American Psychological Association

Journal of Experimental Psychology:
Human Perception and Performance

© 2019 American Psychological Association
0096-1523/19/$12.00

2019, Vol. 45, No. 4, 523–536
http://dx.doi.org/10.1037/xhp0000624

Contrasting Episodic and Template-Based Guidance During Search Through Natural Scenes

Brett Bahle and Andrew Hollingworth
The University of Iowa

Visual search through natural scenes can be guided by knowledge of where a target object has been observed previously (*episodic guidance*) and knowledge of that object's visual properties (*template guidance).* In the present experiments, we compared the relative contributions of these two sources of guidance. Episodic guidance was implemented in a contextual cuing task: participants searched multiple times through a set of scenes for a target letter that appeared in a consistent location within each scene. Template guidance was implemented by the color match between a critical distractor in each scene and a secondary visual working memory (VWM) load. There were four main findings. First, search time decreased with increasing scene repetition; episodic memory guided search. Second, the critical distractor was fixated more frequently on match compared with mismatch trials, consistent with automatic template guidance. Third, the VWM-match effect persisted in blocks with strong episodic guidance. Finally, VWM-match effects were observed from the first saccade during search, whereas episodic guidance to the target developed only later in the trial. The results support a view of natural search in which template-based mechanisms operate early during search in a manner that is not strongly constrained by scene-based forms of guidance, such as episodic knowledge.

Public Significance Statement

Real-world searches are guided by knowledge of where a target object has been observed previously and knowledge of that object's visual features. The present study investigates the interaction between these two sources of guidance during search. By better understanding how these searches are performed, vital tasks in the real world that rely on similar sources of knowledge (e.g., a baggage screener looking for dangerous items or a radiologist looking for tumors) can be potentially improved.

Keywords: visual search, attentional guidance, eye movements, visual attention, contextual cuing

Supplemental materials: http://dx.doi.org/10.1037/xhp0000624.supp

How is attention guided to task-relevant objects in natural scenes? The physical salience of natural objects is only weakly correlated with task relevance (Henderson, Brockmole, Castelhano, & Mack, 2007), so strategic forms of guidance are required. These include knowledge of the visual properties of the target object (template guidance; e.g., Zelinsky, 2008), knowledge of the typical locations of object types (semantic guidance; e.g.,Torralba, Oliva, Castelhano, & Henderson, 2006), and knowledge of the previous location(s) where a target was observed (episodic guidance; e.g., Brockmole, Castelhano, & Henderson, 2006). Traditional theories of visual search (Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Wolfe, 1994) have focused on template guidance; the arrays used in standard search tasks typically contain no semantic or episodic structure. However, recent theories of search through natural scenes have stressed semantic and episodic components, with two prominent accounts holding that these *scene-based* forms of guidance rapidly constrain the spatial scope of search (Torralba et al., 2006; Wolfe, Võ, Evans, & Greene, 2011). In this view, template guidance is limited, primarily, to scene regions that are likely to, or have been observed to, contain the target.

In contrast with these views, Bahle, Matsukura, and Hollingworth (2018) demonstrated that template guidance is applied broadly across a scene in a manner that does not initially distinguish between semantically plausible and implausible regions. A color maintained in visual working memory (VWM; the system typically used to implement template guidance), led to increased fixation of color-matching distractors, even when those objects appeared in implausible locations for the target. This finding suggests that traditional models of visual search (e.g., Wolfe,

---

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

This research was supported by a National Institute of Health grant (R01EY017356) to Andrew Hollingworth.

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

1994) may have greater relevance to this domain than proposed recently (e.g., Wolfe et al., 2011). However, semantic guidance may provide weaker constraint than episodic guidance. We have ample opportunity to learn the locations of specific objects within familiar scenes. The rapid recognition of a scene (Rousselet, Joubert, & Fabre-Thorpe, 2005) could support the efficient retrieval of the known target location (Brockmole & Henderson, 2006), potentially implementing the type of scene-based guidance proposed to dominate the early stages of search (Torralba et al., 2006; Wolfe et al., 2011).

Here, we probed the interaction between episodic guidance and template-based guidance over the course of episodic learning. The first question was whether episodic guidance would come to dominate template guidance with increasing episodic learning (consistent with recent models prioritizing scene-based guidance mechanisms) or whether template guidance would remain potent despite strong episodic guidance (consistent with Bahle, Matsukura et al., 2018). The second question concerned the within-trial time course of the application of the two forms of guidance. In Bahle, Matsukura, et al. (2018), we found that template-based guidance was implemented earlier within a trial (from the very first saccade on the scene) than was sematic guidance. Thus, in the present study, we predicted that template-based guidance would be implemented earlier in a trial than episodic guidance, and that this temporal advantage would be preserved even in the presence of strong episodic learning.

Answering these questions requires a method that can dissociate the scene regions specified by template and episodic guidance. Previous studies have investigated scene-based and template-based sources of guidance simultaneously (Malcolm & Henderson, 2010), but both sources of guidance drove attention to the same scene regions, and thus the method could not distinguish the relative contributions of the two guidance mechanisms, nor could it determine their independent time courses of implementation. In the present two experiments, we monitored eye movements in a search task that supported guidance dissociation. Episodic learning was implemented in a repeated search, contextual cuing design (Brockmole et al., 2006). On each trial, participants searched for a small letter *F* that was superimposed at a different location in each scene. A subset of scenes was repeated across blocks. We expected episodic learning of the letter locations within repeated scenes, leading to a robust contextual cuing effect (Brockmole et al., 2006; Brockmole & Henderson, 2006).

To implement template guidance to different scene locations, we used an attention capture method based on the content of VWM (Bahle, Matsukura, et al., 2018). Simultaneously with letter search, participants maintained a color value in VWM for a memory test at the end of the trial. On some trials, a critical distractor was present in the scene that either matched the color held in VWM (*match* trials) or did not match (*mismatch* trials). A higher probability of fixating matching distractors served as the measure of template guidance. Note that although the color in VWM was not a template per se, VWM is the typical substrate of search templates (Carlisle, Arita, Pardo, & Woodman, 2011), and VWM content guides attention in a similar manner regardless of whether or not that item is the search target (Soto, Heinke, Humphreys, & Blanco, 2005). Thus, we engaged the mechanisms of template guidance for a feature that was not actually the target of search, allowing us to guide attention to scene locations that did not contain the target letter. This implementation of template guidance is conservative, in that capture effects from VWM tend to be smaller than the effects of strategic guidance (e.g., Beck, Luck, & Hollingworth, 2018), which could only work against our prediction of early template guidance throughout the course of episodic learning. Finally, color was chosen for the VWM task because eye movements are more strongly guided by color than by values on other feature dimensions, such as size, shape, or orientation (Rutishauser & Koch, 2007; Williams, 1967). Thus, the types of templates formed during real-world search are likely to contain a representation of color.

To preview the results, we observed robust learning of scene–letter relationships, with searches becoming highly efficient through repetition. We also observed strong capture by VWM-matching distractors. Importantly, this capture effect was not reduced by the increasing influence of episodic guidance. Moreover, the application of episodic guidance within a trial was delayed relative to template-based guidance. The results challenge the central claim of current theories of search through natural scenes (Torralba et al., 2006; Wolfe et al., 2011) and point toward a revised conceptualization of natural search in which template-based processes are applied rapidly and broadly across a scene.

Experiment 1

In Experiment 1, participants searched through both repeated (80%) and nonrepeated (20%) scenes in eight blocks of search for a small, target letter *F*. Simultaneously, they maintained a color in VWM. On 25% of trials, a critical distractor object matched the color held in VWM (Match trials); on the remainder, it did not (Mismatch trials). As in Bahle, Matsukura, et al. (2018), we expected a robust effect of VWM-match on distractor fixation in early blocks of the experiment, before target locations were learned. The critical question was whether this template effect would persist when episodic guidance became available in later blocks. In addition, we probed the within-trial time course of template and episodic guidance to examine their relative influence as search progressed.

