aspects of memory performance (e.g., reducing the probability of random guessing) while harming other aspects of memory performance (e.g., increasing the probability of reporting a non-probed item, i.e., a swap error); however, earlier studies examining the effects of multiple simultaneously presented retrospective cues on WM performance have relied on general measures such as change detection accuracy or absolute average recall error that are opaque to this possibility. Thus, we used a parametric modeling approach (Bays et al., 2009) to examine the effects of multiple simultaneously presented retrocues on different components of memory performance, namely, recall precision, non-target reports (i.e., swap errors), and random guessing. In three experiments, we retrospectively cued participants to prioritize one, two, or four of four items stored in memory, then probed a single item for recall. Average absolute recall errors were significantly lower during cue-one relative to cue-two or cue-four trials, and this effect was driven primarily by a reduction in random guesses (Figures 1E, 3E, 5E) and a reduction in swap errors during Experiment 3 (Figure 5D). Critically, the Bayesian testing found evidence supporting the conclusion that there were no differences in recall precision, non-target reports, or guessing during cue-two and cue-four trials, refuting the hypothesis that multiple simultaneously retrospective cues improve some aspects of memory performance while harming others.

One earlier study reported that multiple simultaneously presented retrospective cues improved WM performance when the cued items appeared in different visual hemifields, but not in the same visual hemifield (Delvenne & Holt, 2016). We were unable to replicate this finding in any of the experiments reported here (Figures 2, 4, and 6). However, we cautioned that our experiments were neither designed nor optimized to capture these effects. Locations of the retrospectively cued items were randomly selected during each cue-two trials. Since there are four possible different-hemifield cue combinations (i.e., upper and lower visual fields as in Figure 1A, or across the diagonals as in Figure 3A) and only two possible same-hemifield cue combinations (i.e., the two left or two right stimuli), the latter were underrepresented in our analysis. This, in turn, may have led to biased or inaccurate estimates of precision, swap rates, and guess rates in the same-hemifield condition.

The lack of a performance difference between cue-two relative to cue-four trials could reflect participants' inability to use multiple simultaneous retrospective cues. We think this unlikely for several reasons. First, there is ample evidence showing that human observers can successfully use multiple simultaneously presented cues to allocate attention in the external environment (e.g., Awh & Pashler, 2000; Müller et al., 2003; Franconeri et al., 2007; Ester et al., 2012; Ester et al., 2014) and to gate access to WM (e.g., Makovski & Jiang, 2007). Second, several studies have documented improved WM performance following the presentation of multiple sequentially presented cues (e.g., Li & Saiki, 2014; Maxcey et al., 2015; Souza et al., 2016). Thus, the limiting factor that precludes performance benefits following multiple simultaneously presented retrospective cues must involve selecting and prioritizing the appropriate items already stored in memory, rather than processing or interpreting the retrospective cue.

To summarize, we examined whether and how multiple simultaneous retrospective cues influenced different aspects of WM performance, including recall precision, the probability of reporting a non-probed item, and randomly guessing. We found superior WM performance during cue-one relative to cue-two or cue-four trials, replicating several prior findings (e.g., Griffin & Nobre, 2003; Landman et al., 2003; Makovski & Jiang, 2007). Conversely, multiple simultaneously presented retrospective cues had no influence on WM performance compared to a neutral cue condition, belying the hypothesis that these cues improve some aspects of memory performance while harming others. Thus, our findings suggest that – barring special circumstances not investigated here – cue-driven access to information stored in WM is limited to a single item at a time.

REFERENCES

Awh, E., Pashler, H. (2000) Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception & Performance*, *26*, 834-846

Bays, P.M., Catalao, R.F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of vision*, *9*(10), 7-7.

Delvenne, J.F., & Holt, J.L. (2012). Splitting attention across the two visual fields in visual short-term memory. *Cognition*, *122*, 258-263.

Ester, E.F., Drew, T., Klee, D., Vogel, E.K., & Awh, E. (2012) Neural measures reveal a fixed item limit in subitizing. *Journal of Neuroscience*, *32*, 7169-7177.

Ester, E.F., Fukuda, K., May, L.M., Vogel, E.K., & Awh, E. (2014) Evidence for a fixed capacity limit in attention multiple locations. *Cognitive, Affective, & Behavioral Neuroscience*, *14*, 62-77.

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

Franconeri, S.L., Alvarez, G.A., & Enns, J.T. (2007) How many locations can be selected at once? *Journal of Experimental Psychology: Human Perception & Performance*, *32*, 7169-7177.

Griffin, I.C., & Nobre, A.C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*, 1176-1194.

Gunseli, E., van Moorselaar, D., Meeter, M., & Olivers, C.N. (2015). The reliability of retro-cues determines the fate of noncued visual working memory representations. *Psychonomic Bulletin & Review*, *22*, 1334-1341.

Heuer, A., & Schubö, A. (2016). Feature-based and spatial attentional selection in visual working memory. *Memory & Cognition*, *44*, 621-632.

Jonides, J., Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*, 346-354.

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

Li, Q., Saiki, J. (2014) The effects of sequential attention shifts within visual working memory. *Frontiers in Psychology, 5,* 965.

Luck, S.J., & Vogel, E.K. 2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17,* 391-400.

Ma, W. J., Husain, M., & Bays, P.M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*, 347-356.

Makovski, T., & Jiang, Y.V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review*, *14*, 1072-1078.

Makovski, T., Pertzov, Y. (2015). Attention and memory protection: Interactions between retrospective attention cueing and interference. *Quarterly Journal of Experimental Psychology*, *68*, 1735-1743.

Matsukura, M., & Vecera, S.P. (2015) Selection of multiple cued items is possible during visual short-term memory maintenance. *Attention, Perception, & Psychophysics*, *77*, 1625-1646.

Maxcey, A.M., Fukuda, K., Song, W.S., & Woodman, G.F. (2015). Using electrophysiology to demonstrate that cueing affects long-term memory over the short term. *Psychonomic Bulletin & Review*, *22*, 1349-1357.

Müller, M.M., Malinowski, P., Gruber, T., & Hillyard, S.A. (2003) Sustained division of the attentional spotlight. *Nature*, *424*, 309-312.

Murray, A.M., Nobre, A.C., Clark, I.A., Cravo, A.M., Stokes, M.G. (2013) Attention restores discrete items to visual short-term memory. *Psychological Science, 24,* 550-556.

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

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

Oberauer, K., & Bialkova, S. (2009). Accessing information in working memory: Can the focus of attention grasp two elements at the same time? *Journal of Experimental Psychology: General*, *138*(1), 64.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.* DOI:10.3758/s13428-018-01193-y

Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1224.

Souza, A.S., Rerko, L., Lin, H.-Y., & Oberauer, K. (2014) Focused attention improves working memory: Implications for flexible-resource and discrete-capacity models. *Attention, Perception, & Psychophysics*, *76*, 2080-2102.

Souza, A.S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, *1339*, 20-31.

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

Souza, A.S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, *78*(7), 1839-1860.

Sprague, T.C., Ester, E.F., Serences, J.T. (2016). Restoring latent visual working memory representations in human cortex. *Neuron*, *91*, 694-707.