Journal of Experimental Psychology:
Human Perception and Performance
2018, Vol. 44, No. 7, 992–1011

0096-1523/18/$12.00 http://dx.doi.org/10.1037/xhp0000509

The Architecture of Interaction Between Visual Working Memory and Visual Attention

Brett Bahle, Valerie M. Beck, and Andrew Hollingworth
The University of Iowa

In five experiments, we examined whether a task-irrelevant item in visual working memory (VWM) interacts with perceptual selection when VWM must also be used to maintain a template representation of a search target. This question is critical to distinguishing between competing theories specifying the architecture of interaction between VWM and attention. The single-item template hypothesis (SIT) posits that only a single item in VWM can be maintained in a state that interacts with attention. Thus, the secondary item should be inert with respect to attentional guidance. The multiple-item template hypothesis (MIT) posits that multiple items can be maintained in a state that interacts with attention; thus, both the target representation and the secondary item should be capable of guiding selection. This question has been addressed previously in attention capture studies, but the results have been ambiguous. Here, we modified these earlier paradigms to optimize sensitivity to capture. Capture by a distractor matching the secondary item in VWM was observed consistently across multiple types of search task (abstract arrays and natural scenes), multiple dependent measures (search reaction time (RT) and oculomotor capture), multiple memory dimensions (color and shape), and multiple search stimulus dimensions (color, shape, common objects), providing strong support for the MIT.

Public Significance Statement

Many real-world tasks require that participants search or monitor for multiple targets. For example, a baggage screener might need to search for multiple different types of weapons. In the present study, we examined whether participants can maintain multiple representations in visual working memory that guide attention simultaneously to different objects. The results have implications both for theoretical accounts of visual search and for the practical application of these theories to real-world contexts.

Keywords: attentional guidance, visual attention, visual working memory

Real-world behavior requires that people efficiently direct attention and gaze to goal-relevant objects. There are multiple mechanisms by which goal-directed control of attention can be implemented, including template guidance by knowledge of the perceptual features of the target object (Wolfe, Cave, & Franzel, 1989), contextual guidance by knowledge of the structure of the environment in which the object appears (Torralba, Oliva, Castelhano, & Henderson, 2006), and guidance by long-term learning based on selection history and reward (Awh, Belopolsky, & Theeuwes, 2012). Of these mechanisms, guidance by a target template is central to most theories of visual search, and the effects of other forms guidance tend to be relatively small compared with the guidance provided by a representation of the target's visual features.[1]

Most theories of attention (e.g., Bundesen, 1990; Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Hamker, 2004) implement template guidance by means of a target representation in visual working memory (VWM). Template guidance can also be supported by long-term-memory (LTM) if the target attribute remains constant across trials (Carlisle, Arita, Pardo, & Woodman, 2011), but because search targets frequently change in real-world behavior, we focus here on the guidance of attention from a VWM representation. Indeed, cuing participants on a trial-by-trial basis to a relevant target feature leads to highly selective search (Vickery, King, & Jiang, 2005; Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004), with attention and gaze limited almost exclusively to task-relevant items in the display (Beck, Hollingworth, & Luck, 2012). Moreover, the maintenance of features in VWM interacts with

Brett Bahle, Valerie M. Beck, and Andrew Hollingworth, Department of Psychological and Brain Sciences, The University of Iowa.

This study was supported by a grant from NIH (R01EY017356) and by an NSF Graduate Research Fellowship to Valerie M. Beck. We thank Toby Mordkoff for helpful discussions of the present research.

Correspondence concerning this article should be addressed to Brett Bahle, Department of Psychological and Brain Sciences, The University of Iowa, W311 Seashore Hall, Iowa City, IA 52242. E-mail: brett-bahle@uiowa.edu

---

[1] The exception may be search for common objects in natural scenes, which is sometimes influenced more by scene contextual factors than by a target template (Torralba et al., 2006; Wolfe, Võ, Evans, & Greene, 2011). However, even in this domain, recent evidence suggests that template-based guidance can dominate the early stages of visual search (Bahle, Matsukura, & Hollingworth, 2017).

perceptual selection even when those features are not associated with the target, further indicating a close relationship between VWM maintenance and the guidance of attention (Hollingworth & Luck, 2009; Hollingworth, Matsukura, & Luck, 2013b; Olivers, 2009; Olivers, Meijer, & Theeuwes, 2006; Soto, Heinke, Humphreys, & Blanco, 2005; Soto, Hodsoll, Rotshtein, & Humphreys, 2008; Soto & Humphreys, 2007).

Although the relationship between VWM and online template guidance is well established, the architecture of interaction between VWM and attention is currently under debate. Competing theories differ in their claims about the number of items in VWM that can interact with attention simultaneously. Under the single-item-template hypothesis (*SIT*; Olivers, Peters, Houtkamp, & Roelfsema, 2011), there are two qualitatively different states in VWM: an active *template* state and an *accessory* state. In this view, the template state allows a prefrontal VWM representation to feed back to sensory cortex, biasing sensory processing to increase the probability that matching items will be attended. However, in the accessory state, this type of interaction is blocked; accessory items are inert with respect to attentional guidance. Critically, the SIT holds that only one item in VWM can be maintained in the active template state. This claim is analogous to theories in the general working memory literature holding that only one item is maintained as the "focus of attention" and can control ongoing cognitive operations (McElree, 2006; Oberauer, 2002).

In contrast, according to the multiple-item-template hypothesis (*MIT*; Beck et al., 2012), multiple items in VWM can guide attention simultaneously. This proposal derives from evidence that the maintenance of features in VWM correlates with sustained activation of subpopulations of neurons in visual cortex, consistent with a *sensory recruitment* hypothesis of VWM maintenance (Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009). Critically, multiple items in VWM can be decoded from activity in sensory cortex simultaneously (Emrich, Riggall, LaRocque, & Postle, 2013). If multiple items are maintained via sustained activation in sensory cortex, then it is plausible that they would interact simultaneously with new sensory processing to bias attention toward matching items. This view is consistent with general theories of working memory holding that the "focus of attention" can span multiple items (Cowan, 1999, 2001).[2]

In the present study, we examined closely the central line of evidence supporting the SIT, which comes from studies that have examined guidance by a secondary item in VWM as participants search for targets that change on a trial-by-trial basis (Downing & Dodds, 2004; Houtkamp & Roelfsema, 2006; Olivers, 2009, Experiment 5). A frequently changing target requires that the search template is maintained online and updated. Thus, it should be represented in VWM and occupy the "template" slot, and, under the SIT, the secondary memory item should be relegated to an "accessory" role. In that latter role, the secondary item should not interact with perceptual selection, and when it appears as a distractor during search, it should not necessarily capture attention or otherwise impair performance.[3] Note that this design examines whether the mere maintenance of a secondary item in VWM automatically guides attention. It does not probe whether participants can *strategically* guide attention on the basis multiple feature values simultaneously (evidence of this type is reviewed in the General Discussion). Nevertheless, results from three different attention capture studies are central to the claims of SIT (Olivers et al., 2011).

In the first of these studies (Houtkamp & Roelfsema, 2006), participants conducted two sequential searches. They initially saw two line drawings of real-world objects, one that would be the target of the first search and one that would be the target of the second search. Participants searched for the first target in one array of objects (target preset or absent), and they then searched for the second target in a second array of objects (again, target present or absent). In this design, during the first search, the first target should occupy the template slot. The second target was not relevant during the first search, but it still had to be maintained in VWM in preparation for the second search. The key manipulation concerned the presence or absence of the second target object as a distractor in the first search array. A capture effect—longer mean RT on trials with a matching distractor—would indicate that the representation of the second target interacted with perceptual selection. No capture effect would be consistent with the maintenance of the second target in an inert, accessory state in VWM.

Although the results of these experiments have been interpreted as consistent with the SIT (Olivers et al., 2011), the full pattern of results in Houtkamp and Roelfsema (2006) was mixed and ultimately ambiguous. The key methodological details and experimental results are summarized in Table 1. Four of the six experiments reported either a reliable capture effect or a strong trend toward a capture effect. In three of these, the effect was limited to target-absent trials (i.e., the capture effect was reliably larger for target-absent than for target-present trials, and there was no reliable capture effect when analyzing the target-present trials separately). In one, there was a capture effect, and the magnitude of the capture effect did not reliably differ between target-present and target-absent trials. Houtkamp and Roelfsema treated only the target-present trials as diagnostic of capture. However, it is not clear why this should be so: if the second search target were maintained in an inert, accessory state, it should not have attracted attention, even when the first search target was absent, as it remained irrelevant to the search task.

In addition to a complex empirical pattern that provides at least some evidence consistent with guidance by the secondary item, there are several methodological features of the Houtkamp and Roelfsema (2006) study that limit its interpretation. First, most of the experiments allowed participants to encode the targets verbally or categorically and thus may not have been ideal for testing VWM maintenance of the second target (i.e., the second target could have been maintained as a verbal label or categorical code until required). Second, the two search targets were similar in that they

---

[2] Note that we do not claim that all active items in VWM are *equivalent* with respect to attentional guidance. There is likely to be task-based prioritization, particularly in the current experiments in which only one item was relevant to the search task.

[3] This contrasts with the SIT prediction for capture when that target value is *static* (Olivers et al., 2006; Soto et al., 2005). Under the SIT, when the target is static, search guidance comes to rely on a LTM template (Carlisle et al., 2011) rather than a VWM template. With the template slot in VWM empty, the secondary item is automatically promoted to that state (Olivers et al., 2011). Presumably, the template slot must always be filled, although the precise reason for this has not been discussed in detail. Because the secondary item occupies the vacated template slot, it interacts with perceptual selection, generating capture.

Table 1

Designs and Results of Experiments That Have Examined Capture From a Secondary VWM Item When the Search Target Changes on a Trial-by-Trial Basis

Study	Experiment	Search task	Search stimulus	Matching item a singleton on matching dimension	Dual task	Search dimension	Memory dimension	Memory sample/ search cue presentation	Within- or between-category secondary task	Articulatory suppression	*N*	Capture by secondary memory item
Houtkamp and Roelfsema (2006)												
	Exp. 1A	Pres/Abs	Circular array	No	Search/Search	Common objects	Common objects	Simultaneous	Between	No	10[a]	Trend[c]
	Exp. 1B	Pres/Abs	Circular array	No	Search/Search	Common objects	Common objects	Simultaneous	Between	Yes	10	Yes[c]
	Exp. 2A	Pres/Abs	Circular array	No	Search/Search	Color	Color	Simultaneous	Between	No	4[a]	No[d]
	Exp. 2B	Pres/Abs	Circular array	No	Search/Search	Color	Color	Simultaneous	Between	Yes	7	Yes
	Exp. 3A	Pres/Abs	Circular array	No	Search/Search	Fractal patterns	Fractal patterns	Simultaneous	No structure	Yes	4[b]	No
	Exp. 3B	Pres/Abs	Circular array	No	Search/Search	Fractal patterns	Fractal patterns	Simultaneous	No structure	No	4	Trend[c]
Downing and Dodds (2004)	Exp. 1	Pres/Abs	Circular array	No	Search/Memory	Novel shapes	Novel shapes	Simultaneous	No structure	Yes	8	No
	Exp. 2	Pres/Abs	Circular array	No	Search/Memory	Novel shapes	Novel shapes	Simultaneous	No structure	Yes	10	No
Olivers (2009)	Exp. 5	Discrim: notch location	Circular array	No	Search/Memory	Color	Color	Simultaneous	Within	No	14	No
Present Study	Exp. 1A	Discrim: line orientation	Circular array	Yes	Search/Memory	Shape	Color	Sequential	Within	Yes	40	Yes
	Exp. 1B	Discrim: line orientation	Circular array	Yes	Search/Memory	Shape	Color	Simultaneous	Within	Yes	40	Yes
	Exp. 2	Discrim: line orientation	Circular array	Yes	Search/Memory	Color	Shape	Sequential	Within	Yes	40	Yes
	Exp. 3	Discrim: letter orientation	Scene photo	No	Search/Memory	Common objects	Color	Sequential	Within	No	12	Yes[e]
	Exp. 4	Discrim: line orientation	Circular array	Yes	Search/Memory	Shape	Color	Sequential	Within	No	12	Yes[e]

Note. Search task indicates whether the main search task was target present/absent or discrimination of an incidental target feature (target always present). Search stimulus indicates the type of image used for search. The next column indicates whether, when a distractor matching the secondary value was present in the search array, that item was a singleton on the matching dimension. Dual task indicates the natures of the primary and secondary tasks. Search dimension is the dimension on which search was guided for the primary task. Memory dimension indicates the dimension of the secondary stimulus that was maintained in VWM during the primary search task. Memory sample/search cue presentation lists whether these two stimuli were presented simultaneously or sequentially. The next column indicates whether the secondary task required within-category or between-category memory. Articulatory suppression indicates whether verbal encoding of the stimuli was suppressed by a simultaneous articulation task. *N* lists the number of participants included in the final analyses. The final column indicates whether the study observed capture from the secondary memory item in VWM.

