

COMMENTARY

Task Specificity and the Influence of Memory on Visual Search: Comment on Vö and Wolfe (2012)

Andrew Hollingworth
University of Iowa

Recent results from Vö and Wolfe (2012b) suggest that the application of memory to visual search may be task specific: Previous experience searching for an object facilitated later search for that object, but object information acquired during a different task did not appear to transfer to search. The latter inference depended on evidence that a preview task did not improve later search, but Vö and Wolfe used a relatively insensitive, between-subjects design. Here, we replicated the Vö and Wolfe study using a within-subject manipulation of scene preview. A preview session (focused either on object location memory or on the assessment of object semantics) reliably facilitated later search. In addition, information acquired from distractors in a scene-facilitated search when the distractor later became the target. Instead of being strongly constrained by task, visual memory is applied flexibly to guide attention and gaze during visual search.

Keywords: visual search, scene perception, eye movements, visual memory, episodic memory

What factors control visual search efficiency in natural scenes? Factors that control efficiency in traditional search paradigms influence search through scenes, such as target-distractor similarity (Pomplun, 2006) and set size (Neider & Zelinsky, 2008). However, the strongest factors appear to be knowledge and memory (for reviews, see Hollingworth, 2012; Wolfe, Vö, Evans, & Greene, 2011). Semantic knowledge of the typical locations of objects in scenes allows participants to direct attention to plausible scene regions (Henderson, Weeks, & Hollingworth, 1999; Neider & Zelinsky, 2006; Torralba, Oliva, Castelano, & Henderson, 2006). Memory for scene exemplars also provides strong guidance: A scene preview establishes a memory representation that facilitates subsequent search (Castelano & Henderson, 2007; Hollingworth, 2009), and repeated searches lead to rapid savings (Brockmole, Castelano, & Henderson, 2006; Brockmole & Henderson, 2006).

In a recent article, Vö and Wolfe (2012b) investigated the application of memory to visual search. They asked whether memory acquired from an object when it is not the target of search is used to guide attention and gaze when that object subsequently becomes a target. There were two main tests of memory transfer. First, Vö and Wolfe examined whether a preview session, in which participants saw the relevant scenes but did not engage in search, would facilitate later search for objects in those scenes. Second,

they examined the effect of searching for multiple objects in the same scene to determine whether information acquired from an object when it was a distractor would facilitate search when it later became a target.

Each experiment began with a preview session (or no preview session). The preview instructions focused on spatial memory or semantic content. Participants did not know that they would later search through the scenes. In a surprise search task, participants searched sequentially for 15 objects within each of the 10 scene items, with the scene remaining visible throughout these multiple searches. Participants were instructed to direct their gaze to the target object as quickly as possible and to press a button upon doing so. The primary measure was elapsed time from the removal of the label to the first eye fixation on the target object. After searching through all 10 scenes, the search task was repeated in two additional blocks.

Vö and Wolfe (2012b) reported three findings. First, and most surprising, the preview session had no effect on search efficiency. Search in an experiment with no preview was just as rapid as search in experiments that included a preview, and there was no effect of preview instructions. Second, search times decreased substantially as participants repeated their search for a particular target object in Blocks 2 and 3, replicating contextual cuing studies. Finally, Vö and Wolfe observed trends (some statistically reliable, some not) indicating that search times decreased over the course of the 15 sequential searches within a scene. This effect is consistent with several studies showing that distractor memory facilitates search (Howard, Pharaon, Körner, Smith, & Gilchrist, 2011; Körner & Gilchrist, 2007). However, Vö and Wolfe argued that the effect of multiple searches within a scene was considerably smaller than the effect of target repetition across blocks and thus that the influence of distractor memory was not substantial.

Andrew Hollingworth, Department of Psychology, University of Iowa.
This research was supported by National Institutes of Health Grant R01 EY017356.

Correspondence concerning this article should be addressed to Andrew Hollingworth, Department of Psychology, University of Iowa, 11 Seashore Hall E, IA City, IA 52242-1407. E-mail: andrew-hollingworth@uiowa.edu

On the basis of the absence of a preview effect and the relatively small effect of multiple searches, *Võ and Wolfe (2012b)* concluded that memory representations acquired from nontarget objects had failed to facilitate visual search. Accepting this interpretation for the moment, there are two possible explanations for the absence of transfer. First, the memory representations functional in guiding visual search may be task specific, with the influence of memory limited to information acquired from earlier searches for that particular object. Such a view would be broadly consistent with theoretical approaches holding that visual memory encoding and application are strongly governed by task (*Ballard, Hayhoe, & Pelz, 1995; Droll, Hayhoe, Triesch, & Sullivan, 2005*). Alternatively, memory representations formed from nontargets may have the potential to facilitate search, but this effect was overshadowed in the *Võ and Wolfe* study by other factors. *Võ and Wolfe* raised the possibility that efficient guidance from general knowledge of typical object locations may have dominated guidance from distractor memory and from memory encoded during the preview.