Method

Participants. We selected an *N* of 24, three times the *N* indicated by the effect size from two experiments probing a similar VWM-match effect (Bahle, Beck, et al., 2018, Experiment 3; Bahle, Matsukura, et al., 2018, Experiment 1). In these experiments, the omnibus effect size for VWM match was $p\eta^2 = .585$, requiring eight participants for 80% power (calculated using G-Power, Faul, Erdfelder, Lang, & Buchner, 2007). All participants were between the ages of 18 and 30 and reported normal or correct-to-normal vision without contact lenses. Four participants were replaced who failed to perform reliably above chance (57%) on the memory test. Of the final 24 participants, 22 were female.

Stimuli. The scene stimuli were 144 photographs of different common environments. They subtended 26.32° × 19.53° at a resolution of 1280 × 960 pixels. Most were indoor environments such as kitchens and bedrooms. A complete list of scene items is provided in the Appendix, along with several additional sample images in the online supplementary materials. Forty-eight experimental items were repeated-search scenes that also contained a critical distractor object. Each distractor was chosen to have a relatively uniform color across its surface and was the only object with that particular color in the scene, so that any capture effect would be limited to the distractor object. To do this, in some scenes we used photo-editing software to change the color of the distractor. Distractor colors varied across the set of scene items. The distractors ranged from subtending 1.64° × 1.63 to 7.87° × 9.42°, with a mean of 4.30° × 3.61°. None appeared at the center of the scene (mean eccentricity was 7.51°) so that distractor fixation would require at least one saccade. The remaining 96 scenes were novel items that were never repeated and never contained a matching distractor object, serving as a baseline to assess the contextual cuing effect.

The search target was a left- or right-facing *F* (Arial font, 0.25° × 0.41°). Its location was chosen randomly for each scene item with the following constraints: targets were equally likely to appear in four scene quadrants, and the target could not appear within 3° of the center or 1° of the scene edges. The *F* was either black, white, or gray, chosen to ensure that the target letter was not conspicuous at the chosen location. For each scene item, the target *F* position remained the same across repetition. However, the orientation of the *F* was selected randomly on each trial, requiring participants to find the *F* each time to report its orientation. This type of search target was used for several reasons. First, the target *F* was very small so that it could not be resolved in the periphery, its salience against the local background was relatively low, and its specific perceptual properties (black, white, or gray) varied. These methods ensured that participants would not attempt to guide attention to particular scene regions based on a representation of the target letter itself, which could have produced interference with guidance from the content of VWM. Second, the search target was an *F* on every trial. Under such "static target" conditions, the target representation is typically maintained in long-term memory rather than VWM (Woodman, Carlisle, & Reinhart, 2013), and participants could therefore devote VWM exclusively to maintaining the memory color.

For the memory task, the to-be-remembered color square subtended 1.64° × 1.64°. On Match trials, this color was selected to match the mean color across the surface of the critical distractor object. On Mismatch trials, the memory color was chosen from a different color category that did not closely match any object in the scene. Note that color match was manipulated by changing the remembered color, not the color of the distractor in the scene; the scene items were identical on match and mismatch trials. In the memory test at the end of the trial, two colored squares were presented to the left and right of central fixation (randomly selected). The correct alternative matched the sample color. The foil color square varied from the exact match square by ±20 on each of the three RGB channels, with the +/− direction determined randomly for each channel (if a value was not available due to range restrictions in RGB color space, the value for that channel was chosen in the reverse direction).

Apparatus. Stimuli were presented on an LCD monitor (100 Hz refresh) at a fixed viewing distance of 77 cm, maintained by a chin and forehead rest. The right eye was monitored using an SR Research Eyelink 1000 eyetracker sampling at 1,000 Hz. Manual responses were collected with a USB response pad.

Procedure. After arriving for the experiment session, participants provided informed consent. They were instructed that the primary task was to find the *F* in each scene and to report whether the *F* was normally oriented or mirror-reversed. They were also instructed that there would sometimes be an object in the scene with a color similar to the color of the remembered square, but this object would never contain the *F*. The eyetracker was calibrated at the beginning of the session and was recalibrated if the estimate of gaze positon deviated from the central fixation point by more than approximately 0.75°.

The sequence of events in a trial is illustrated in Figure 1. Participants were instructed to complete the search task as quickly and as accurately as possible. If the search task was not completed within 15 seconds, the search task terminated, and the memory test began. Only accuracy was stressed for the memory task. Trials with incorrect memory response were followed by a 500-ms "Incorrect" message. Feedback was not provided for the search task.

Participants first completed a practice block of four trials, followed by eight experimental blocks of 60 trials each. Each block contained 12 experimental scenes in the Match condition, 36 experimental scenes in the Mismatch condition, and 12 novel scenes. There were fewer Match than Mismatch scenes so that participants would not anticipate a matching item in each display. Trial order within each block was determined randomly. Each experimental scene appeared in the Match condition once in the first four blocks and once in the second four blocks. For the remaining six presentations, the scene item appeared in the Mismatch condition. The assignment of scene items to the two match conditions across blocks was randomized across participants.

Data analysis. Saccades were defined by a combined velocity (30°/s) and acceleration (8,000°/s^2) threshold. Eyetracking data were analyzed with respect to two scoring regions: the target region and the critical distractor region. Both were rectangular and extended approximately 1° beyond the edge of the target and 0.3° beyond the edge of the critical distractor. Trials were eliminated if the target was not fixated, if the participant was fixating either region at the onset of search, if the search response was incorrect, or if a search time was greater than 2.5 *SD* from a subject's condition mean (a total of 11% of trials). Mean search accuracy was high (96.5%), but did differ among conditions (Novel: 95.2%, Match: 97.3%. Mismatch: 97.1%), $F(2, 46) = 9.15$, $p < .001$, $p\eta^2 = .285$. However, this was caused by a greater proportion of trials in the novel condition on which the target was not found (i.e., there was no response before the trial timed out after 15 seconds, which was recorded as an incorrect response), consistent with the search time results reported below. Mean memory accuracy was 63.7% and did not differ among conditions (Novel: 62.3%, Match: 65.0%. Mismatch: 63.7%), $F(2, 46) = 2.526$, $p = .091$, $p\eta^2 = .099$. See Table 1 for complete search accuracy and memory accuracy results.

Results and Discussion

First, we present contextual cuing results diagnostic of episodic guidance. We then present the critical results concerning template-based guidance to VWM-matching distractors and the potential interaction of this effect with episodic guidance. Finally, we examine the within-trial time course of the application of episodic and template-based guidance.

Episodic guidance: Search time. Search time (ST) was the elapsed time from scene onset to the beginning of the first fixation