[a] Two were authors. [b] One was an author. [c] Effect limited to target-absent trials. [d] Distractors matching the secondary item more likely to be fixated. [e] Primary dependent measure was probability of distractor fixation.

came from the same dimension (i.e., both were object line drawings or both were colored disks, etc.). Thus, capture during the first search could have been attributable to a failure to keep track of which object was relevant for each of the subsequent search arrays. Evidence of capture therefore cannot provide unambiguous support for the alternative hypothesis that multiple items were maintained in a state that guided attention. Finally, the experiments often produced trend-level effects using a very small number of participants, raising concerns about whether they had sufficient power.

A second line of evidence supporting the SIT comes from a similar study by Downing and Dodds (2004). In this method (see Table 1), participants saw two novel silhouette shapes on each trial. They searched for one of these in an array of novel shapes (target present or absent) and remembered the second for a memory test at the end of the trial. The memory item could sometimes appear as a distractor in the search display. Downing and Dodds found no effect of the presence of a memory-matching distractor on search times. However, other evidence within their results suggests that the presence of a matching distractor interacted with search. Across the two experiments, accuracy was higher on target-present trials when the matching distractor was also present. Moreover, accuracy was lower on target-absent trials when the matching distractor was present. This pattern suggests that participants sometimes responded "present" based on distractor object presence rather than target presence. That is, participants did not always keep track of which object was relevant for each task. This complicates interpretation of the response time data. For example, mean RT on target-present trials may have been artificially reduced when a matching distractor was present, because participants sometimes responded "present" if either of the two objects was present in the array.

A final capture experiment supporting the SIT was reported in a study by Olivers (2009). In Olivers' Experiment 5, he implemented a *varied mapping* condition similar to the task of Downing and Dodds (2004). This method used color stimuli for both the search and memory tasks (see Table 1). The search task involved discrimination of a secondary target feature (notch location in a colored target disk) so that all trials were target present, alleviating some of the concerns about the present/absent method of Downing and Dodds (2004) and Houtkamp and Roelfsema (2006). Olivers likewise found no evidence of capture by the presence of an item matching the secondary memory color, despite finding capture when the target attribute remained static across trials (*consistent mapping* condition).

In contrast with these studies, evidence indicating simultaneous capture by multiple items in VWM comes from a recent experiment in which the number of items in VWM was manipulated along with the number of matching items in the display (Hollingworth & Beck, 2016). This study used a fixed search target across the experiment but addressed the same issue as in the studies reviewed above. Reliable memory-based capture was observed when two colors were maintained in VWM (cf., van Moorselaar, Theeuwes, & Olivers, 2014), and this effect scaled with the number of matching distractors in the display (one vs. two), suggesting that when two colors were held in VWM, both interacted with selection (see also Chen & Du, 2017).

In sum, the extant evidence provides no clear resolution to the question of whether a secondary item in VWM interacts with perceptual selection to capture attention during search. The Houtkamp and Roelfsema (2006) results provides some, albeit inconsistent, evidence supporting guidance by a secondary item, but this study has nevertheless been interpreted as consistent with the SIT. The Downing and Dodds (2004) experiments provided no evidence for secondary item capture. However, each of these studies is difficult to interpret given the nature of the present/absent search task and other methodological considerations. The Olivers (2009) experiment provides the strongest evidence to date in favor of the SIT, but this constitutes a single null effect. The Hollingworth and Beck (2016) study found capture by multiple VWM items in a search paradigm similar to that of Olivers, but this again was a single experiment, and it differed from the other experiments reviewed above in that the target value remained constant across trials. Resolving this question is central to resolving the theoretical debate between the SIT and MIT, with major implications for understanding the architecture of attentional guidance by VWM representations.

Present Study

In five experiments (1A, 1B, 2, 3, and 4), we examined whether attention is captured by a stimulus matching a secondary item maintained in VWM. The major design features were as follows (see Figure 1 for illustration and Table 1 for summary):

- In all experiments, the search target value changed on a trial-by-trial basis and thus should have occupied the "template" slot under the SIT.
- Each search task involved finding a particular target object and reporting a secondary feature of that object. Thus, all trials were target-present trials, removing any need to consider whether different strategies or mechanisms are involved in search when a target is present or absent.
- In all experiments (except Experiment 3), when a distractor matching the secondary memory item was present in the search array, the distractor was a singleton on the matching dimension. This method has produced robust memory-based capture in earlier experiments (Hollingworth & Beck, 2016; Hollingworth & Hwang, 2013; Hollingworth & Maxcey-Richard, 2013) but has not yet been applied to examine attentional capture when the search target value changes on a trial-by-trial basis.
- Each trial included two tasks: *search* and *memory*. Participants saw a search cue and a memory item. They searched for the cued target item, and then completed a two-alternative memory test.
- In all experiments, the relevant dimensions for the search and memory tasks were different (e.g., color memory and shape search). The memory dimension was irrelevant to the search task, minimizing the possibility that participants would confuse which of the two stimuli was required for each of the two tasks.
- The memory item and search cue stimuli were presented sequentially rather than simultaneously (except in Experiment 1B). Simultaneous presentation has been used in all previous experiments of this type (see Table 1). We used sequential presentation for two reasons. First, it created an additional, temporal dimension on which the memory item and search cue could be distinguished and segregated in

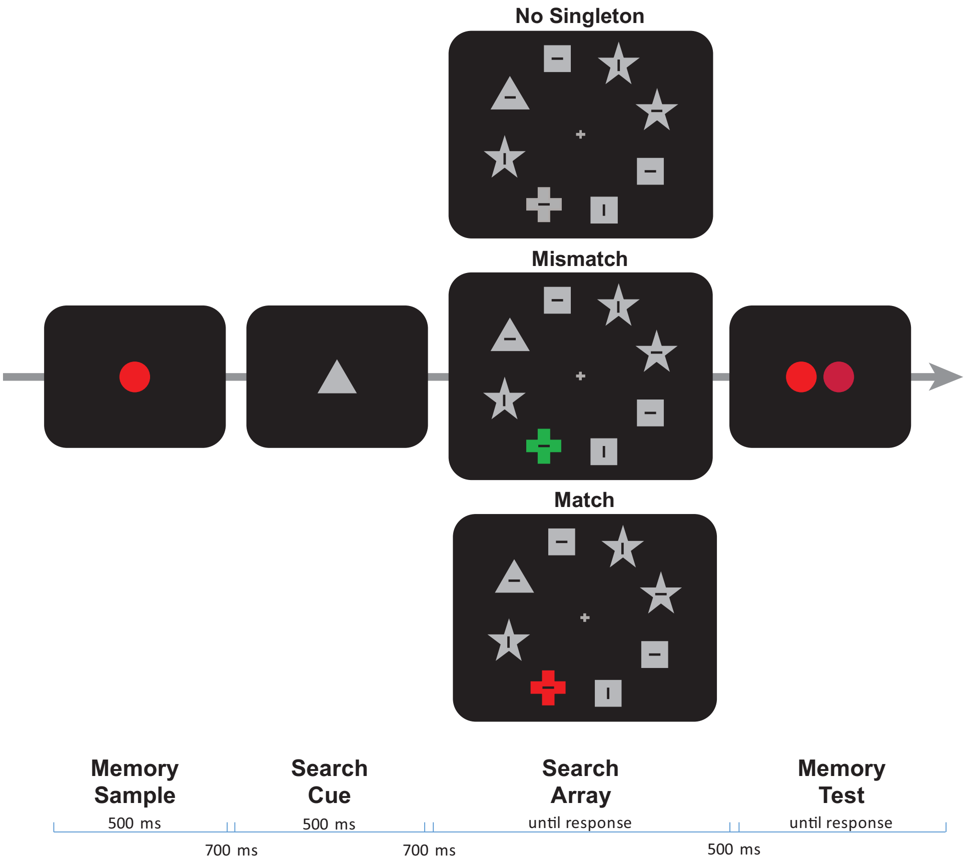

Figure 1. Sequence of events in a trial and singleton-match manipulation for Experiment 1A. See the online article for the color version of this figure.

memory. Second, the search cue always appeared after the memory item, facilitating the assignment of the search cue to the relevant "template" role, because it was presented more recently than the memory item and immediately before search. This is a conservative approach that should only work against the prediction of the MIT.

- The memory task required a relatively precise representation of the color or shape of the remembered object. That is, all memory tests included a foil from the same category as the remembered item, forcing participants to retain specific visual information in VWM and minimizing the possibility that stimuli were encoded verbally or categorically.
- As an additional safeguard against verbal encoding, all experiments included simultaneous articulatory suppression (except Experiments 3 and 4, in which eyetracking precluded this).
- Compared with previous experiments, the present experiments were designed to have substantially more power to detect an effect of secondary item memory match.

To preview the results, we found reliable capture by a secondary memory item in all four experiments, using three different search dimensions (color, shape, real objects), two different memory dimensions (shape and color), and two markedly different search tasks (search through arrays and search through natural scenes). In addition, we implemented the capture paradigm using two different dependent measures: manual RT and oculomotor capture. Although effects on manual RT were statistically reliable, effect sizes fell in the medium range (Experiments 1 and 2). In contrast, the oculomotor capture measure, which provides a direct index of the behavior of interest, produced large effects of secondary memory match (Experiments 3 and 4).

Experiments 1A and 1B

In Experiments 1A and 1B, the search task required participants to find a shape-defined target (randomly selected on each trial) and report the orientation of an embedded bar (see Figure 1). Simultaneously, they maintained a color in VWM in preparation for a within-category memory test at the end of the trial. There were three main conditions. In the *Match* condition, the search array contained one uniquely colored item (never the target) that matched the color category of the item maintained for the memory task. In the *Mismatch* condition, the search array contained one uniquely colored item (never the target) that *did not* match the color category of the item maintained for the memory task. This condition controlled for possible capture effects generated by the mere presence of a color singleton in the array. Finally, in the *No-Singleton* condition, the array did not contain a uniquely colored item. Longer search RT in the *Match* compared with the

Mismatch condition would indicate memory-based capture driven by the secondary memory item in VWM, consistent with the MIT. Equivalent search RTs in the *Match* and *Mismatch* conditions would indicate that the secondary memory item was inert with respect to attentional guidance, consistent with the SIT. In Experiment 1A, the memory sample and search cue were presented sequentially. In Experiment 1B, they were presented simultaneously, as in previous studies testing this question (see Table 1).

Method

Participants. Given the inconsistent extant data, and methods that were not always optimized to observe capture, it was not possible to estimate the required sample size from previous experiments. Instead, we used a relatively large sample size (40) to ensure sufficient power to detect a medium sized effect. For the type of within-subjects contrast of relevance here (distractor match vs. mismatch), a sample of 40 has 80% power to detect an effect of $\eta_p^2 = .18$ (calculated using G*power, Faul, Erdfelder, Lang, & Buchner, 2007). Each participant was recruited from the University of Iowa community, was between the ages of 18 and 30, reported normal or corrected to normal vision, and completed the experiment for course credit. In Experiment 1A, three participants were replaced, two who failed to perform significantly above chance on the memory task and one who failed to perform significantly above chance on the search task. Of the final 40 participants, 32 were female. In Experiment 1B, two participants who failed to perform significantly above chance on the memory task were replaced. Of the final 40, 26 were female. All human subjects procedures were approved by the University of Iowa Institutional Review Board. Each participant completed only one of the experiments in this series.

Apparatus. The apparatus was the same for Experiments 1 and 2. The stimuli were presented on an LCD monitor (resolution: 1280 × 960 pixels) with a refresh rate of 100 Hz at a viewing distance of 77 cm, maintained by a chin and forehead rest. Manual responses to both the search and memory test were collected with a response pad. The experiment was controlled by E-prime software (Schneider, Eschmann, & Zuccolotto, 2002).

Stimuli and procedure, Experiment 1A. After arriving for the experiment session, participants provided informed consent and were given oral and written instructions. They were tested individually in the presence of an experimenter. The sequence of events in a trial in Experiment 1A is illustrated in Figure 1. At the beginning of each trial, participants saw a screen with a "Press Button" message and four digits (not pictured in Figure 1). They began repeating the digits aloud at a rate of approximately two digits per second (monitored by the experimenter). After pressing a button to begin the trial, there was a 500-ms delay (fixation cross only), followed by a memory stimulus for 500 ms, a 700-ms ISI, a search cue for 500 ms, a 700-ms ISI, a search array until response, a 500-ms ISI, and a memory test stimulus until response. Participants responded to the search array to indicate whether a bar in the target shape was oriented horizontally (left button) or vertically (right button). They responded to the memory test array, using the same buttons, to indicate whether the exact match to the memory color was on the left or right. On trials with an incorrect memory test response, a message with a central, red "incorrect" was displayed for 300 ms. Feedback was not provided for the search task. There was an inter-trial-interval of approximately 500 ms before the appearance of the next "Press Button" screen.