It is not clear, however, that the *Võ and Wolfe* data are sufficient to accept the claim that distractor and preview memory failed to facilitate search in their study. Although the effect of distractor memory during multiple searches was relatively small, it was present nevertheless, and the magnitude of the effect is similar to other learning effects, such as contextual cuing (*Chun & Jiang, 1998*). Moreover, *Võ and Wolfe (2012a)* have subsequently reported a robust effect of multiple searches. Thus, distractor memory does facilitate search when the distractor becomes a target, even when search could be guided by general knowledge.

This still leaves the absence of a preview effect. In the *Võ and Wolfe (2012b)* preview session, each of the scenes was viewed for 30 s. This is easily sufficient to have encoded into memory the identities, visual details, and locations of many of the objects in each scene (*Brady, Konkle, Alvarez, & Oliva, 2008; Hollingworth, 2004, 2005; for a review, see Hollingworth, 2006*). It does not seem likely that such memory would be completely overshadowed by guidance from general knowledge of typical object locations. Thus, if the absence of a preview effect were robust and replicable, it would provide clear evidence that the application of memory to visual search is task specific and that memory representations formed during nonsearch tasks fail to transfer reliably to visual search.

However, there are three considerations that reduce confidence in the null effect of preview observed by *Võ and Wolfe (2012b)*. First, the result contrasts with other studies in which scene previews have facilitated search. *Hollingworth (2009)* implemented a preview manipulation very similar to that of *Võ and Wolfe* and found a robust preview advantage in search. However, participants in the *Hollingworth* study knew during each preview that they would later search through the scene, and this may have influenced the nature of the memory representations formed.

Second, *Võ and Wolfe (2012b)* may not have had enough power to detect a preview effect. They reported sufficient power to detect an effect similar in magnitude to the effect of repeated search across blocks (~350 ms), but preview effects in earlier studies have been substantially less than 350 ms. In addition, the effect of repeated search across blocks in *Võ and Wolfe* was driven both by memory for target location and by the opportunity to associate the visual properties of a target object with its corresponding label (*Wolfe, Alvarez, Rosenholtz, Kuzmova, & Sherman, 2011*). The latter advantage

could not contribute to a preview effect, because objects were not associated with labels in the preview. Thus, the tests of power in *Võ and Wolfe* were not necessarily appropriate for the magnitude of the effect plausibly generated by a preview.

Finally, preview manipulations in *Võ and Wolfe (2012b)* were implemented across experiments and may have suffered from group differences in baseline search performance. *Võ and Wolfe* compared elapsed time to target fixation in an experiment with no preview (Experiment 1) the two experiments that included a preview (Experiments 3 and 4). Not only was there no preview benefit, there was a substantial trend toward a preview cost of approximately 120 ms in each case. It is not clear why a scene preview should produce a cost in visual search. However, a comparison of Experiments 1 and 5 in *Võ and Wolfe* indicates that group differences in baseline search performance may have limited these cross-experiment analyses. In the first block of search in Experiment 5, there were two conditions, a cued condition (in which a spatial cue indicated the target object) and an uncued condition. Searches in the uncued condition were identical to those in Experiment 1; neither was preceded by a preview session. Elapsed time to target fixation was more than 300 ms shorter for the Experiment 1 group than for the Experiment 5 group, and manual response time was nearly 400 ms shorter.¹

The potential for group differences in baseline search performance and the relatively low power of between-subjects designs highlight the need for within-subject manipulations in this type of paradigm. To resolve the discrepancy between *Võ and Wolfe (2012b)* and earlier demonstrations of a preview advantage (*Hollingworth, 2009*), we replicated their study using a within-subjects manipulation of scene preview. In addition, we tested whether transfer is modulated by the overlap in informational demands between preview and search tasks. The preview task either focused on object location memory (strongly related to visual search) or on the semantic properties of the objects (not strongly related to visual search).

The method, illustrated in *Figure 1*, followed closely the basic method of *Võ and Wolfe (2012b)*. There were 12 scene items. Twelve objects were chosen as search targets in each scene. In a preview session, participants viewed 6 of the 12 scenes for 20 s each. One group of participants received preview instructions that focused on object location memory, and a second group received instructions that focused on assessing the semantic relationship between objects and scenes. The preview session was followed by two blocks of visual search. In each block, participants viewed each of the 12 scenes (half previewed, half novel) and searched sequentially for each of the 12 target objects.

¹ *Võ and Wolfe (2012b)* speculated that because search trials in Experiment 5 were intermixed with trials on which the target was cued directly, participants may have delayed the initiation of search as they waited to see whether a cue would appear. However, there were large differences between Experiments 1 and the uncued trials of Experiment 5 on measures of search efficiency that should not have depended on the time taken to initiate the search, such as the elapsed number of fixations to the target, the path ratio (ratio of the eye movement scanpath to a direct path), and decision time (the time taken from fixation of the target to manual response). On all measures of performance, participants in Experiment 1 were substantially faster and more efficient than those in Experiment 5.