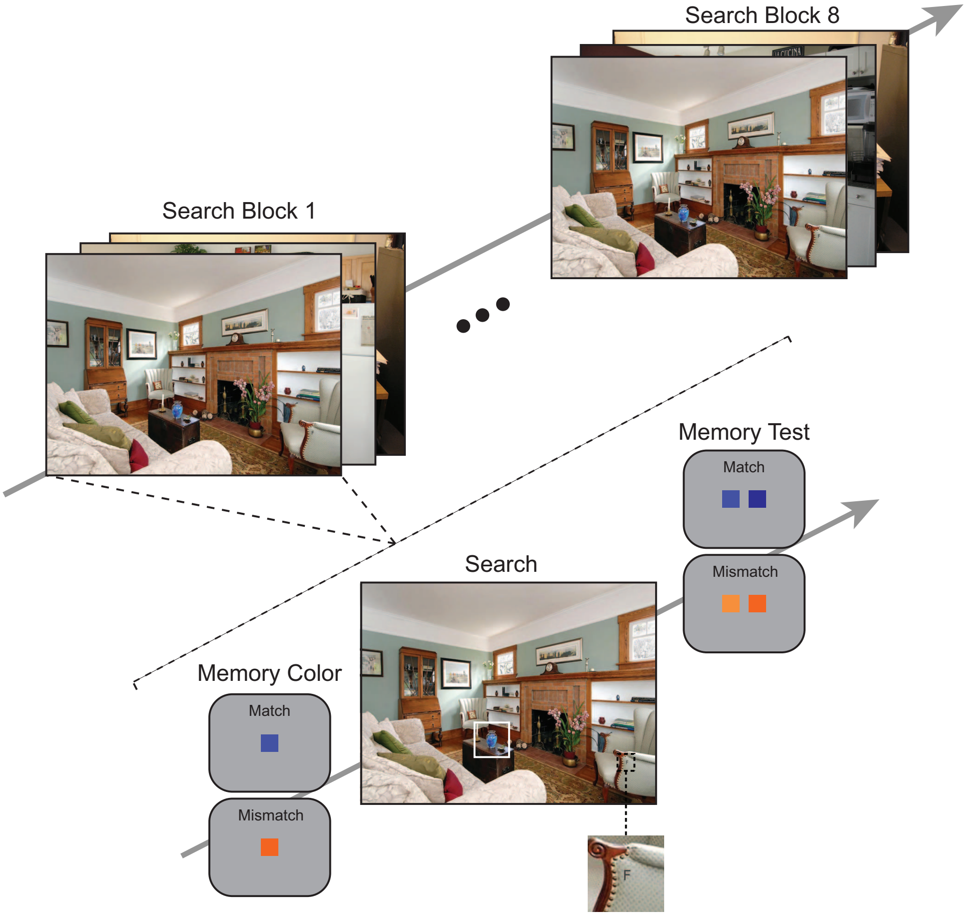

Figure 1. Experimental design. Participants completed eight search blocks, in which 48 scene items were repeated. Each trial of search was initiated by the experimenter. There was a 1,000-ms fixation cross (not pictured), followed by the memory color (500 ms), another fixation cross for 700 ms (not pictured), and the search scene. The scene always contained a target *F* (outlined by black square, not present in experimental image). For each scene item, the *F* was always presented at the same location. Also, the memory color could either match or mismatch a critical distractor object (marked by white square, not present in experimental image). Participants responded by pressing one of two buttons (left/right) to indicate the orientation of the *F*. The trial terminated after 15 s if no response was registered. Following search response, there was a 500-ms delay before the presentation of the memory test display. Participants used the same two buttons to indicate the location (left/right) of the matching color square. See the online article for the color version of this figure.

on the target (see Figure 2). An equivalent pattern of results was observed for both manual response time and average number of fixations before the first target fixation. To examine change in search efficiency across blocks, we fit each participant's ST data for Repeated Mismatch and for Novel scenes with a power function of the form $\mathrm{ST} = ix^{-s}$, where i is the intercept parameter, s the learning rate parameter, and x the search block (Brooks, Rasmussen, & Hollingworth, 2010; Chun & Jiang, 2003).[1] This is consistent with evidence that contextual cueing is well explained by instance theory (Chun & Jiang, 1998) and by the resulting power law function of learning (Logan, 1992). Fits were implemented using the nlsList module of R (Pinheiro & Bates, 2000). We then conducted inferential statistics over the parameter estimates in each of the conditions (see online supplemental materials for individual participant fits).[2]

There was a reliable difference in the intercept parameter, $t(23) = 2.21$, $p = .037$, $p\eta^2 = .175$, with a higher intercept for Repeated Mismatch compared with Novel scenes. This was likely caused by idiosyncratic differences between scene items. Critically, there was a significant difference in learning rate for Repeated Mismatch ($s = .303$) and Novel scenes ($s = .021$), $t(23) = 9.06$, $p < .001$, $p\eta^2 = .781$. Participants learned the repeated target

---

[1] Repeated scenes in the match condition were excluded from this analysis to avoid contamination by the influence of template guidance from VWM. However, an analysis including the match condition produced equivalent results.

[2] Note that the power function was chosen for theoretical reasons. We also fit the data in Experiments 1 and 2 with a simple linear function. The results did not differ materially using the two methods.

Table 1
Search and Memory Accuracy (%) for Experiment 1 as a Function of Distractor Color Match Condition and Block

Condition	Search Accuracy (*SE*)	Memory Accuracy (*SE*)
Block 1		
Repeated Match	93.1 (1.64)	67.7 (2.52)
Repeated Mismatch	95.3 (.86)	63.1 (1.49)
Novel	93.1 (1.48)	60.8 (2.53)
Block 2		
Repeated Match	96.9 (.98)	61.8 (3.25)
Repeated Mismatch	96.2 (.52)	64.4 (1.74)
Novel	94.4 (1.08)	62.5 (2.07)
Block 3		
Repeated Match	97.9 (.90)	62.2 (3.95)
Repeated Mismatch	96.6 (.91)	59.8 (3.09)
Novel	95.8 (1.32)	62.5 (3.79)
Block 4		
Repeated Match	98.3 (.87)	66.3 (3.19)
Repeated Mismatch	97.8 (.45)	65.3 (1.82)
Novel	97.2 (.96)	63.2 (3.01)
Block 5		
Repeated Match	97.6 (1.06)	61.8 (2.92)
Repeated Mismatch	97.5 (.62)	64.4 (1.39)
Novel	95.8 (1.23)	61.1 (2.39)
Block 6		
Repeated Match	98.3 (.87)	62.8 (3.33)
Repeated Mismatch	97.7 (1.01)	63.5 (1.88)
Novel	95.5 (1.42)	58.3 (3.21)
Block 7		
Repeated Match	99.3 (.48)	68.8 (2.94)
Repeated Mismatch	97.8 (.40)	65.2 (1.65)
Novel	96.2 (1.50)	63.9 (2.87)
Block 8		
Repeated Match	96.9 (1.11)	68.4 (2.79)
Repeated Mismatch	98.2 (.46)	64.6 (1.90)
Novel	93.8 (1.53)	65.3 (3.00)

locations and used this episodic knowledge to improve search efficiency dramatically.

Template-based guidance: Search time. To probe whether the match between the critical distractor color and the VWM color influenced overall search times, we first examined ST in the Repeated Match and Repeated Mismatch conditions, collapsing across blocks. ST was not reliably different on Match (1,957 ms) compared with Mismatch (1,973 ms) trials, $t(23) = 0.42$, $p = .681$, $p\eta^2 = .007$. To examine this effect across learning, we fit the ST data in the Repeated Match condition, comparing these parameter estimates to those in the Repeated Mismatch condition (see Figure 2). Interestingly, there was no effect of distractor color match on either the intercept parameter, $t(23) = 0.25$, $p = .801$, $p\eta^2 = .003$, or the learning-rate parameter, $t(23) = 0.53$, $p = .600$, $p\eta^2 = .012$. Thus, any tendency for attention to be guided to VWM-matching objects was not ultimately reflected in overall search time in this experiment. Note that a significant effect of memory match on overall search time was observed in Experiment 2.

Template-based guidance: Distractor fixation probability. Next, we examined the primary measure of template-based guidance: capture by a VWM-matching distractor (Bahle, Matsukura, et al., 2018). For the Match and Mismatch conditions, we calculated the probability of critical distractor fixation by block (see Figure 3). These probabilities were then arcsine transformed and analyzed with a 2 (Match type) × 8 (Block) ANOVA.