All stimuli were presented against a black background with a central grayscale fixation cross subtending 0.61° × 0.61°. For the memory task, there were 12 possible colors, with three values in each of four color categories, reported in the Commission Internationale de l'Eclairage (CIE) 1931 color coordinate system: reds ($x = 0.63$, $y = 0.33$, 48.6 cd/m^2; $x = 0.61$, $y = 0.32$, 20.8 cd/m^2; $x = 0.61$, $y = 0.33$, 29.9 cd/m^2), blues ($x = 0.16$, $y = 0.09$, 52.7 cd/m^2; $x = 0.15$, $y = 0.05$, 28.2 cd/m^2; $x = 0.17$, $y = 0.11$, 25.8 cd/m^2), greens ($x = 0.31$, $y = 0.60$, 31.5 cd/m^2; $x = 0.27$, $y = 0.46$, 47.4 cd/m^2; $x = 0.31$, $y = 0.61$, 59.8 cd/m^2), and yellows ($x = 0.38$, $y = 0.49$, 176.6 cd/m^2; $x = 0.41$, $y = 0.47$, 163.0 cd/m^2; $x = 0.39$, $y = 0.49$, 116.1 cd/m^2). The memory item color was selected randomly from this set of 12 on each trial. The memory stimulus was a disk (a shape not used in the search arrays) with a diameter of 1.34° of visual angle, presented centrally. For the memory test, two colored disks (each subtending 1.34°) were presented to the left and right of central fixation at an eccentricity of 2.0° (measured to the center). The color of one of the test disks was identical to the initial memory disk (correct alternative). The foil color was drawn randomly from the remaining two colors in the same category. The locations of the matching and foil colors, and thus the correct response, were determined randomly on each trial.

For the search task, the search cue was a single, grayscale ($x = 0.29$, $y = 0.29$, 74.2 cd/m^2) shape selected randomly from four possible shapes that could appear in the search array: square, triangle, star, and cross. The shapes subtended, on average, 1.46° × 1.46°. The eight search items in the array were presented on a virtual circle around central fixation with a radius of 4.06°. The location of the first shape was selected randomly within a range from 1° to 45°. The remaining shapes were offset, each progressively by 45° around the virtual circle. Each shape contained a small, black bar (0.35° × 0.04°), oriented either vertically or horizontally (randomly selected for each item). One shape in the array matched the cue shape (location randomly selected). The other seven shapes in the array were chosen randomly from the remaining three nontarget shapes.

In the *No-singleton* condition, all of the shapes in the array were grayscale. In the remaining conditions, one of the distractor shapes was uniquely colored (randomly selected from the seven distractor items). In the *Match* condition, this singleton color matched the color category of the memory item color. *Match* trials were further divided into *Exact* and *Inexact* matches. On *Exact Match* trials, the singleton color was the same color as the memory item color. On *Inexact Match* trials, the singleton color was the color (from the same memory category) that would appear as the foil in the memory test. This discouraged participants from strategically attending to a matching array color in order to improve performance on the memory test (for similar methods, see Hollingworth & Beck, 2016; Hollingworth et al., 2013b; Olivers et al., 2006). In the *Mismatch* condition, the color singleton value was selected randomly from the values in the three color categories not used for the memory task on that trial. Participants were instructed that on some trials, one of the objects in the search array would have a unique color, but that this object would never be the search target.

Participants first completed a practice block of 12 trials. Then they completed two experimental blocks, with a short break between blocks. Each experimental block contained 160 trials: 50%

(80) in the *No-singleton* condition, 25% (40) in the *Mismatch* condition, and 25% (40) in the *Match* condition (evenly divided between *Exact Match* and *Inexact Match*). The *No-singleton* trials were included so that participants would not anticipate the presence of a color singleton on every trial. Trials from the various conditions were randomly intermixed. The entire experiment lasted approximately 45 min.

Stimuli and procedure, Experiment 1B. The stimuli and procedure were the same as in Experiment 1A, except that the memory sample and search cue were combined into a single display. The memory item was always presented above the central fixation cross, and the shape cue below the central fixation cross. Each was 1.54° from screen center. After initiating digit repetition and pressing a button to start the trial, there was a 500-ms delay, followed by the combined memory item and search cue display presented for 1500 ms, followed by a 1000-ms ISI before appearance of the search array.

Results

Search accuracy. Overall search task accuracy was 92.5% correct in Experiment 1A (see Table 2) and 94.4% correct in Experiment 1B. There was no effect of singleton-match condition (*No-singleton, Mismatch, Match*) in either experiment: Experiment 1A, $F(2, 78) = 0.633$, $p = .534$, $\eta_p^2 = .016$; Experiment 1B, $F(2, 78) = 0.056$, $p = .946$, $\eta_p^2 = .001$.

Manual RT. The critical measure was mean search time as a function of singleton-match condition (Figure 2 and Table 2). The analysis was limited to correct search trials, and trials with RTs more than 2.5 *SD* from the participant's mean in each condition were removed (2.8% of the correct search trials in Experiment 1A; 2.6% in Experiment 1B). The pattern of results was not influenced by trimming in any of the experiments reported in this study. The analyses included trials on which the participant was either correct or incorrect on the ultimate memory test, since the memory test required a much more precise representation of color than would have been needed to guide attention (Hollingworth et al., 2013b). Analyses limited to memory-correct trials produced the same pattern of results as the full analysis in all experiments.

Table 2
Search Accuracy (%), Search RT (ms), and Memory Accuracy (%) for the Array-Based Search Experiments (1A, 1B, 2, and 4) as a Function of Singleton-Match Condition

Experiment	Search Accuracy (*SE*)	Search RT (*SE*)	Memory Accuracy (*SE*)
Experiment 1A			
No singleton	92.8 (0.96)	1240 (37.3)	70.4 (1.14)
Mismatch	92.3 (1.16)	1258 (36.3)	70.0 (1.21)
Match	92.1 (1.21)	1287 (40.8)	67.9 (1.51)
Experiment 1B			
No singleton	94.5 (0.65)	1137 (32.9)	75.8 (1.39)
Mismatch	94.4 (0.83)	1171 (33.7)	74.5 (1.54)
Match	94.3 (0.86)	1193 (35.8)	71.7 (1.40)
Experiment 1 (no memory control)			
No singleton	95.3 (0.60)	1098 (39.1)	—
Mismatch	95.6 (0.65)	1110 (42.9)	—
Match	95.2 (0.63)	1091 (39.0)	—
Experiment 2			
No singleton	95.4 (0.76)	977 (31.7)	79.4 (1.45)
Mismatch	95.5 (0.82)	990 (31.5)	79.9 (1.63)
Match	95.8 (0.69)	1013 (37.0)	77.9 (1.50)
Experiment 4			
No singleton	95.5 (1.40)	1171 (41.0)	83.6 (1.72)
Mismatch	95.7 (1.00)	1222 (56.0)	85.1 (1.83)
Match	96.3 (0.79)	1274 (54.4)	82.1 (2.80)

There were two planned contrasts. The first was between the *No-singleton* and *Mismatch* trials to assess the effect of the mere presence of a color singleton in the array. The second was between the *Mismatch* and *Match* trials to assess the critical effect of match to the memory color, controlling for the presence of a color singleton. For Experiment 1A, there was no reliable difference between mean RT in the *No-singleton* (1240 ms) and *Mismatch* (1258 ms) conditions, $t(39) = 1.07$, $p = .290$, $\eta_p^2 = .029$, although the numerical difference was in the direction expected by singleton capture. Critically, mean RT was reliably higher in the *Match* condition (1287 ms) than in the *Mismatch* condition, $t(39) = 2.44$, $p = .019$, $\eta_p^2 = .132$, indicating attentional capture by the array item matching the secondary color maintained in VWM. For Experiment 1B, there was a reliable difference between mean RT in the *No-singleton* (1135 ms) and *Mismatch* (1170 ms) conditions (see Figure 2 and Table 2), $t(39) = 4.32$, $p < .001$, $\eta_p^2 = .324$, consistent with capture by any color singleton. Critically, mean RT was again reliably higher in the *Match* condition (1191 ms) than in the *Mismatch* condition, $t(39) = 2.25$, $p = .030$, $\eta_p^2 = .115$, indicating memory-based attentional capture by the secondary color maintained in VWM.

We also examined the RT data in the *Match* condition as a function of whether the match was exact or inexact. There was no reliable difference between these subconditions in either experiment: Experiment 1A, $t(39) = 1.63$, $p = .111$, $\eta_p^2 = .064$; Experiment 1B, $t(39) = 1.22$, $p = .228$, $\eta_p^2 = .037$. The absence of an effect of *Exact/Inexact Match* is consistent with previous reports (e.g., Hollingworth & Beck, 2016; Hollingworth et al., 2013b).

Memory accuracy. Mean percent correct on the memory task was 69.7% in Experiment 1A and 74.4% in Experiment 1B (see Table 2). In Experiment 1A, accuracy did not reliably differ between the three conditions (no singleton, match, mismatch), $F(2, 78) = 2.31$, $p = .106$, $\eta_p^2 = .056$. In Experiment 1B, there was a reliable effect of condition on memory accuracy, $F(2, 78) = 9.92$, $p < .001$, $\eta_p^2 = .203$. There was a nonsignificant trend for lower accuracy in the *Mismatch* condition (74.3%) than in the *No-singleton* condition (75.8%), $t(39) = 1.95$, $p = .059$, $\eta_p^2 = .087$, and accuracy was reliably lower in the *Match* condition (71.7%) than in the *Mismatch* condition, $t(39) = 2.63$, $p = .012$, $\eta_p^2 = .150$. This pattern is consistent with the possibility that capture of attention during search impaired memory performance. Note that a similar numerical pattern was observed in Experiment 1A.

We also examined memory accuracy in the *Match* condition as a function of whether the match was exact or inexact. In Experiment 1A, memory accuracy was reliably higher on *Exact Match* (70.0%) than on *Inexact Match* (64.9%) trials, $t(39) = 3.10$, $p = .004$, $\eta_p^2 = .198$. A similar effect was observed in Experiment 1B: *Exact Match* (73.3%), *Inexact Match* trials (70.1%), $t(39) = 2.17$, $p = .036$, $\eta_p^2 = .108$. This pattern is consistent with the capture of attention by the memory matching item, leading to a higher probability of memory encoding for the color of the search item (Schmidt, Vogel, Woodman, & Luck, 2002; Scholl, 2000) or to the

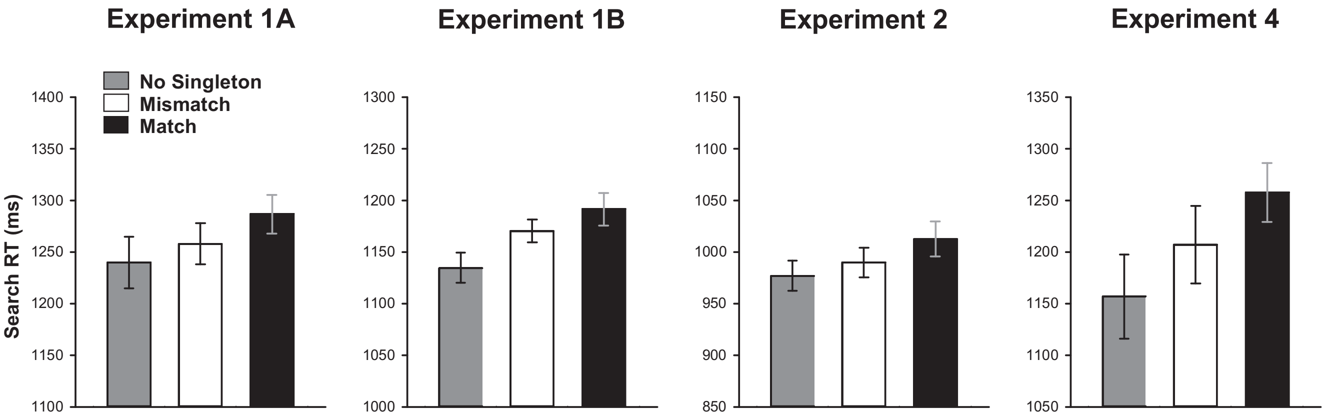

Figure 2. Mean search RT results as a function of singleton-match condition for Experiments 1A, 1B, 2, and 4. Error bars are condition-specific, within-subject 95% confidence intervals (Morey, 2008).

biasing of the memory color value toward the array color value (Schneegans, Spencer, Schöner, Hwang, & Hollingworth, 2014). At test, there would be a greater probability that when retrieving the relevant color, participants would retrieve a color more similar to the array item than to the original memory item. This would increase performance on *Exact Match* trials and decrease performance on *Inexact Match* trials. It is also possible that the accuracy difference reflected a strategy of attending to matching items in an effort to improve performance on the memory test. However, there are two reasons that this is unlikely. First, if participants based their response to the memory test on the color of the array item rather than the color of the memory item, they should have performed significantly *below* chance on the memory test in the *Inexact Match* condition, since the inexact color matched the memory foil color. Second, if a strategy of attending to similar-color array items was the cause of the memory-based capture effect, then there should have been a positive relationship between a participant's memory accuracy difference between *Exact* and *Inexact Match* trials (*Exact* accuracy − *Inexact* accuracy) and the size of a participant's capture effect (*Match* RT – *Mismatch* RT). Yet, there was no such relationship in either experiment: Experiment 1A, $r = .031$, $t(38) = 0.190$, $p = .851$; Experiment 1B, $r = -.091$, $t(38) = -0.561$, $p = .578$.