Participants were more likely to fixate the distractor on Match (.37) compared with Mismatch (.26) trials, $F(1, 23) = 30.128$, $p < .001$, $p\eta^2 = .567$. The distractor was also less likely to be fixated as the blocks progressed, $F(7, 161) = 20.412$, $p < .001$, $p\eta^2 = .470$ (Block 1: .48, Block 8: .28). Critically, there was no reliable interaction between these two factors, $F(7, 161) = 1.412$, $p = .204$, $p\eta^2 = .058$. To verify that the match effect remained robust in the later stages of the experiment, when episodic guidance was available, we computed the match effect in Block 1 (no episodic guidance), Block 8 (strong episodic guidance), and an average of Blocks 5–8 (i.e., the second half of the experiment, to provide greater power during the period of strong episodic guidance). The match effect was statistically reliable in all three cases (Dunn-Sidak corrected for eight comparisons): For Block 1, $t(23) = 3.68$, $p = .008$, $p\eta^2 = .370$; for Block 8, $t(23) = 3.26$, $p = .024$, $p\eta^2 = .316$; and for Blocks 5–8, $t(23) = 4.55$, $p < .001$, $p\eta^2 = .474$. These results indicate that template-based guidance was functional even in the presence of strong episodic guidance.

As a converging test of the potential influence of episodic guidance on template-based guidance, we examined the magnitude of the match effect on distractor fixation probability as a function of the distance between the critical distractor and the target *F*. If template-based guidance was restricted or biased to regions of the scene that coincided with the locus of episodic guidance, the match effect should have been larger when the critical distractor was relatively close to the target *F* than when it was relatively far from the target *F*. However, in an analysis of the data from Blocks 5–8, we found no relationship between the size of the match effect and the distance between the target and critical distractor, $r = .093$, $t(47) = 0.632$, $p = .531$, $p\eta^2 = .009$, suggesting that template guidance was not constrained by episodic guidance. Similarly, in Bahle, Matsukura, et al., 2018, we found that the size of the match effect was not influenced by the *semantic* plausibility of the distractor location as a location for the target.

Time course of guidance during a trial. Finally, we examined the evolution of episodic and template-based guidance during a trial, focusing on the how quickly each was applied to the search process. To probe template-based guidance, we calculated the probability of critical distractor fixation for each ordinal fixation in the Match and Mismatch conditions (Figure 4A). This was done separately for Block 1, Block 8, and Blocks 5–8. As is evident from the figure, VWM match had a substantial effect on distractor fixation probability at the earliest stages of search. Despite an overall reduction in capture from Block 1 to Block 8, the early capture effect was evident both before and after episodic learning, with a reliable difference in distractor fixation probability for the fixation following the very first saccade on the scene: Block 1, $t(23) = 4.62$, $p < .001$, $p\eta^2 = .482$; Block 8, $t(23) = 3.67$, $p = .001$, $p\eta^2 = .369$; Blocks 5–8, $t(23) = 6.52$, $p < .001$, $p\eta^2 = .649$.

To probe the time course of episodic guidance during a trial, we calculated, for each ordinal fixation, the probability of target fixation for Repeated Mismatch scenes and for Novel scenes (Figure 4B). In the first block, we expected no difference in target fixation probability, as there was no opportunity for learning. In contrast, there was a large difference in target fixation probability in Block 8. However, this guidance did not emerge until the third participant-controlled fixation on the scene. There was no reliable

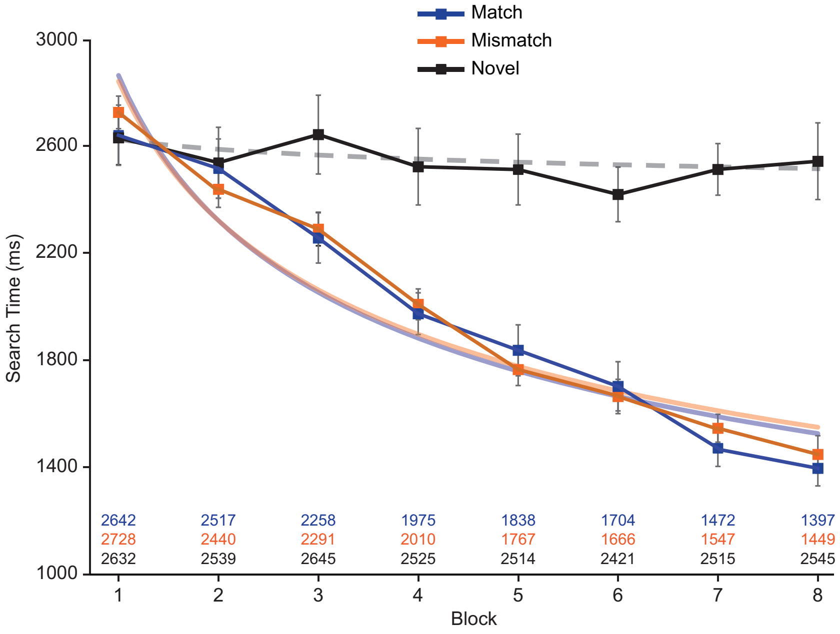

Figure 2. Search time as a function of scene repetition and VWM match. Data points indicate observed values (exact values presented numerically at the bottom of the figure). Error bars indicate within-subjects 95% confidence intervals (Morey, 2008). Faded lines indicate mean parameter values for the power function fits, plotted for the Repeated Match (blue [dark gray], solid lines, RMSE = 145.3), Repeated Mismatch (orange [light gray], solid lines, RMSE = 118.6), and Novel (black, dashed, RMSE = 53.8) conditions. Although the trend for these aggregate data may appear to conform better to a linear function than to a power function, the individual participant data were well fit by the power function (see online supplemental materials), and fitting the data to a linear function produced an equivalent pattern of results. See the online article for the color version of this figure.

difference between repeated and novel trials for Fixations 1 (p = .318) and 2 (p = .768). A reliable effect was observed only starting at Fixation 3, $t(23) = 2.98$, $p = .007$, $p\eta^2 = .279$. The same held when considering the combined data from Blocks 5–8: There was no reliable difference for Fixations 1 ($p = .088$) and 2 ($p = .648$); a reliable effect was observed only starting at Fixation 3, $t(23) = 4.69$, $p < .001$, $p\eta^2 = .488$. Thus, the application of episodic guidance was delayed relative to template-based guidance from VWM.

In sum, we observed robust episodic learning and contextual cuing. Yet, this scene-based guidance had minimal influence on template guidance, which was observed even when episodic guidance was strongest. In addition, template-based guidance was implemented earlier during search than episodic guidance at all stages of learning.

Experiment 2

A potential concern with the results of Experiment 1 is that search times may not have reached asymptote by the eighth repetition of the 48 experimental items (see Figure 2). Thus, although there was a substantial contextual cuing effect, there may have been opportunity for additional episodic learning that, ultimately, could have led to reduction or elimination of the VWM-match effect. This issue was addressed in Experiment 2.

Experiment 2 was identical to Experiment 1 in most respects, except we limited the scene items to 24 repeated scenes, the novel scenes were eliminated, and there were 16 (rather than eight) blocks of search. These changes were expected to improve episodic learning and bring search performance closer to floor by the end of the experiment, as there were fewer scenes to learn, shorter elapsed time between scene repetitions, and a greater number of repetitions.

Method

Participants. Twenty-four new participants (14 female) from the University of Iowa completed the experiment for course credit. All participants were between the ages of 18 and 30 and reported normal or correct-to-normal vision without contact lenses.

Stimuli and procedure. Twenty-four scenes from the set of 48 repeated scenes in Experiment 1 were randomly selected to be the scene items in Experiment 2 (scenes with an asterisk in the scene list in the Appendix were the scenes used in Experiment 2). Each participant searched through the same 24 scenes. The procedure was the same as in Experiment 1, except there were no novel scenes in each block, and there were 16 blocks of search. Each block contained 24 scenes, with 6 in the Match and 18 in the Mismatch condition. Each scene was presented once in each block.