Discussion

Experiment 1 was designed to optimize sensitivity to memory-based capture from a secondary item maintained in VWM while also replicating the basic methods of earlier capture studies (Downing & Dodds, 2004; Houtkamp & Roelfsema, 2006; Olivers, 2009). Indeed, there was reliable memory-based capture in the *Match* condition (compared with the *Mismatch* condition), both when the memory item and search cue were presented sequentially (Experiment 1A) and simultaneously (Experiment 1B). Color was the dimension on which the distractor matched the secondary memory item. This serves to clarify the results of Houtkamp and Roelfsema (2006), who also found color-based capture, but only in a subset of conditions. The results are consistent with the MIT, holding that multiple items in VWM (the search target and a secondary color) can interact with selection simultaneously.[4]

Experiment 2

In Experiments 1A and 1B, reliable memory-based capture was observed when the match to VWM was on the dimension of color. In contrast, the ambiguous effects of memory match observed by Houtkamp and Roelfsema (2006) and the null effects observed by Downing and Dodds (2004) came primarily from paradigms in which the match to VWM was on the dimension of shape. Thus, it is possible that shape memory is less efficient in capturing attention. To test this, in Experiment 2 we reversed the roles of shape and color in the Experiment 1 paradigm (see Figure 3). Participants remembered a shape for a within-category memory test. For the search task, they saw a color cue and searched for the target color among a set of colored disks. In the *Match* and *Mismatch* conditions, a shape-singleton distractor either matched or mismatched the category of the shape held in VWM.

---

[4] In memory-based capture experiments, it is important to eliminate the possibility that the match effect was attributable to low-level priming from the mere appearance of the memory color before search. To do this, we ran a control experiment ($N = 40$, 20 female) that was identical to Experiment 1A, except there was no memory test at the end of the trial; the trial ended upon response to the search array. Thus, although participants saw the colored disk at the beginning of the trial, there was no demand to maintain the color in VWM across the search task. Participants were informed that the appearance of the colored square signaled that the trial was about to begin. The data are reported in Table 2. Overall search accuracy was 95.3% correct, and there was no effect of singleton-match condition, $F(2, 78) = 0.389$, $p = .679$, $\eta_p^2 = .010$. For Manual RT, there was no reliable difference between the *No-singleton* (1098 ms) and *Mismatch* (1110 ms) conditions, $t(39) = 1.10$, $p = .278$, $\eta_p^2 = .030$. Critically, there was no reliable difference between the *Match* (1091 ms) and *Mismatch* conditions, $t(39) = 1.54$, $p = .131$, $\eta_p^2 = .058$. In sum, when the demand to remember the color disk was eliminated, the memory-based capture effect was also eliminated. Thus, capture in the main experiment was likely to have been caused by active memory maintenance and was unlikely to have been caused by low-level priming. Dependence of capture on active maintenance in VWM is consistent with several additional studies that have used similar tests (Bahle et al., 2017; Hollingworth & Luck, 2009; Hollingworth, Matsukura, & Luck, 2013a; Hollingworth et al., 2013b).

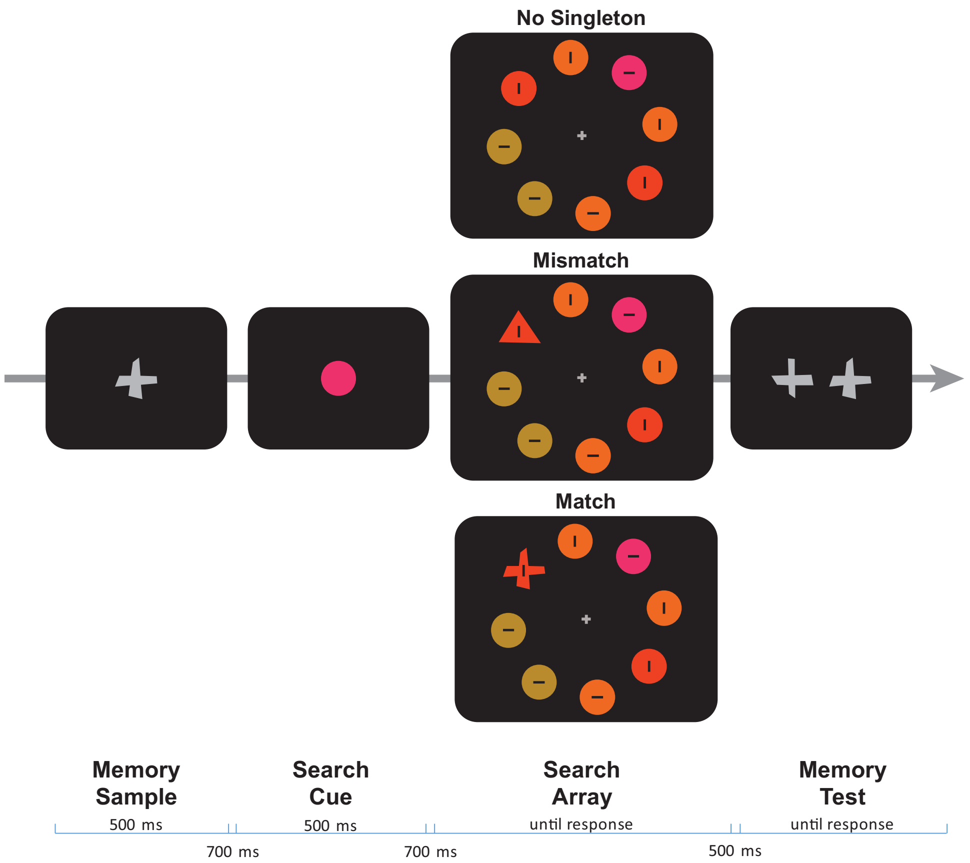

Figure 3. Sequence of events in a trial and singleton-match manipulation for Experiment 2. See the online article for the color version of this figure.

Method

Participants. Forty new participants from the University of Iowa community completed the experiment for course credit. Each was between the ages of 18 and 30 and reported normal or corrected to normal vision. Two participants were replaced, one who failed to perform significantly above chance on the memory task and one who failed to perform significantly above chance on the search task. Of the final 40, 32 were female.

Stimuli and procedure. For the memory task, there were 12 possible shapes, with three similar versions in each of four shape categories: cross, diamond, star, and triangle. The different versions of each shape were derived from a canonical version: Each vertex of the canonical shape was jittered by a random value between −.12° and .12° visual angle. The memory shape was selected randomly from this set of 12 on each trial, and the foil shape in the memory test was selected using the same method as in Experiment 1.

For the search task, the search cue was a single colored disk, with a color selected randomly from four possible colors that could appear in the search array ($x = 0.44$, $y = 0.47$, 59.6 cd/m^2; $x = 0.57$, $y = 0.38$, 48.6 cd/m^2; $x = 0.61$, $y = 0.24$, 48.0 cd/m^2; $x = 0.51$, $y = 0.25$, 50.3 cd/m^2). These same four colors appeared on every trial. They were selected from a circular color space (CIE L^*a^*b space centered at $L = 60$, $a = 18$, $b = 36$; radius = 65), with each color value evenly spaced along a continuous section of the wheel, separated by 35°. Thus, interitem similarity was relatively high. This selection was developed based on pilot testing to equate, approximately, search times in this experiment and in Experiment 1. One color in the array matched the cue color. The colors of the remaining seven disks were chosen randomly from the three nontarget colors. In the *No-singleton* condition, all of the array objects were disks. In the *Match* condition, one of the colored items was a singleton shape that matched the category of the remembered shape (again, the match could be exact or inexact). In the *Mismatch* condition, the singleton shape was drawn from one of the two shape categories not used for the memory test.

The procedure was the same as in Experiment 1A, except participants remembered a grayscale shape and searched through the array for a colored object that matched the search target color.

Results

Search accuracy. Overall search accuracy was 95.5% correct (see Table 2), and there was no effect of singleton-match condition (*No-singleton, Mismatch, Match*), $F(2, 78) = 0.449$, $p = .640$, $\eta_p^2 = .011$.

Manual RT. The analysis was limited to correct search trials, and trials with RTs more than 2.5 *SD* from the participants' mean in each condition were removed (2.9% of the correct search trials).

There was no reliable difference between mean RT in the *No-singleton* (977 ms) and *Mismatch* (990 ms) conditions (see Figure 2 and Table 2), $t(39) = 1.38$, $p = .175$, $\eta_p^2 = .047$. Mean RT was reliably higher in the *Match* condition (1013 ms) than in the *Mismatch* condition, $t(39) = 2.03$, $p = .049$, $\eta_p^2 = .096$, indicating memory-based attentional capture by the secondary shape maintained in VWM.[5]

We also examined the RT data in the *Match* condition as a function of whether the match was exact or inexact. There was no reliable difference between these subconditions, $t(39) = 1.30$, $p = .201$, $\eta_p^2 = .042$, with mean RT of 1021 ms in the *Exact Match* condition and 1004 ms in the *Inexact Match* condition.

Memory accuracy. Mean percent correct on the memory task was 79.1% (see Table 2). There was a marginal effect of singleton-match condition, $F(2, 78) = 3.08$, $p = .051$, $\eta_p^2 = .073$. There was no reliable difference between the *Mismatch* condition (79.9%) and the *No-singleton* condition (79.4%), $t(39) = 0.74$, $p = .466$, $\eta_p^2 = .014$. There was a trend toward lower accuracy in the *Match* condition (77.9%) than in the *Mismatch* condition, $t(39) = 1.94$, $p = .060$, $\eta_p^2 = .080$, again consistent with memory impairment in conditions with greater capture.

As in previous experiments, memory accuracy was reliably higher on *Exact Match* (79.8%) than on *Inexact Match* (76.0%) trials, $t(39) = 2.77$, $p = .009$, $\eta_p^2 = .164$. There was no reliable relationship between the size of this *Exact-Inexact* difference and the size of the participants' memory-based capture effect (*Match* RT − *Mismatch* RT), $r = -.068$, $t(38) = -0.418$, $p = .679$.

Discussion

A secondary shape stimulus maintained in VWM for a later memory test produced reliable capture when that shape appeared as a distractor in a search array, even though the search target color changed from trial to trial and, under the SIT, should have been maintained in VWM as the template. The results indicate that the secondary item influenced selection, consistent with the MIT. Moreover, the difference between our results and those of previous studies is unlikely to have been caused by differences in the remembered feature dimension.

Omnibus Analysis of Experiments 1 and 2

The key effect (higher RT in the *Match* condition than in the *Mismatch* condition) was observed in each of Experiments 1A, 1B, and 2. To obtain a precise estimate of the effect size for these array-based capture experiments, we combined the data in an omnibus analysis, treating experiment as a between-subjects factor. There was a reliable difference between *Match* and *Mismatch* conditions, $F(1, 117) = 15.0$, $p < .001$, $\eta_p^2 = .114$, and no interaction between match condition and experiment, $F(2, 117) = 0.141$, $p = .869$, $\eta_p^2 = .002$. The omnibus effect size ($\eta_p^2 = .114$) falls in the "medium" range (Cohen, 1988). Note that with this effect size, an N of 63 would be required achieve 80% power. Thus, even with an N of 40, our experiments were, individually, somewhat underpowered.[6] This highlights the difficulty in making confident inferences from previous studies that observed null or ambiguous effects using much smaller sample sizes (see Table 1).

Experiment 3

In Experiments 1 and 2, we eliminated two plausible explanations for the difference between our capture effects and the null or ambiguous effects reported by Houtkamp and Roelfsema (2006); Downing and Dodds (2004), and Olivers (2009, Experiment 5). There are several differences remaining between our method and the individual methods of one or more of these experiments. However, trying to pinpoint the precise source could be a long and ultimately futile endeavor, particularly if earlier null and ambiguous effects were simply the result of insufficient power. Moreover, our results are at least partially consistent with those of Houtkamp and Roelfsema, who reported some evidence indicating guidance by a secondary memory item. Thus, we took a different approach in Experiment 3. We attempted to establish the generality of the capture effect in a different search paradigm (search through natural scenes) and using a different dependent measure (oculomotor capture). This method had the further benefit of eliminating one of the final remaining differences between our experiments and the methods of all previous studies: when there was a memory-matching distractor in the search display, it was no longer a singleton on the matching dimension (see Table 1).