Data analysis. Trials were eliminated for the same reasons as in Experiment 1 (a total of 9% of trials). Mean search accuracy was high (97.7%) and did not differ among conditions (Match: 97.9%. Mismatch: 97.6%), $t(23) = 0.802$, $p = .431$, $p\eta^2 = .027$. Mean memory accuracy was 66.1, and there was slightly better memory

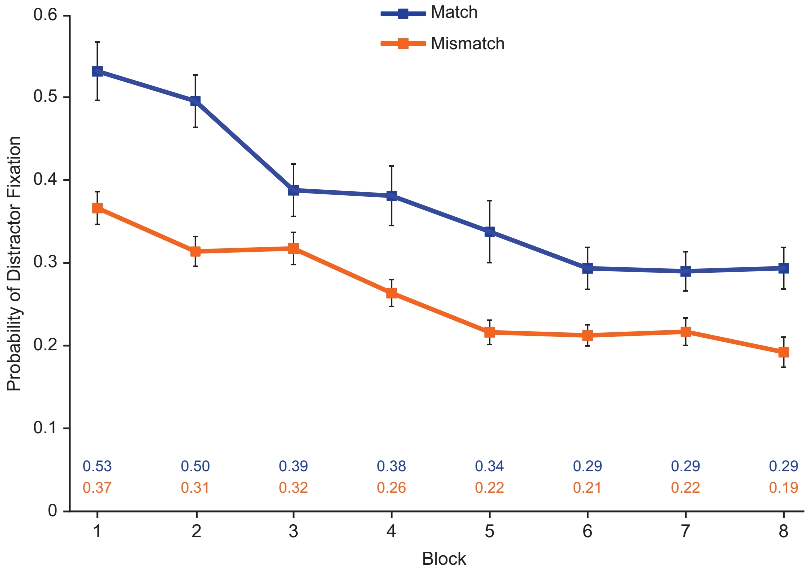

Figure 3. Probability of distractor fixation as a function of VWM match. Data points indicate observed values (exact values presented numerically at the bottom of the figure). Error bars indicate within-subjects 95% confidence intervals (Morey, 2008). See the online article for the color version of this figure.

performance on Match compared with Mismatch trials (Match: 69.4%. Mismatch: 65.1%), $t(23) = 4.05$, $p < .001$, $p\eta^2 = .417$. It is possible that the accuracy difference between match and mismatch trials reflected strategic orienting to matching distractors in order to improve memory performance. If this was the case, then there should have been a positive relationship between a participant's memory advantage and the size of their memory match effect on critical distractor fixation. There was no such relationship, $r = .149$, $t(23) = 0.707$, $p = .487$, $p\eta^2 = .022$. See Table 2 for complete search accuracy and memory accuracy results.

For analysis, the 16 search blocks were combined into eight, two-block epochs. This was done to both equate the number of observations for each condition type in Experiments 1 and 2 and to ensure there was adequate power to analyze individual epochs.

Results and Discussion

Episodic and template-based guidance: Search time. As is evident in Figure 5, search times decreased significantly for Repeated Mismatch searches over the course of the experiment. By the end of the experiment, mean ST was less than 1,000 ms and approximately 500 ms faster than ST for the corresponding condition at the end of Experiment 1. In addition, ST in Experiment 2 improved only modestly in the final two epochs (i.e., the final four blocks) of the experiment, with a mean decrease of 40 ms/repetition[3] compared with a 107 ms/repetition in Experiment 1. Thus, the modifications to the stimuli and design were successful in producing more robust episodic learning, with ST that approached asymptote by the end of the experiment.

As in Experiment 1, we investigated the extent to which ST was affected by distractor color match. Collapsing across epochs, ST was reliably faster on Mismatch (1,434 ms) compared with Match (1,606 ms) trials, $t(23) = 6.12$, $p < .001$, $p\eta^2 = .619$, consistent with the oculomotor capture results reported below. To characterize the data across epochs, we fit the data and compared parameter estimates for the Repeated Match and Repeated Mismatch conditions as in Experiment 1 (see Figure 5). There was no effect of distractor color match on the intercept parameter, $t(23) = 1.00$, $p = .329$, $p\eta^2 = .041$, but a reliable effect on the learning-rate parameter, $t(23) = 2.18$, $p = .040$, $p\eta^2 = .171$ ($s = 0.674$ for Mismatch; $s = 0.587$ for Match). However, these results appear to have been driven by performance in the first block of the experiment that was not characteristic of the pattern in the rest of the experiment (see Figure 5). Overall, the pattern indicates that capture by the matching distractor led to longer ST, as indicated by the ST data collapsed across epochs (above). Note that we did not observe a pattern of this type in Experiment 1. The source of the difference between experiments is not clear, except we note that oculomotor capture can often produce weak or inconsistent effects on end-of-trial measures, such as ST (Bahle, Beck, et al., 2018).

Template-based guidance: Distractor fixation probability. As in Experiment 1, we conducted an 8 (Epoch) × 2 (Match type) ANOVA on arcsine transformed probabilities of critical distractor fixation (see Figure 6). Participants were more likely to fixate the distractor on Match (.36) compared with Mismatch (.20) trials, $F(1, 23) = 47.123$, $p < .001$, $p\eta^2 = .672$. The distractor was also less likely to be fixated as the epochs progressed, $F(7, 161) = 50.363$, $p < .001$, $p\eta^2 = .686$ (Epoch 1: .48, Epoch 8: .28). Critically, there was no reliable interaction between these two factors, $F(7, 161) = 1.121$, $p = .352$, $p\eta^2 = .046$. However, to

---

[3] The mean decrease in the final two epochs was calculated by fitting a linear fit to the final two epochs of the experiment and extracting the slope parameter.

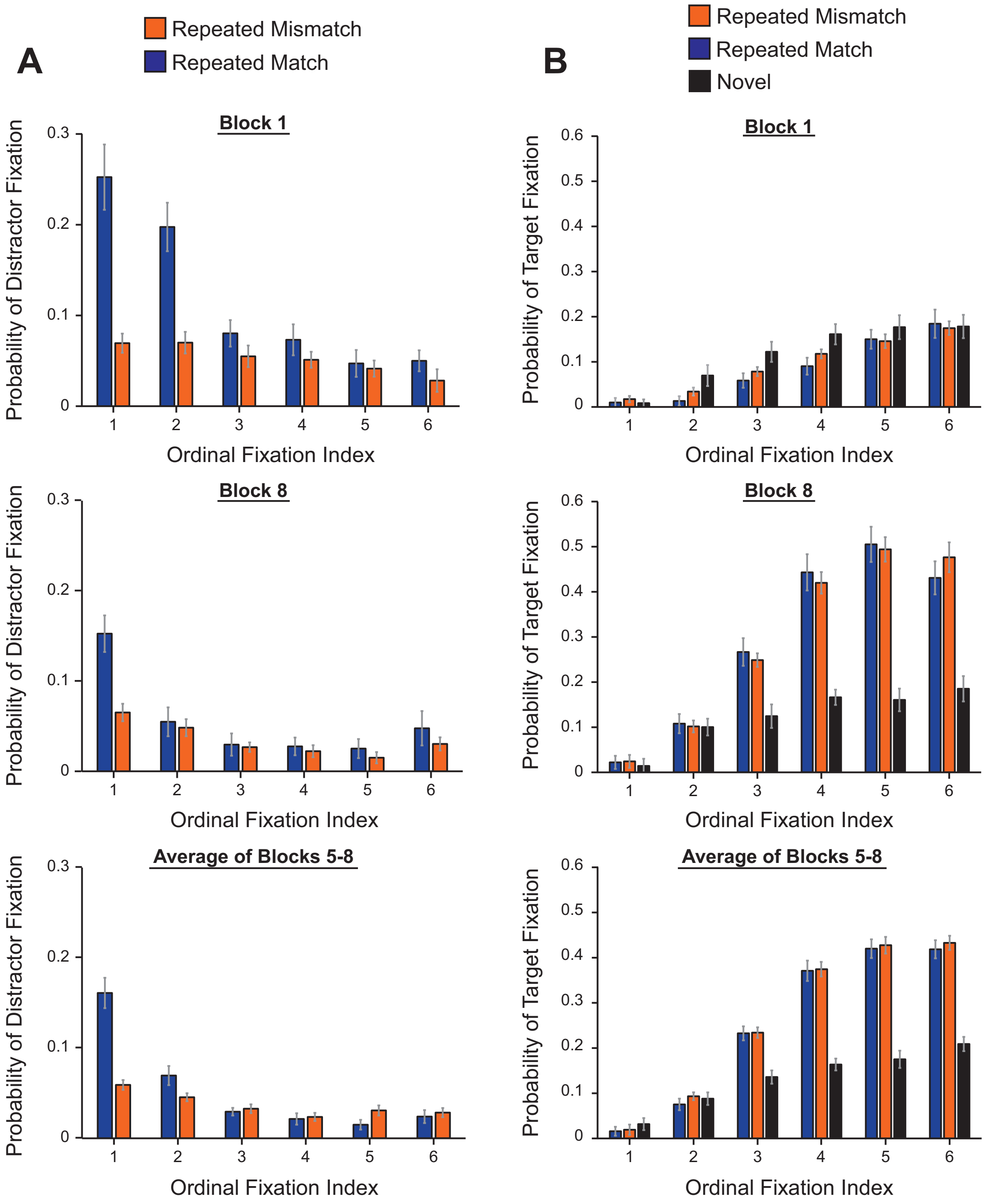

Figure 4. (A) Probability of distractor fixation as a function of ordinal fixation index for the Repeated Mismatch (orange [light gray]) and Repeated Match (blue [dark gray]) condition for Block 1 (top), Block 8 (middle), and the average of Blocks 5–8 (bottom). Note that Fixation 1 is the first participant-controlled fixation on the scene, following the first saccade. (B) Probability of target fixation as a function of ordinal fixation index for the Repeated Mismatch (orange [light gray]), Repeated Match (blue [dark gray]) and Novel (black) condition for Block 1 (top), Block 8 (middle), and the average of Blocks 5–8 (bottom). Note that the critical comparison is between the Repeated Mismatch and Novel conditions. Error bars indicate within-subjects 95% confidence intervals (Morey, 2008). See the online article for the color version of this figure.

Table 2
Search and Memory Accuracy (%) for Experiment 2 as a Function of Distractor Color Match Condition and Block

Condition	Search Accuracy (*SE*)	Memory Accuracy (*SE*)
Epoch 1		
Repeated Match	94.4 (1.39)	69.8 (3.21)
Repeated Mismatch	95.6 (.92)	63.9 (1.83)
Epoch 2		
Repeated Match	99.0 (.58)	67.4 (2.40)
Repeated Mismatch	97.8 (.53)	63.7 (1.59)
Epoch 3		
Repeated Match	99.3 (.48)	70.1 (2.60)
Repeated Mismatch	97.7 (.64)	65.1 (1.62)
Epoch 4		
Repeated Match	98.6 (.82)	69.1 (2.86)
Repeated Mismatch	98.7 (.41)	65.5 (2.33)
Epoch 5		
Repeated Match	98.6 (.65)	70.8 (3.09)
Repeated Mismatch	98.8 (.37)	65.4 (1.80)
Epoch 6		
Repeated Match	97.2 (1.19)	71.5 (2.83)
Repeated Mismatch	96.6 (1.47)	66.6 (1.82)
Epoch 7		
Repeated Match	97.6 (1.23)	67.0 (2.77)
Repeated Mismatch	97.8 (.98)	65.9 (1.68)
Epoch 8		
Repeated Match	98.6 (.65)	69.1 (2.72)
Repeated Mismatch	98.2 (.60)	64.7 (1.79)

again verify that the match effect was present in the later stages of the experiment when episodic guidance was available, we computed the match effect in Epoch 1, Epoch 8, and an average of Epochs 5–8. We found that the match effect was statistically reliable in all three cases (again, Dunn-Sidak corrected for eight comparisons): For Epoch 1, $t(23) = 7.68$, $p < .001$, $p\eta^2 = .720$; For Epoch 8, $t(23) = 2.94$, $p = .050$, $p\eta^2 = .274$; and For Epochs 5–8, $t(23) = 6.20$, $p < .001$, $p\eta^2 = .626$. Thus, the effect of template-guidance remained robust even in the presence of strong episodic guidance.

As in Experiment 1, we also considered whether the size of the match effect was modulated by the distance between the target *F* and the critical distractor in the last half of the experiment. As before, there was no relationship between these two factors, $r = .188$, $t(23) = 0.896$, $p = .380$, $p\eta^2 = .034$. This result, together with Experiment 1 and our previous work (Bahle, Matsukura, et al., 2018), suggests that template-based guidance is not restricted to regions of the scene likely to contain the target.

Time course of guidance during a trial. Finally, we examined the time course of template-based guidance during a trial, as in Experiment 1 (see Figure 7). This was done separately for Epoch 1 (no episodic guidance), Epoch 8 (strong episodic guidance), and Epochs 5–8 collapsed. As in Experiment 1, initial eye movements were preferentially directed to the critical distractor on match compared with mismatch trials in the earliest stages of search. Although the magnitude of capture decreased from Epoch 1 to Epoch 8, the early capture effect was evident both before and after episodic learning, with a reliable difference in distractor fixation probability for the fixation following the very first saccade on the scene: Epoch 1, $t(23) = 8.87$, $p < .001$, $p\eta^2 = .774$; Epoch 8, $t(23) = 2.92$, $p = .008$, $p\eta^2 = .271$; Epoch 5–8, $t(23) = 5.34$, $p < .001$, $p\eta^2 = .553$.

General Discussion

Well-learned scene-to-target-location associations provide a potentially limiting case of efficient, scene-based guidance of atten-

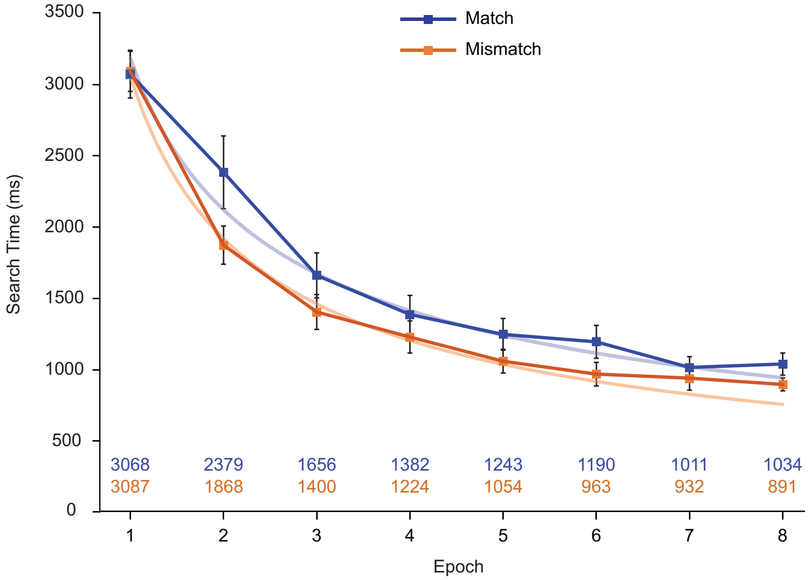

Figure 5. Search time as a function of distractor color match and epoch. Data points indicate observed values (exact values presented numerically at the bottom of the figure). Error bars indicate within-subjects 95% confidence intervals (Morey, 2008). Faded lines represent values for the power function fits, plotted for the Match (RMSE = 110.8) and Mismatch (RMSE = 72.0) conditions. See the online article for the color version of this figure.