As illustrated in Figure 4, participants searched for a target object in a natural scene, reporting the orientation of a letter superimposed upon it. Before search, they saw a cue displaying a picture of the target object. As in previous experiments, the target changed on each trial; under the SIT, the cue representation should have occupied the active, "template" role in VWM. The search task was flanked by a color memory task similar to that in Experiment 1. The primary manipulation was the match between the remembered color and the color of a critical distractor object in the scene. We examined two measures of capture. The primary measure was the probability of fixating the critical distractor. Capture of attention by a memory-matching item should increase the probability that the distractor was fixated. The secondary measure was overall search time, operationalized as the time until the first fixation on the target object. This measure is limited because total search time through a scene is influenced by many factors, includ-

---

[5] Consistent with the standard of full disclosure, we report the results from two partial experiments conducted while refining the Experiment 2 method. The first used a set of highly discriminable and categorically unique colors for the search task. This experiment was discontinued after 18 participants who met performance criteria, because the difficulty of the search task was not well equated with that in Experiment 1. That is, mean search RT was much lower in this experiment than in Experiment 1: *No-singleton* = 812 ms, *Mismatch* = 827 ms, *Match* = 841 ms. The contrast between *Mismatch* and *Match* conditions was not reliable in this sample, $t(17) = 1.43$, $p = .173$, $\eta_p^2 = .106$. We then selected the search colors from a small region of color space to increase the difficulty of the search task. This second partial experiment was discontinued after 17 participants, because the search task became too difficult, leading to longer RTs and substantially lower search accuracy ($M = 86\%$ correct) than in Experiment 1. Accuracy fell below the standard criterion for RT experiments in our laboratory ($M = 90\%$ correct). The condition means were as follows: *No-singleton* = 1371 ms, *Mismatch* = 1388 ms, *Match* = 1373 ms. The contrast between *Mismatch* and *Match* conditions was not reliable in this sample, $t(16) = .628$, $p = .540$, $\eta_p^2 = .024$. The region of color space for selection of the search stimuli was then expanded in Experiment 2.

[6] In the two partial experiments reported in Footnote 3—which used smaller sample sizes (18 and 17) and were rejected for reasons unrelated to the effect of interest—we did not observe reliable capture effects.

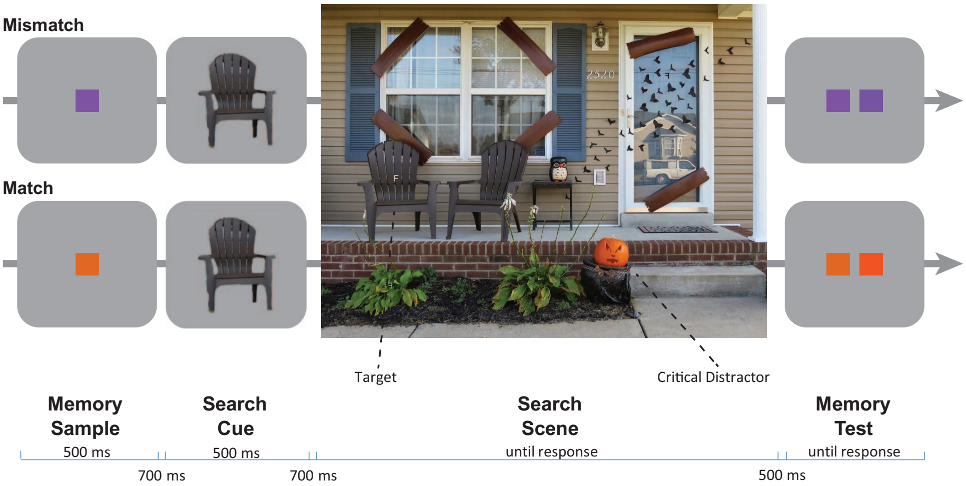

Figure 4. Sequence of events in a trial and match manipulation for Experiment 3. See the online article for the color version of this figure.

ing idiosyncratic differences between scene items. Moreover, in the present implementation, there were only 24 observations per cell (scenes were not viewed more than once) compared with 80 observations per cell in Experiment 1 and 2. Nevertheless, attentional capture by a memory-matching item should generally produce longer overall search times in the *Match* condition than in the *Mismatch* condition.

Method

Participants. Twelve new participants from the University of Iowa community completed the experiment for course credit. This *N* was chosen on the basis of a similar experiment (Bahle et al., 2017, Experiment 1 picture-cue condition, $N = 20$).[7] The effect size from this previous study ($\eta_p^2 = .684$) indicated that, to detect the effect of memory match on the probability of distractor fixation, a minimum of six participants would be necessary to achieve 80% power. Thus, the choice of 12 participants was conservative. Each was between the ages of 18 and 30 and reported normal or corrected to normal vision (we excluded participants who needed contact lenses to achieve normal vision). Two participants were replaced because they did not perform significantly above chance on the memory task. Of the final 12 participants, eight were female.

Apparatus. The position of the right eye was monitored using an SR Research Eyelink 1000 eyetracker, sampling at 1000 Hz. Otherwise, the apparatus was the same as in Experiments 1 and 2.

Stimuli and procedure. Ninety-six photographs of real-world scenes constituted the stimulus set (primarily indoor environments; see Bahle et al. (2017) for a complete list of scene items and target objects), each subtending 26.32° × 19.53°. The target object in each scene was chosen as clearly visible and identifiable (e.g., heavily occluded objects, objects in deep shadow, or objects with atypical appearance for their category were not considered as potential targets). None of the targets appeared at the center of the scene. Targets subtended between 1.63° × 1.59° and 9.30° × 7.56° visual angle, with a mean of 3.27° × 3.34°. The mean eccentricity of targets (screen center to object center) was 7.62°. Superimposed on the search target object in each scene was a left or right facing ‘F’ in Arial font, subtending 0.25° × 0.41°. The ‘F’ was either black, white, or gray, chosen to ensure visibility when superimposed over each target. This also ensured that participants could not search for the target letter solely based on luminance or contrast. To ensure that participants searched for the depicted target object (rather than just searching for an ‘F’), distractor ‘F’s (also black, white, or gray) were superimposed on two other objects in each scene. These objects were randomly selected from a set of eight objects in each scene that could plausibly serve as a target. The orientations of the target ‘F’ and the two distractor ‘F’s were randomly determined.

Half (48) of the scenes were experimental items and half filler items. The experimental items contained a critical distractor object. VWM content was manipulated so that this object either matched or mismatched a color maintained in VWM. In the filler scenes, no object was a close match for the color maintained in VWM. The critical distractor in each experimental scene was chosen to have a relatively uniform color across its surface, which varied across the set of scene items. The distractors ranged from subtending 1.63° × 1.59° to 9.40° × 8.70°, with a mean of 3.34° × 3.86°. The critical distractor never had a superimposed ‘F.’ Note that the images for a particular scene item were identical in the *Match* and *Mismatch* conditions: VWM-match was manipulated by changing the remembered color, not the scene.

The memory color square was presented centrally, subtending 1.64° × 1.64°. On *Match* trials, the memory color was the average RGB color value across all pixels of the critical distractor. Thus, there was rarely any major part of the object that was an exact match with the remembered color, and it is therefore unlikely that participants attended to the distractor strategically to improve

---

[7] This experiment was designed to answer a different question (concerning the relative roles of template- and gist-based guidance during search through scenes). The present Experiment 3 differed in the order of the presentation of the memory item and search cue stimuli, in the absence of a category label accompanying the picture cue, in the presence of distractor “F”’s in the display, and in minor aspects of stimulus timing (see Bahle et al., 2017).

memory performance. On *Mismatch* trials, the memory color was selected from a different color category than on *Match* trials. The memory colors associated with a scene item (*Match* and *Mismatch*) were consistent across participants. For the end-of-trial memory test, two colored squares were presented to the left and right of central fixation (randomly assigned). One of the colored squares was an exact match to the color presented for memorization. The other colored square varied from the exact match square by ± 25 on each of the three RGB channels, with the ± direction determined randomly for each channel. If an increment of + or −25 was not possible because of boundary limitations, the value was selected in the reverse direction. The magnitude of the color difference was piloted to generate memory accuracy similar to that observed in Experiments 1 and 2.

The search cue consisted of an image of the target object presented at the center of the display. The target image was extracted from the scene itself and was presented at the same size as it would appear in the scene.

Participants were instructed to search each scene for the object that matched the picture cue and to report whether the "F" was normally oriented or mirror-reversed. They were also instructed that if there was an object in the scene with a color similar to the memory square, this object would never contain the target "F". The eyetracker was calibrated at the beginning of the session and was recalibrated as necessary throughout the experiment.

The experimenter waited until the participant fixated the screen center and then pressed a silent button to initiate each trial. After a delay of 400 ms, a fixation cross appeared for 400 ms, followed by the memory item for 500 ms, a 700-ms ISI, the search cue for 500 ms, a 700-ms ISI, the search scene until response, a 500-ms ISI, and the memory test display until response. Participants pressed the right button to indicate a normally oriented "F" and the left button to indicate a mirror-reversed "F". They used the same buttons to indicate whether the color square on the left or right was an exact match for the memory color. On trials with an incorrect memory test response, a message with a central, red "incorrect" was displayed for 300 ms. Feedback was not provided for the search task.

Participants first completed a six-trial practice session. Then, they completed an experimental session of 96 trials: 24 *Match* trials, 24 *Mismatch* trials, and 48 filler trials. Trial order was randomly determined. Participants saw each scene once. Across the experiment, each scene appeared in each condition an equal number of times, with the assignment of scenes to the two match conditions counterbalanced across pairs of participants. The entire session lasted approximately 40 min.

Results

The primary analyses concerned eye movement measures. Saccades were defined by a combined velocity (30°/s) and acceleration (8000°/s^2) threshold. Eyetracking data were analyzed with respect to two regions of interest: the *target region* and the *critical distractor region*, which never overlapped. Both regions were rectangular and extended approximately 0.3° beyond the edges of the target and critical distractor objects, respectively. Trials were eliminated from further analysis if the very first fixation on the scene fell in one of the two regions of interest (rather than at the center of the screen), if the target object was not fixated during search, if the search response was incorrect, or if the search time on a trial (elapsed time until target fixation) was more than 2.5 *SD* from the participant's condition mean. A total of 8.7% of trials was eliminated. Trial elimination did not alter the pattern of results.

We first report oculomotor capture results, quantified as the probability of fixating a critical distractor on *Match* compared with *Mismatch* trials. Then, we report global measures of target fixation and search time.

Distractor fixation. Participants were more likely to fixate the critical distractor when they were remembering a color that matched the color category of that object than when they were remembering a mismatching color (Figure 5A). The mean probability of fixating the critical distractor object during search was reliably higher on *Match* trials (.236) than on *Mismatch* trials (.110), $t(11) = 4.45$, $p < .001$, $\eta_p^2 = .643$.

To examine the evolution of this effect across the course of search, we calculated the probability of fixating the critical distractor at each ordinal fixation number (Figure 5B). Fixation 1 was defined as the first participant-controlled fixation (after the first saccade following scene onset). This analysis was limited by the fact that there was substantial variation in the number of fixations in a trial before the target was found and the trial ended, yielding a smaller and smaller number of observations as ordinal fixation increased. Thus, we included an ordinal fixation bin in the Figure 5 only if a participant contributed at least 16 of 24 trials to the bin and only if 9 of the 12 participants' data were available. The relatively small number of depicted bins reflects the fact that targets were generally found after only a few saccades. As is evident from the figure, the effect of memory match was observed from the very first subject-controlled fixation.

Target fixation and search RT. Figure 5C shows the cumulative probability of having fixated the target by each ordinal fixation. On most trials, gaze was directed to the target in just a few saccades, which is unsurprising given that participants were given a pictorial cue. Of greater interest, however, there was no obvious influence of match condition on the time-course of target fixation, with similar cumulative functions in the *Match* and *Mismatch* conditions. Further, an analysis of the mean number of fixations until the first fixation on the target revealed no reliable effect of match condition [*Match* = 2.61; *Mismatch* = 2.59; $t(11) = 0.256$, $p = .803$, $\eta_p^2 = .006$], and there was no effect of match condition on the elapsed time until the first fixation on the target [*Match* = 548 ms; *Mismatch* = 558 ms; $t(11) = 0.421$, $p = .682$, $\eta_p^2 = .016$] or on Manual RT [*Match* = 1320 ms; *Mismatch* = 1336 ms; $t(11) = 0.390$, $p = .705$, $\eta_p^2 = .014$].[8] Thus, and intriguingly, robust differences in the probability of critical distractor fixation early in the trial did not produce observable differences in the overall time necessary to complete search. This is likely caused by the fact that the effect of distractor fixation probability was most prominent for the very first saccade on the scene (see Figure 5B). Yet, this saccade was very rarely directed to

---

[8] Note that the sample size in Experiment 4 was chosen to have sufficient power to detect a match effect on the probability of oculomotor capture and may not have had sufficient power to detect effects on more variable, end-of-trial dependent measures, such as RT and elapsed time to target fixation. Nevertheless, we reported inferential statistics to connect the present results with the results of Experiments 1–3.

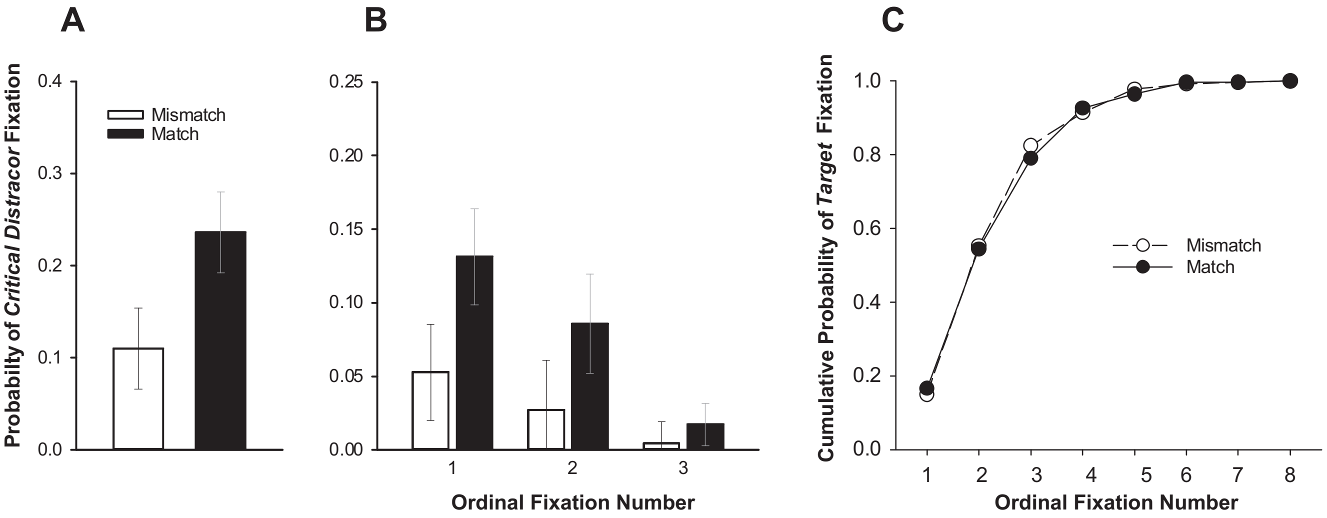

Figure 5. Results of Experiment 3. A: Overall probability of fixating the critical distractor as a function of match condition. B: Probability of fixating the critical distractor for each ordinal fixation during search (fixation 1 is the first participant-controlled fixation after the first saccade on the scene). C: Cumulative probability of fixating the target object for each ordinal fixation as a function of match condition. Error bars are condition-specific, within-subject 95% confidence intervals (Morey, 2008).

the target, even in the *Mismatch* condition (only ~15% of trials, see Figure 5C). That is, the first saccade was typically not guided accurately to the target in either condition. Participants were then able to direct subsequent saccades efficiently to the target in a manner that was very similar in the two match conditions.

Search and memory accuracy. Overall search accuracy was very high and did not differ as a function of match condition, with 96.5% in the *Match* condition and 97.2% in the *Mismatch* condition, $t(11) = 0.320$, $p = .755$, $\eta_p^2 = .009$. In addition, memory accuracy did not differ between the *Match* (72.9%) and *Mismatch* (76.0%) conditions, $t(11) = 1.06$, $p = .313$, $\eta_p^2 = .092$.

Omnibus estimate of effect size. The key effect (higher probability of distractor fixation in the *Match* condition than in the *Mismatch* condition) generated an effect size of $\eta_p^2 = .643$. To obtain a more precise estimate of the effect size for this type of oculomotor, scene-based capture experiment, we combined the data from Experiment 3 with data from the similar experiment reported in Bahle et al. (2017, Experiment 1 picture-cue condition), treating experiment as a between-subjects factor. There was a reliable difference between match conditions, $F(1, 30) = 42.3$, $p < .001$, $\eta_p^2 = .585$. This effect size can be used to guide future research seeking to replicate or extend the present effect.

Discussion

In Experiment 3, we tested the generality of capture by an object matching a secondary item in VWM, using a different search task (search for objects in scenes) and a different primary dependent measure (probability of critical distractor fixation). There was a robust effect of memory match, with the probability of distractor fixation approximately doubled when that object matched the color of a secondary VWM item. Interestingly, the capture effect did not produce observable differences in overall search time: the effect was most prominent for the very first saccade on the scene, which was not itself strongly guided to the target, even in the *Mismatch* condition. This raises the possibility that capture effects early in a search trial will not always be reflected reliably in end-of-trial measures, highlighting the value of using techniques that provide a continuous window on selection processes, such as eye tracking.

Experiment 4

What might account for the difference in effect size between Experiments 1–2 and Experiment 3? Experiment 3 differed from Experiments 1 and 2 on several methodological dimensions, including the oculomotor dependent measure, the use of natural scene stimuli, and the naturalistic search task (real objects instead of abstract colors and shapes). Results from Experiment 3 suggest that the dependent measure may be a key difference: The robust oculomotor capture effect did not produce a reliable difference in overall search time, indicating that oculomotor capture may be more sensitive than end-of-trial measures of search RT. Note that in the similar experiment reported in Bahle et al. (2017), there was a large oculomotor capture effect as a function of memory match ($\eta_p^2 = .648$), but a substantially smaller effect on overall search times ($\eta_p^2 = .281$), although this latter effect was statistically reliable.

Thus, the primary purpose of Experiment 4 was methodological: to identify the experimental conditions most sensitive to capture from a secondary memory item. We returned to the array-based search task of Experiments 1 and 2 but with oculomotor capture as the primary dependent measure. Specifically, the basic method of Experiment 4 was the same as Experiment 1, except that the internal target feature (oriented bar) was made smaller to require target fixation before response, and eye movements were monitored to observe the probability of fixating the critical distractor in the three singleton-match conditions (*No-singleton*, *Mismatch*, and *Match*). A robust effect of match condition on oculomotor capture would provide converging evidence in favor of the MIT and would confirm the utility of oculomotor measures in this type of paradigm (see also Bahle et al., 2017; Le Pelley, Pearson, Griffiths, & Beesley, 2015; Soto, Humphreys, & Heinke, 2006; Theeuwes, Kramer, Hahn, & Irwin, 1998; Theeuwes, Kramer, Hahn, Irwin, & Zelinsky, 1999).

Method

Participants. We used a sample size of 12 (six female), with the expectation that we would observe an effect on oculomotor capture similar to that observed in Experiment 3. Each participant was between the ages of 18 and 30 and reported normal or corrected to normal vision (we excluded participants who needed contact lenses to achieve normal vision).

Stimuli, apparatus, and procedure. The stimuli were the same as in Experiment 1, with the following exceptions. The internal target feature (oriented bar) was made substantially smaller, subtending 0.18° × 0.03°, so that discrimination required fixation of the target shape. In addition, the size of the fixation cross was reduced to 0.38° × 0.38° to promote precise fixation at the screen center before search onset. Finally, the color disks for the memory task were made slightly larger, 1.38° × 1.38°. The apparatus was the same as in Experiment 3. The procedure was the same as in Experiment 1A, with two exceptions. First, there was no articulatory suppression component, as this would have interfered with accurate eye tracking. Second, each trial was initiated by the experimenter (rather than the participant). The experimenter waited until the participant fixated the screen center and then pressed a silent button to initiate the trial. There was a delay of 500 ms before the events depicted in Figure 1. The eyetracker was calibrated at the beginning of the session and was recalibrated as necessary throughout the experiment.

Results

We first report the eye movement results using the same analytical method as employed in Experiment 3. Then we report standard search results (manual RT, accuracy) using the same analytical method as in Experiments 1 and 2 (to allow direct comparison with the results of previous experiments).

Eyetracking data were analyzed with respect to two regions of interest: the *target region* and the *critical distractor region*. Both regions were circular, with a diameter of 2.76°. In the *Match* and *Mismatch* conditions, the critical distractor region corresponded to the color singleton. In the No-singleton condition, the “critical” distractor was selected randomly from the set of 7 distractors. Trials were eliminated from further analysis if the very first fixation on the array fell in one of the two regions of interest (rather than at the center of the screen), if the target object was not fixated during search, if the search response was incorrect, or if the search time on a trial (elapsed time until target fixation) was more than 2.5 *SD* from the participant’s condition mean. A total of 21.8% of trials was eliminated. Trial elimination did not alter the pattern of results.

Distractor fixation. Consistent with Experiment 3, mean fixation probability was reliably higher in the *Match* condition (.312) than in the *Mismatch* condition (.153), $t(11) = 9.76$, $p < .001$, $\eta_p^2 = .897$, indicating capture by an item matching the secondary VWM value (see Figure 6A). Distractor fixation probability was not reliably different in the *Mismatch* condition compared with the *No-singleton* condition (.134), $t(11) = 1.10$, $p = .293$, $\eta_p^2 = .100$.

To examine the evolution of this effect across the course of search, we calculated the probability of fixating the critical distractor at each ordinal fixation number (Figure 6B). Fixation 1 was defined as the first participant-controlled fixation (after the first saccade following array onset). We included an ordinal fixation

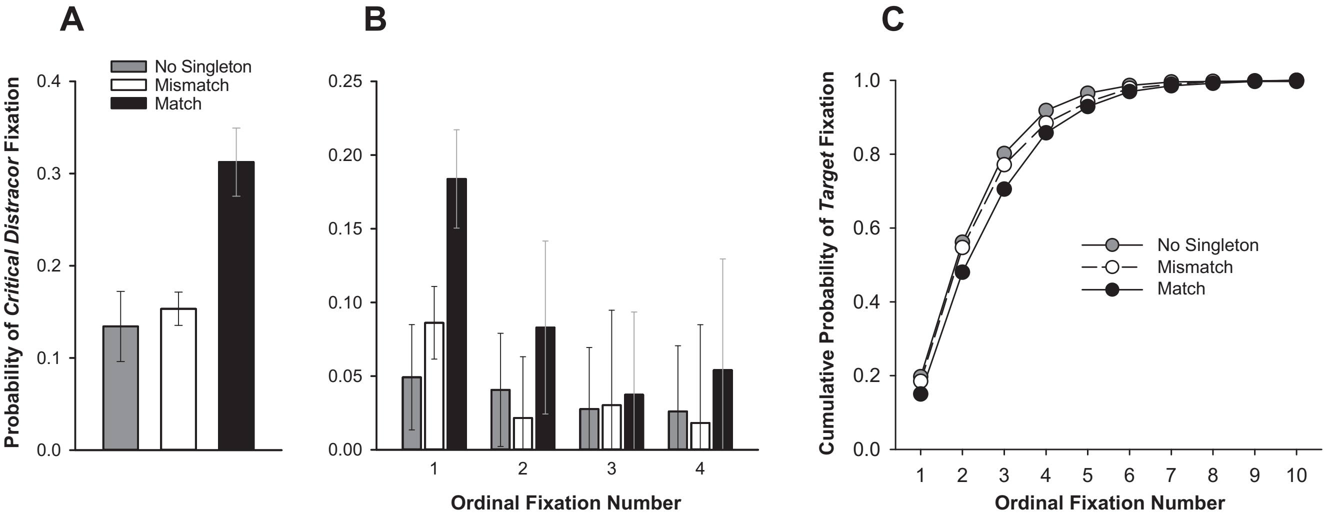

Figure 6. Results of Experiment 4. A: Overall probability of fixating the critical distractor as a function of match condition. B: Probability of fixating the critical distractor for each ordinal fixation during search (fixation 1 is the first participant-controlled fixation after the first saccade on the scene). C: Cumulative probability of fixating the target object for each ordinal fixation as a function of match condition. Error bars are condition-specific, within-subject 95% confidence intervals (Morey, 2008).

bin in the Figure 6 only if a participant contributed at least 20 trials to the bin and if 9 of the 12 participants' data were available. As is evident from the figure, the effect of memory match was observed from the very first subject-controlled fixation.

Target fixation. Figure 6C shows the cumulative probability of having fixated the target by each ordinal fixation. Overall, gaze was directed less efficiently to the target in the *Match* condition compared with the *Mismatch* condition, consistent with the oculomotor capture results. An analysis of the mean number of fixations until the first fixation on the target revealed a reliable effect of match condition [*Match* = 2.99; *Mismatch* = 2.72; $t(11) = 3.05$, $p = .011$, $\eta_p^2 = .457$], and there was a reliable effect of match condition on the elapsed time until the first fixation on the target [*Match* = 604 ms; *Mismatch* = 547ms; $t(11) = 2.87$, $p = .015$, $\eta_p^2 = .427$]. We also compared target fixation measures in the *Mismatch* and *No-singleton* conditions. There was no reliable differences on the elapsed number of fixations measure [*No-singleton* = 2.61; $t(11) = 1.53$, $p = .154$, $\eta_p^2 = .175$] or on the elapsed time to target fixation measure [*No-singleton* = 522 ms; $t(11) = 1.39$, $p = .193$, $\eta_p^2 = .149$]. Note that the sample size in Experiment 4 was chosen to have sufficient power to detect a match effect on the probability of oculomotor capture and may not have had sufficient power to detect effects on more variable, end-of-trial dependent measures, such as elapsed time to target fixation. This applies as well to the manual RT analysis, reported subsequently.

Search accuracy. Search accuracy data are reported in Table 2. Overall search accuracy was 95.8% correct, and there was no effect of singleton-match condition (*No-singleton, Mismatch, Match*), $F(2, 22) = 0.255$, $p = .778$, $\eta_p^2 = .023$.