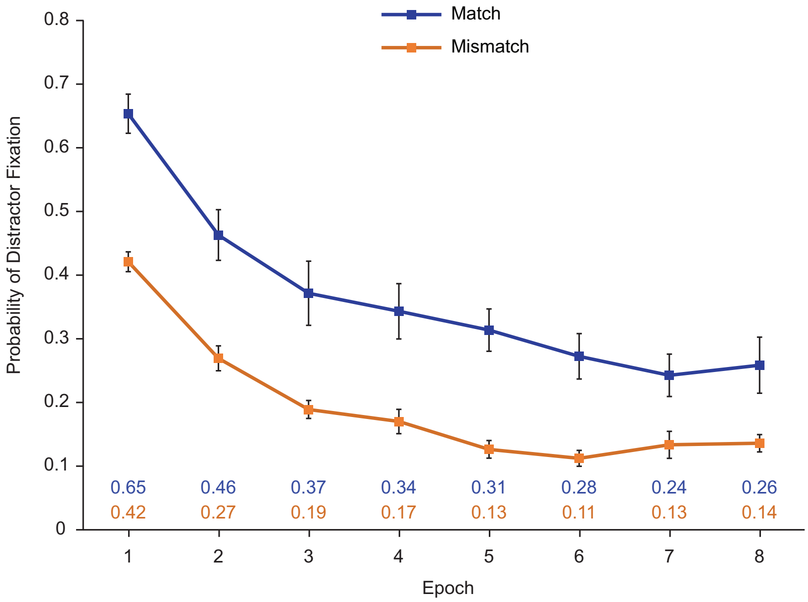

Figure 6. Probability of distractor fixation as a function of VWM match. Data points indicate observed values (exact values presented numerically at the bottom of the figure). Error bars indicate within-subjects 95% confidence intervals (Morey, 2008). See the online article for the color version of this figure.

tion and gaze (Brockmole & Henderson, 2006): Rapid scene recognition allows one to simply retrieve the episodically associated target location, providing maximal spatial constraint on search. Thus, the present results showing early, template-based guidance from VWM, even after substantial episodic learning, strongly suggest that search through scenes need not be characterized by a strictly ordered structure (scene-based guidance followed by template guidance) or a strictly hierarchical structure (template guidance applied only to plausible scene regions). With the results of Bahle, Matsukura, et al. (2018), this line of work challenges the core assumption of current models (Torralba et al., 2006; Wolfe et al., 2011) that scene-based knowledge is applied at the onset of search, constraining other forms of guidance. Template-based guidance was implemented earlier during the course of search than was episodic guidance (the present study) or semantic guidance (Bahle, Matsukura, et al., 2018), and template-based guidance was not constrained spatially by either form of scene-based guidance.

Specifically, the present results indicate a need for reconceptualization of how different sources of guidance are combined in a priority map of the current scene. For example, in the contextual guidance model of Torralba et al. (2006), saliency and global contextual factors of a scene are integrated before search commences, leading to a contextually modulated priority map that restricts even the very first eye movement during search to those regions of the scene that are consistent with the scene knowledge. Later work (Ehinger, Hidalgo-Sotelo, Torralba, & Oliva, 2009) expanded this model to include guidance from a target template representation, but this model also integrates all sources of information before search commences, such that early saccades will be driven primarily by scene knowledge. In contrast, the present results demonstrate that scene gist takes longer to incorporate into the search process, whereas initial guidance from the contents of VWM is available from the onset of search. Thus, future models of visual search through natural scenes must allow that different sources of guidance become available at different times during the search process, and not necessarily in a manner that prioritizes scene-based components.

Because of this difference in time-course for different sources of guidance, attentional priority maps should not be conceptualized as static. Static priority maps are the classic formulation in most models of attentional priority (Torralba et al., 2006; Wolfe, 1994). This is true both in models created to explain search in real-world scenes (e.g., Torralba et al., 2006) and in abstract arrays (e.g., Wolfe, 1994), with the only change to the topography of the map occurring after a location has been attended and rejected (typically implemented as location-specific inhibition of return). However, this conceptualization misses crucial aspects of the search process. As demonstrated in the present experiments and by Bahle, Matsukura, et al. (2018), scene-based guidance comes online only following the first one to two saccades on the scene. Thus, map topography will change as some regions are identified as plausible and others implausible for the target. Moreover, as gaze traverses the visual scene, the topography of the priority map will necessarily change as the eccentricity of various scene elements changes (Zelinsky, 2008). For example, features that may have been difficult to identify in the periphery due to poor resolution and crowding become easier to identify when gaze position changes and eccentricity is reduced. Such dynamics will alter low-level physical salience, the goodness of match to a template representation, and perhaps even scene-based guidance if local information is necessary to apply scene-level knowledge.

In contrast with the present results, previous work has found that scene-based guidance was implemented from the very first saccade on a scene and tended to dominate other forms of guidance (Torralba et al., 2006). However, this finding may have been caused by the fact that the search task was the same on every trial

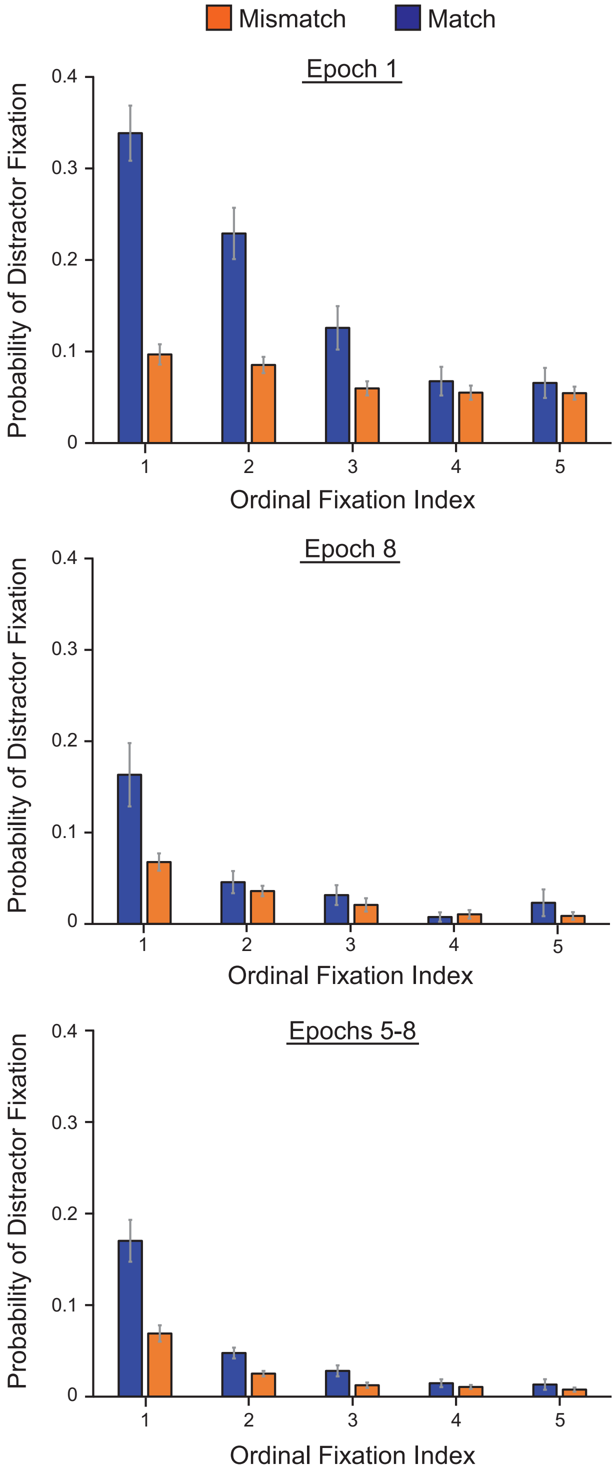

Figure 7. Probability of distractor fixation as a function of ordinal fixation index for the Mismatch (orange [light gray]) and Match (blue [black]) condition for Epoch 1 (top), Epoch 8 (middle), and the average of Epochs 5–8 (bottom). Note that Fixation 1 is the first participant-controlled fixation on the scene, following the first saccade. See the online article for the color version of this figure.