Manual RT. RT data are reported in Figure 2 and Table 2. The analysis was limited to correct search trials, and trials with RTs more than 2.5 *SD* from the participants' mean in each condition were removed (2.7% of the correct search trials).[9] There was a trend toward a difference between mean RT in the *No-singleton* (1157 ms) and *Mismatch* (1207 ms) conditions, $t(11) = 1.87$, $p = .089$, $\eta_p^2 = .240$. Mean RT was reliably higher in the *Match* condition (1258 ms) than in the *Mismatch* condition, $t(11) = 2.57$, $p = .026$, $\eta_p^2 = .376$, again indicating memory-based capture of attention by the secondary color maintained in VWM. Thus, robust oculomotor capture translated into differences in end-of-trial search measures, in contrast with Experiment 3. This may simply derive from the larger number of observations per cell in Experiment 4 and from greater similarity in the search stimulus from trial to trial, reducing variability in overall search time. Note, however, that in the scene-based search experiment in Bahle et al. (2017), we did observe an effect of memory match on end-of-trial measures, in addition to the oculomotor capture effect. Thus, the absence of end-of-trial effects in Experiment 3 may have been anomalous. Nevertheless, the experiments converge on the conclusion that oculomotor capture is substantially more sensitive to secondary item guidance than is end-of-trial RT.

We also examined the RT data in the *Match* condition as a function of whether the match was exact or inexact. There was no reliable difference between these subconditions, $t(11) = 0.504$, $p = .624$, $\eta_p^2 = .023$, with mean RT of 1264 ms in the *Exact Match* condition and 1250 ms in the *Inexact Match* condition.

Memory accuracy. Mean percent correct on the memory task was 83.6% (see Table 2). There was no effect of singleton-match condition, $F(2, 22) = 2.27$, $p = .127$, $\eta_p^2 = .171$. In addition, there was no reliable difference in memory accuracy on *Exact Match* (84.2%) and *Inexact Match* (80.0%) trials, $t(11) = 1.52$, $p = .157$, $\eta_p^2 = .174$.

Discussion

Using the array-based search method of Experiments 1 and 2, we found robust oculomotor capture by a distractor matching a secondary color in VWM, replicating the oculomotor capture results observed in search through natural scenes (Experiment 3). In addition, there was a reliable effect of match condition on end-of-trial measures of search (elapsed time to target fixation and manual RT), with the RT effects replicating Experiments 1 and 2. Thus, the results provide strong converging support for the MIT. Moreover, they make a substantial methodological contribution: The effect size for oculomotor capture was far larger than that for end-of-trial measures, confirming the efficacy of eyetracking to assess memory-based capture.

General Discussion

The present study provides substantial support for the hypothesis that multiple items in VWM can be maintained in a state that guides perceptual selection. We tested whether a secondary item in VWM captures attention in a search task where the target changes on each trial. The MIT (Beck et al., 2012) holds that both the target and the secondary item have the capability to guide attention. However, under the SIT (Olivers et al., 2011), the target should occupy the single template slot, leaving the secondary item in an inert, accessory state. Our method drew from previous experiments that have found either null or ambiguous secondary-item capture effects in this type of paradigm (Downing & Dodds, 2004; Houtkamp & Roelfsema, 2006; Olivers, 2009, Experiment 5), but we modified these designs to optimize capture sensitivity. Capture by the secondary memory item was observed in five experiments, across multiple types of search task, multiple feature dimensions for the memory task, multiple feature dimensions for the search task, and multiple dependent measures.

Furthermore, several recent studies provide converging evidence in support of the MIT. In the study most similar to the present method, Hollingworth and Beck (2016) used an array-based search task in which the target remained the same across trials. The number of colors retained in VWM was manipulated, as well as the number of VWM-matching distractors in the display. Reliable capture was observed when more than one color was maintained in VWM (Chen & Du, 2017; cf. van Moorselaar et al., 2014). The critical results came from a comparison of two conditions: (a) participants remembered one color, and there was one matching distractor in the array (1Mem/1Match); and (b) participants remembered two colors, and there were two matching distractors in the array (2Mem/2Match). Under the SIT, the latter condition should not generate a larger capture effect than the former, because only one of the two VWM items should be maintained in the template state (and thus there should be only

---

[9] An analysis over manual RT that was limited to the trials used for the eyetracking analysis produced the same pattern of results and the same pattern of statistical significance.

one functional distractor match in the array). Yet, capture magnitude was reliably larger in the 2Mem/2Match condition than in the 1Mem/1Match condition, indicating that both memory items were maintained in a state that interacted with selection, producing capture. This effect was replicated recently by Chen and Du (2017).

Additional converging evidence comes from a gaze-correction study examining the influence of a secondary VWM item on feature-based guidance following a saccade (Hollingworth & Luck, 2009). In the primary task, participants executed a saccade to a target disk (cued by abrupt expansion and contraction) in a circular array of colored disks. On a subset of trials, the array was rotated during the saccade to the target by one half of the angular difference between array items, causing the eyes to land between the target and a differently colored distractor. Because the array was spatially regular and the rotation itself was masked by saccadic suppression, gaze correction required the maintenance of the target color in VWM across the saccade and feature based-guidance of attention to support an accurate corrective saccade (Hollingworth, Richard, & Luck, 2008). With the saccade target representation occupying the putative "template" role in VWM, the critical data came from the manipulation of a secondary VWM color maintained for a concurrent memory task (similar to the present manipulation). Gaze correction to the target was substantially impaired when the color of the adjacent distractor object changed during the saccade to match the color of the secondary VWM item. That is, with a distractor memory match, a substantial proportion of corrective saccades were directed to the distractor rather than to the target (i.e., oculomotor capture), and the mean latency of corrective saccades to the target increased, further suggesting competition from the distractor. These effects indicate that both items in VWM interacted with perceptual selection, consistent with the MIT.

All of the studies discussed so far have probed the key architectural question—whether multiple VWM representations can influence attentional guidance simultaneously—using paradigms designed to detect the capture of attention by memory-matching distractors. These paradigms test whether a secondary item in VWM *automatically* biases selection, which is a particularly strong test. They do not necessarily probe whether it is possible to *strategically* guide attention using multiple VWM items. Studies examining strategic guidance have also supported the MIT. In Beck et al. (2012), participants were asked to attend selectively to items drawn in two different colors within search arrays composed of four colors. They were also instructed, in different blocks, to implement feature-based selection either sequentially (e.g., first the red items and then the blue items) or simultaneously. In the sequential condition, participants fixated relatively long sequences of items in a particular color, and when they switched colors, there was a saccade latency switch cost, potentially indicating template reconfiguration. However, in the simultaneous condition, participants switched between the two relevant colors more frequently, and there was no switch cost, consistent with the use of a multiple-item template.[10] Similar evidence of minimal switch costs have been reported by Grubert, Carlisle, and Eimer (2016) and by Johannesson, Thornton, Smith, Chetverikov, and Kristjánsson (2016).

Beck and Hollingworth (2017) extended this line of work to test a different prediction of the MIT: if participants must select between two saccade targets, both of which match a template value, the two objects should generate substantial competition for selection, as both match an active VWM representation. Participants saw two pairs of stimuli sequentially and, for each pair, made a saccade to the item that matched one of the two target colors on a particular trial (e.g., red and blue). In the first pair, a target color (red) appeared with a distractor color (green). Correct oculomotor selection of the red item then led to the presentation of the second pair, which could contain the same target color (red) along with a new distractor color (*same* condition), the second target color (blue) along with a new distractor color (*switch* condition), or both target colors (red and blue) presented together (*both* condition). First, selection accuracy for the second pair on *same* and *switch* trials was relatively high (>70% correct) and did not differ between conditions. Saccade latency also did not differ between these conditions: that is, there was no switch cost, replicating Beck et al. (2012). Critically, in the *both* condition, selection probability for the second cued color (blue) was approximately equivalent with selection probability for the first cued color (red). That is, the two cue-matching colors were approximately equal candidates for selection, even though the first cued color had guided selection in the first pair and, under the SIT, should have occupied the sole template slot, leading to its efficient selection when paired with the second cued color.

A final line of evidence comes from studies examining the detection of multiple targets in rapid serial visual presentation (RSVP) streams. Houtkamp and Roelfsema (2009) asked participants to detect targets (common objects and colors) in RSVP streams, manipulating whether there were two potential targets or only one. Detection accuracy dropped significantly from the 1-target to the 2-target condition in a manner suggesting that participants maintained only a single target representation. However, this approach concerns the *comparison* of VWM representations to perceptual inputs rather than the *guidance* of attention by VWM. Guidance entails only that perceptual competition is biased by the current state of VWM; it does not entail explicit recognition of particular objects, as in the Houtkamp and Roelfsema task. In an RSVP study that specifically probed attentional guidance, Roper and Vecera (2012, Experiment 3) had participants search for two possible color targets (the target colors changed on a trial-by-trial basis) in an RSVP stream. Before the appearance of the target in the stream, a flanking display was presented that could contain a colored flanker. Capture of attention by the colored flanker should impair detection of the closely following target (Folk, Leber, & Egeth, 2002). The colored distractor either matched one of the two target colors or a nontarget color. Roper and Vecera observed substantial capture by either of the two target colors, relative to the non-target-color control. Although these results are consistent with

---

[10] In a very recent study, Ort, Fahrenfort, and Olivers (2017) found switch costs when only one memory matching alternative was available in the current search array, potentially consistent with template reconfiguration. However, Beck and Hollingworth (2017) found no such effect in a similar design. One major difference between these studies is that Ort et al. had participants search for the same pair of colors over a sequence of 40 searches, and thus their method may not have probed control by VWM, because target repetition leads to a transfer of control from VWM to LTM (Carlisle et al., 2011). In addition, their effects are potentially consistent with a MIT model in which the two representations control attention simultaneously but are maintained at different levels of priority.

the guidance of attention by multiple template items in VWM, unambiguous support for the MIT would require further evidence that participants searched for the two colors simultaneously and did not switch between single-item target representations.

In sum, the results from studies probing the strategic guidance of attention indicate that participants can maintain multiple items in VWM that bias selection simultaneously. The results from the present study indicate that secondary items in VWM have the capability to interact automatically with selection, producing capture. These latter data are critical to the larger theoretical debate, because empirical support for the SIT has typically come from studies probing attention capture.

Although there is now strong evidence to support the MIT, several caveats are in order. First, we do not claim that all items maintained in VWM are necessarily equivalent with respect to attentional guidance. The capture effects observed here were quite small relative to the guidance of attention to the target object. Thus, it is clearly possible to prioritize certain VWM representations for strategic guidance. Moreover, such prioritization can generate circumstances in which the interaction between a deprioritized item and selection can no longer be observed. Hollingworth and Hwang (2013) had participants remember two colors. A postcue indicated which color that was likely to be tested. Subsequently, during the retention interval, the prioritized or deprioritized color could appear as a distractor during a search task. Deprioritized colors produced no observable attentional capture, even on trials when a continuous report procedure indicated that the deprioritized color had been remembered accurately (see also van Moorselaar et al., 2014). Thus, we do not challenge the core assumption of the SIT that visual information can be retained over brief periods of time in states that do and do not interact with attentional selection. Our specific claim is that the active state, which interacts with selection, can span multiple items.

Interestingly, a similar debate has emerged from the general literature concerning guidance of attention during visual search. One of the central assumptions of Wolfe's guided search model (Wolfe, 2007) was that attention can be guided by only one feature from a particular dimension at a time. However, a series of experiments has indicated that participants can implement simultaneous guidance from multiple features on a dimension (Grubert & Eimer, 2015, 2016; Irons, Folk, & Remington, 2012; Stroud, Menneer, Cave, & Donnelly, 2012). The key difference between these studies and present literature on *VWM* guidance is that in the former, the target values were static across the entire experiment or for large blocks of the experiment and were likely to have reflected guidance from a LTM template rather than a VWM template (Carlisle et al., 2011). Thus, VWM and LTM templates appear to share a common principle of multiple-item guidance.[11]

Our results also may have implications for general theories of working memory. One of the key differences between competing theories is whether the active component, or "focus of attention," is limited to a single representation (McElree, 2006; Oberauer, 2002) or can contain multiple representations (Cowan, 2001). These theories were developed primarily in the domain of verbal working memory, and we can draw conclusions only about VWM from the present experiments. Nevertheless, in the domain of VWM, there appears to be no hard, single-item limit on the number of items maintained in an active state, where "active" refers to the ability to interact with sensory processing to guide attention and gaze. This is consistent with recent evidence that multiple remembered feature values can be decoded simultaneously from sustained activation in visual cortex (Emrich et al., 2013). Of course, there may be other visual processes utilizing VWM in which a single-item limit is operational (possibilities include perceptual comparison, long-term memory encoding, and so on), so we cannot conclude from these data that all operations involving VWM can involve multiple, simultaneously active representations. With respect to guidance, single-item, discrete selection is ultimately instantiated by the oculomotor system via fixation; only one object can be fixated at a time. This enables item-level specificity in the mapping of visual objects to internal operations, such as specification of the target of a grasping behavior (see Ballard, Hayhoe, Pook, & Rao, 1997). However, the VWM system that guides selection has the capability to do so based on multiple representations.

Unlike previous experiments (Downing & Dodds, 2004; Houtkamp & Roelfsema, 2006; Olivers, 2009, Experiment 5), we consistently found capture from a secondary item maintained in VWM. What might explain the difference? We implemented almost all the major design features used by at least one of these previous studies. The only remaining, consistent difference was that our memory and search dimensions were always different (e.g., color memory and shape-based search), whereas the memory and search dimensions were always the same in previous studies (see Table 1). However, this potential explanation is unlikely, because in Hollingworth and Beck (2016) and Hollingworth and Luck (2009) we found reliable capture from multiple VWM items on the same dimension (color); the capture effect does not appear to be limited to the case when the two items in VWM come from different feature dimensions. Thus, we cannot identify any single cause for the difference between our effects and those of previous studies. Note, however, that the effect size for RT differences in the array-based capture experiments (Experiments 1 and 2) was only in the medium range. Given the relatively small sample sizes in previous experiments, it is plausible that they were simply underpowered, especially given that the most comprehensive study (Houtkamp & Roelfsema, 2006) used very small sample sizes and found a mixture of reliable, trend-level, and null capture effects.

In comparison with the end-of-trial RT effects, oculomotor capture during search through real-world scenes (Experiment 3, $\eta_p^2 = .643$) and abstract arrays (Experiment 4, $\eta_p^2 = .897$) produced far larger effects. With eye tracking, a capture event can be observed to occur or not occur on each trial, whereas variability in end-of-trial measures, such as RT, does not allow such direct correspondence; RTs can be influenced by many factors in addition to the factor(s) of interest. Moreover, eyetracking allows the researcher to estimate directly the proportion of trials on which capture occurred, which is not possible from aggregate RTs. With these advantages of sensitivity, precision, and transparency, the field might consider moving away from end-of-trial RT measures and toward oculomotor measures (see related discussions in Beck et al., 2017; Zelinsky, Rao, Hayhoe, & Ballard, 1997). In this

---

[11] VWM and LTM templates can differ functionally in other respects, however, such as whether each type of template supports feature-based avoidance (Beck, Luck, & Hollingworth, 2017; Gaspelin, Leonard, & Luck, 2015; Moher, Lakshmanan, Egeth, & Ewen, 2014).

particular domain, our work suggests that the most sensitive and flexible paradigm is the oculomotor capture method implemented in Experiment 4, producing robust effects of memory-based capture while retaining tight control over stimulus and task parameters.

Finally, Experiment 3 informs understanding of the generality of VWM-based attentional guidance and capture (Olivers et al., 2006; Soto et al., 2005). Previously, this area has been studied almost exclusively using simple search stimuli (geometric shapes, colors) presented using either fixed locations or randomly arranged arrays. Experiment 3 indicates that these effects generalize to more complex, real-world stimuli and to a more naturalistic search task. Consequently, theoretical accounts that have been developed to explain attentional guidance using highly controlled stimuli (Bundesen, 1990; Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Hamker, 2004; Schneegans et al., 2014; Wolfe, 1994) may also be likely to generalize to more naturalistic contexts (see also Bahle et al., 2017). Moreover, it may be fruitful to consider the role of VWM content in attentional guidance (both target related VWM content and incidental content) when conducting research in complex, applied domains, such as baggage screening, driving, and interface design, and in clinical manifestations of attention bias.

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16,* 437–443. http://dx.doi.org/10.1016/j.tics.2012.06.010

Bahle, B., Matsukura, M., & Hollingworth, A. (2017). Contrasting gist-based and template-based guidance during real-world visual search. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. http://dx.doi.org/10.1037/xhp0000468

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20,* 723–742. http://dx.doi.org/10.1017/S0140525X97001611

Beck, V. M., & Hollingworth, A. (2017). Competition in saccade target selection reveals attentional guidance by simultaneously active working memory representations. *Journal of Experimental Psychology: Human Perception and Performance, 43,* 225–230. http://dx.doi.org/10.1037/xhp0000306

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science, 23,* 887–898. http://dx.doi.org/10.1177/0956797612439068

Beck, V. M., Luck, S. J., & Hollingworth, A. (2017). Whatever you do, don't look at the . . .: Evaluating guidance by an exclusionary attentional template. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. http://dx.doi.org/10.1037/xhp0000485

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547. http://dx.doi.org/10.1037/0033-295X.97.4.523

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *The Journal of Neuroscience, 31,* 9315–9322. http://dx.doi.org/10.1523/JNEUROSCI.1097-11.2011

Chen, Y., & Du, F. (2017). Two visual working memory representations simultaneously control attention. *Scientific Reports, 7,* 6107. http://dx.doi.org/10.1038/s41598-017-05865-1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cowan, N. (1999). An embedded process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139174909.006

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114. http://dx.doi.org/10.1017/S0140525X01003922

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222. http://dx.doi.org/10.1146/annurev.ne.18.030195.001205

Downing, P. E., & Dodds, C. M. (2004). Competition in visual working memory for control of search. *Visual Cognition, 11,* 689–703. http://dx.doi.org/10.1080/13506280344000446

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. http://dx.doi.org/10.1037/0033-295X.96.3.433

Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *The Journal of Neuroscience, 33,* 6516–6523. http://dx.doi.org/10.1523/JNEUROSCI.5732-12.2013

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics, 64,* 741–753. http://dx.doi.org/10.3758/BF03194741

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychological Science, 26,* 1740–1750. http://dx.doi.org/10.1177/0956797615597913

Grubert, A., Carlisle, N. B., & Eimer, M. (2016). The control of single-color and multiple-color visual search by attentional templates in working memory and in long-term memory. *Journal of Cognitive Neuroscience, 28,* 1947–1963. http://dx.doi.org/10.1162/jocn_a_01020

Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-color and multiple-color visual search. *Journal of Experimental Psychology: Human Perception and Performance, 41,* 86–101. http://dx.doi.org/10.1037/xhp0000019

Grubert, A., & Eimer, M. (2016). All set, indeed! N2pc components reveal simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 1215–1230. http://dx.doi.org/10.1037/xhp0000221

Hamker, F. H. (2004). A dynamic model of how feature cues guide spatial attention. *Vision Research, 44,* 501–521. http://dx.doi.org/10.1016/j.visres.2003.09.033

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458,* 632–635. http://dx.doi.org/10.1038/nature07832

Hollingworth, A., & Beck, V. M. (2016). Memory-based attention capture when multiple items are maintained in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 911–917. http://dx.doi.org/10.1037/xhp0000230

Hollingworth, A., & Hwang, S. (2013). The relationship between visual working memory and attention: Retention of precise colour information in the absence of effects on perceptual selection. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*. Advance online publication. http://dx.doi.org/10.1098/rstb.2013.0061

Hollingworth, A., & Luck, S. J. (2009). The role of visual working memory (VWM) in the control of gaze during visual search. *Attention, Perception, & Psychophysics, 71,* 936–949. http://dx.doi.org/10.3758/APP.71.4.936

Hollingworth, A., Matsukura, M., & Luck, S. J. (2013a). Visual working memory modulates low-level saccade target selection: Evidence from rapidly generated saccades in the global effect paradigm. *Journal of Vision, 13,* 4. http://dx.doi.org/10.1167/13.13.4

Hollingworth, A., Matsukura, M., & Luck, S. J. (2013b). Visual working memory modulates rapid eye movements to simple onset targets. *Psychological Science, 24,* 790–796. http://dx.doi.org/10.1177/0956797612459767

Hollingworth, A., & Maxcey-Richard, A. M. (2013). Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1047–1058. http://dx.doi.org/10.1037/a0030238

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General, 137,* 163–181. http://dx.doi.org/10.1037/0096-3445.137.1.163

Houtkamp, R., & Roelfsema, P. R. (2006). The effect of items in working memory on the deployment of attention and the eyes during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 423–442. http://dx.doi.org/10.1037/0096-1523.32.2.423

Houtkamp, R., & Roelfsema, P. R. (2009). Matching of visual input to only one item at any one time. *Psychological Research, 73,* 317–326. http://dx.doi.org/10.1007/s00426-008-0157-3

Irons, J. L., Folk, C. L., & Remington, R. W. (2012). All set! Evidence of simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 758–775. http://dx.doi.org/10.1037/a0026578

Jóhannesson, O. I., Thornton, I. M., Smith, I. J., Chetverikov, A., & Kristjánsson, Á. (2016). Visual foraging with fingers and eye gaze. *i-Perception, 7,* 2041669516637279. http://dx.doi.org/10.1177/2041669516637279

Le Pelley, M. E., Pearson, D., Griffiths, O., & Beesley, T. (2015). When goals conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *Journal of Experimental Psychology: General, 144,* 158–171. http://dx.doi.org/10.1037/xge0000037

McElree, B. (2006). Accessing recent events. *Psychology of Learning and Motivation: Advances in Research and Theory, 46,* 155–200).

Moher, J., Lakshmanan, B. M., Egeth, H. E., & Ewen, J. B. (2014). Inhibition drives early feature-based attention. *Psychological Science, 25,* 315–324. http://dx.doi.org/10.1177/0956797613511257

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4,* 61–64. http://dx.doi.org/10.20982/tqmp.04.2.p061

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 411–421. http://dx.doi.org/10.1037/0278-7393.28.3.411

Olivers, C. N. L. (2009). What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1275–1291. http://dx.doi.org/10.1037/a0013896

Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1243–1265. http://dx.doi.org/10.1037/0096-1523.32.5.1243

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences, 15,* 327–334.

Ort, E., Fahrenfort, J. J., & Olivers, C. N. L. (2017). Lack of free choice reveals the cost of having to search for more than one object. *Psychological Science, 28,* 1137–1147. http://dx.doi.org/10.1177/0956797617705667

Roper, Z. J. J., & Vecera, S. P. (2012). Searching for two things at once: Establishment of multiple attentional control settings on a trial-by-trial basis. *Psychonomic Bulletin & Review, 19,* 1114–1121. http://dx.doi.org/10.3758/s13423-012-0297-8

Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics, 64,* 754–763. http://dx.doi.org/10.3758/BF03194742

Schneegans, S., Spencer, J. P., Schöner, G., Hwang, S., & Hollingworth, A. (2014). Dynamic interactions between visual working memory and saccade target selection. *Journal of Vision, 14,* 9. http://dx.doi.org/10.1167/14.11.9

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

Scholl, B. J. (2000). Attenuated change blindness for exogenously attended items in a flicker paradigm. *Visual Cognition, 7,* 377–396. http://dx.doi.org/10.1080/135062800394856

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20,* 207–214. http://dx.doi.org/10.1111/j.1467-9280.2009.02276.x

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 248–261. http://dx.doi.org/10.1037/0096-1523.31.2.248

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences, 12,* 342–348. http://dx.doi.org/10.1016/j.tics.2008.05.007

Soto, D., & Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 730–737. http://dx.doi.org/10.1037/0096-1523.33.3.730

Soto, D., Humphreys, G. W., & Heinke, D. (2006). Working memory can guide pop-out search. *Vision Research, 46,* 1010–1018. http://dx.doi.org/10.1016/j.visres.2005.09.008

Stroud, M. J., Menneer, T., Cave, K. R., & Donnelly, N. (2012). Using the dual-target cost to explore the nature of search target representations. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 113–122. http://dx.doi.org/10.1037/a0025887

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science, 9,* 379–385. http://dx.doi.org/10.1111/1467-9280.00071

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1595–1608. http://dx.doi.org/10.1037/0096-1523.25.6.1595

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113,* 766–786. http://dx.doi.org/10.1037/0033-295X.113.4.766

van Moorselaar, D., Theeuwes, J., & Olivers, C. N. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1450–1464. http://dx.doi.org/10.1037/a0036229

Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision, 5,* 81–92. http://dx.doi.org/10.1167/5.1.8

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238. http://dx.doi.org/10.3758/BF03200774

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433. http://dx.doi.org/10.1037/0096-1523.15.3.419

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44,* 1411–1426. http://dx.doi.org/10.1016/j.visres.2003.11.024

Wolfe, J. M., Võ, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15,* 77–84. http://dx.doi.org/10.1016/j.tics.2010.12.001

Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science, 8,* 448–453. http://dx.doi.org/10.1111/j.1467-9280.1997.tb00459.x

Received July 5, 2017
Revision received September 15, 2017
Accepted October 19, 2017 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Call for Nomination: APA Journals Open Science and Methodology Chair

APA's Publications and Communications (P&C) Board has opened nominations for a new Open Science and Methodology (OSM) Chair advisory position.

The OSM Chair will work with the Chief Editorial Advisor and a small ad-hoc committee to recommend refinements and developments to APA Journal policy and procedures relevant to OSM.

The OSM Chair will serve the main roles of communicating and educating authors and editors on APA policies related to OSM, bringing recurring issues to the Council of Editors (COE) and P&C Board for policy consideration, and consulting in cases related to publication issues related to OSM.

Candidates should be members of APA and should be available to serve a two-year term beginning in January 2019.

Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

The search committee is comprised of the P&C Board Chair and Chair-Elect, the COE Chair and Chair-Elect, the Chief Editorial Advisor, and the Journals Publisher.

Qualified candidates will have experience with or knowledge of open science practices (e.g., data-sharing, reproducibility, registered studies), editorial experience, strong publication record, experience with data management, diverse research methodology, clinical trials, and/or service on an Institutional Review Board.

Nominate candidates through APA's Editor Search website.

Prepared statements of one page or less in support of a nominee can also be submitted by email to Rose Sokol-Chang, PhD, Journals Publisher.

Deadline for accepting nominations is **Monday, July 30, 2018**, after which phase one vetting will begin.