(e.g., find pedestrians), and the scene images were always drawn from a single category (e.g., street scenes). Such a design would have allowed participants to apply the same scene-based knowledge throughout the entire experiment (i.e., that pedestrians tend to be found near roads and in the lower portions of the images). In the present method, the constraint from episodic knowledge changed on a trial-by-trial basis, as target locations were chosen randomly for each scene. And, in Bahle, Matsukura, et al. (2018), each search involved a different scene type and a different target object, requiring the participant to apply different semantic information for every search. We believe that our experimental conditions are more characteristic of typical real-world search behavior that those of Torralba et al. (2006). In real-world searches, target objects change frequently, requiring one to modify the application of scene-based knowledge. Only in relatively rare, repetitive tasks (e.g., industrial quality control) do the target and scene type remain constant. However, we acknowledge that under conditions where the same scene-based constraints can be applied consistently, one is likely to find earlier scene-based guidance, and this may place constraints on the application of template-based guidance.

Finally, although we observed consistent differences in distractor fixation probability as function of VWM match in both experiments, only in Experiment 2 did this capture effect ultimately manifest in end-of-trial ST differences. One possible cause for the lack of an effect in Experiment 1 is that the first saccade on the scene was strongly influenced by VWM match but was not influenced by episodic guidance to the target; early saccades that were directed to the distractor on match trials were not being directed to the target on mismatch trials. As a result, capture produced minimal search cost. In contrast, in Experiment 2, where search performance was much more efficient because of greater episodic learning, some initial saccades that may have been directed to the target on mismatch trials were directed to the distractor on match trials. In any case, the discrepancy in ST between match and mismatch trials in these two experiments highlights the importance of using techniques, such as eye tracking, that provide a direct measure of capture and do not depend on inference from end-of-trial measures of search (see also Bahle, Beck, et al., 2018; Beck et al., 2018).

References

Bahle, B., Beck, V. M., & Hollingworth, A. (2018). The architecture of interaction between visual working memory and visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 44,* 992–1011. http://dx.doi.org/10.1037/xhp0000509

Bahle, B., Matsukura, M., & Hollingworth, A. (2018). Contrasting gist-based and template-based guidance during real-world visual search. *Journal of Experimental Psychology: Human Perception and Performance, 44,* 367–386. http://dx.doi.org/10.1037/xhp0000468

Beck, V. M., Luck, S. J., & Hollingworth, A. (2018). Whatever you do, don't look at the . . .: Evaluating guidance by an exclusionary attentional template. *Journal of Experimental Psychology: Human Perception and Performance, 44,* 645–662. http://dx.doi.org/10.1037/xhp0000485

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 699–706. http://dx.doi.org/10.1037/0278-7393.32.4.699

Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition, 13,* 99–108. http://dx.doi.org/10.1080/13506280500165188

Brooks, D. I., Rasmussen, I. P., & Hollingworth, A. (2010). The nesting of search contexts within natural scenes: Evidence from contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1406–1418. http://dx.doi.org/10.1037/a0019257

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

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71. http://dx.doi.org/10.1006/cogp.1998.0681

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 224–234. http://dx.doi.org/10.1037/0278-7393.29.2.224

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

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

Ehinger, K. A., Hidalgo-Sotelo, B., Torralba, A., & Oliva, A. (2009). Modeling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition, 17,* 945–978. http://dx.doi.org/10.1080/13506280902834720

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

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford, UK: Elsevier. http://dx.doi.org/10.1016/B978-008044980-7/50027-6

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 883–914. http://dx.doi.org/10.1037/0278-7393.18.5.883

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision, 10,* 4. http://dx.doi.org/10.1167/10.2.4

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4,* 61–64.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS.* New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-1-4419-0318-1

Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes? *Visual Cognition, 12,* 852–877. http://dx.doi.org/10.1080/13506280444000553

Rutishauser, U., & Koch, C. (2007). Probabilistic modeling of eye movement data during conjunction search via feature-based attention. *Journal of Vision, 7,* 5. http://dx.doi.org/10.1167/7.6.5

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

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

Williams, L. G. (1967). The effects of target specification on objects fixated during visual search. *Acta Psychologica, 27,* 355–360. http://dx.doi.org/10.1016/0001-6918(67)90080-7

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., 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

Woodman, G. F., Carlisle, N. B., & Reinhart, R. M. G. (2013). Where do we store the memory representations that guide attention? *Journal of Vision, 13,* 1. http://dx.doi.org/10.1167/13.3.1

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115,* 787–835. http://dx.doi.org/10.1037/a0013118

Appendix

Scene List

Experimental scenes		Filler scenes
Scene category	Critical distractor	Scene category
Bathroom 1	Lotion	Bathroom 1
Bathroom 2	Towel	Bathroom 2
Bedroom 1	Bag	Bathroom 3
Bedroom 2*	Chair	Bathroom 4
Bedroom 3	Lamp	Bedroom 1
Bedroom 4*	Ottoman	Bedroom 2
Bedroom 5*	Lamp	Bedroom 3
Den 1*	Cushion	Bedroom 4
Den 2*	Vase	Bedroom 5

(*Appendix continues*)

Appendix (*continued*)

Experimental scenes		Filler scenes
Scene category	Critical distractor	Scene category
Den 3*	Cushion	Bedroom 6
Den 4	Cushion	Bedroom 7
Den 5	Garbage Bin	Bedroom 8
Den 6	Vase	Bedroom 9
Kitchen 1*	Strainer	Bedroom 10
Kitchen 2	Orange	Bedroom 11
Kitchen 3*	Bowl	Bedroom 12
Kitchen 4*	Cooler	Bedroom 13
Kitchen 5*	Towel	Bedroom 14
Kitchen 6	Towel	Bedroom 15
Kitchen 7*	Towel	Bedroom 16
Kitchen 8*	Filter	Bedroom 17
Kitchen 9	Skillet	Bedroom 18
Kitchen 10	Placemat	Den 1
Kitchen 11	Tomato	Den 2
Kitchen 12	Plate	Den 3
Kitchen 13	Pot	Den 4
Kitchen 14*	Towel	Den 5
Kitchen 15*	Teapot	Den 6
Office 1	Candle	Den 7
Office 2*	Folder	Den 8
Office 3	Folder	Den 9
Office 4	Lamp	Den 10
Office 5	File Cabinet	Den 11
Office 6	Folder	Den 12
Office 7*	Book	Den 13
Office 8	Folder	Garage 1
Office 9*	Chair	Kitchen 1
Office 10*	Cushion	Kitchen 2
Office 11*	Recyling Bin	Kitchen 3
Office 12*	Paper	Kitchen 4
Office 13	Folder	Kitchen 5
Office 14*	Soda Can	Kitchen 6
Office 15	Tape Dispenser	Kitchen 7
Office 16	Folder	Kitchen 8
Office 17	Paper	Kitchen 9
Porch 1*	Sandbox	Kitchen 10
Porch 2*	Pumpkin	Kitchen 11
Porch 3*	Sign	Kitchen 12
		Kitchen 13
		Kitchen 14
		Kitchen 15
		Kitchen 16
		Kitchen 17
		Kitchen 18
		Kitchen 19
		Kitchen 20
		Kitchen 21
		Kitchen 22
		Kitchen 23
		Kitchen 24
		Kitchen 25
		Kitchen 26
		Kitchen 27
		Kitchen 28
		Kitchen 29
		Kitchen 30
		Office 1
		Office 2
		Office 3

(*Appendix continues*)

Appendix (*continued*)

Experimental scenes		Filler scenes
Scene category	Critical distractor	Scene category
		Office 4
		Office 5
		Office 6
		Office 7
		Office 8
		Office 9
		Office 10
		Office 11
		Office 12
		Office 13
		Office 14
		Office 15
		Office 16
		Office 17
		Office 18
		Office 19
		Office 20
		Office 21
		Office 22
		Office 23
		Office 24
		Office 25
		Office 26
		Office 27
		Porch 1
		Porch 2
		Porch 3

Note. Scenes with an asterisk were used in both Experiment 1 and 2.

Received May 24, 2018
Revision received October 25, 2018
Accepted December 5, 2018 ■