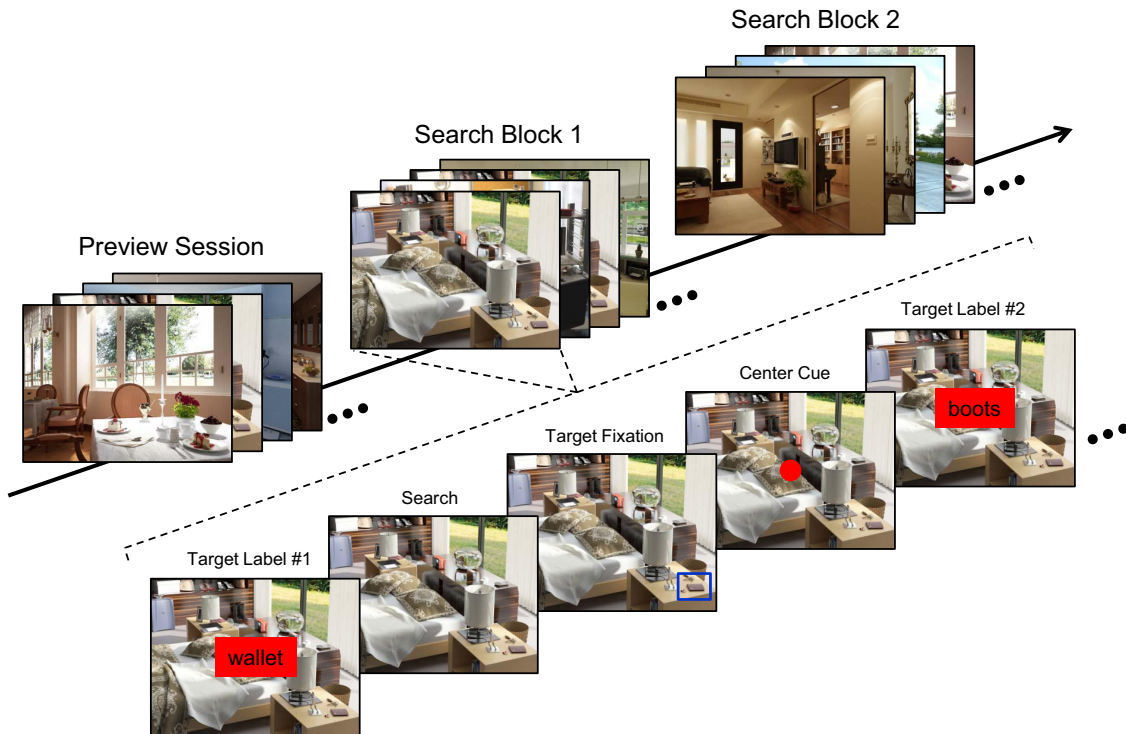


Figure 1. Illustration of the structure of the experiment. There were three sessions: a preview session and two blocks of search. Within each block of search, participants looked for 12 different objects sequentially within each of 12 scenes.

Method

Participants

Forty-eight participants (18–30 years of age) from the University of Iowa community completed the experiment for course credit or for pay. All reported normal or corrected-to-normal vision.

Stimuli

Twelve scenes were created in 3D Studio Max software and were rendered with the V-Ray engine. Each of the 12 target objects was plausibly found in its scene and appeared at a plausible location. Each object was the only exemplar of a particular object category (e.g., the stapler was the only stapler in the office). All 144 of the target object types were unique. Scene images subtended $26.0^\circ \times 19.5^\circ$ at a resolution of 1024×768 pixels. Target objects subtended $1.80^\circ \times 1.83^\circ$, on average. Target labels were presented in black text against a rectangular red background slightly larger than the word. A label of average length subtended $2.2^\circ \times 0.3^\circ$.

Apparatus

Stimuli were displayed on a 17-inch CRT monitor (75-Hz refresh) at a distance of 70 cm. An Eyelink 1000 eyetracker sampled eye position monocularly at 1000 Hz, with the head stabilized by a chin and forehead rest. Manual responses were

made on a serial button box. The experiment was controlled by E-prime software.

Procedure

Participants completed a preview session in which they viewed eight scenes for 20 s each. The first and last scenes were filler items. One group of 24 participants was instructed to remember each scene and the locations of the objects to prepare for a memory test (not administered). The other group of 24 was instructed to decide which object in the scene was least likely to appear in a scene of that type. The latter instructions forced participants to evaluate the semantic properties of every object in the scene, but the task made no demands on remembering the objects or their locations.

After the preview session, there were two blocks of search. In the first, participants completed 14 trials. Each trial consisted of multiple searches through a scene item. On the first two trials, participants searched for six objects in each of two filler items to acclimate them to the multiple search procedure. One item was old (a filler item from the preview session), and one was new. Next, they searched for each of the 12 objects in each of the 12 experimental scenes.

The sequence of events in a trial is illustrated in Figure 1. The scene was presented initially for 1000 ms. Then, the first target label appeared at the center for 500 ms. Participants were free to start searching for the object as soon as they had read the label. They were instructed to press the response button immediately upon target fixation. Upon response, a blue rectangle appeared

around the target object for 200 ms, providing feedback. Next, a red dot (0.3° diameter) appeared at scene center for 1000 ms, cuing return of gaze to the center. This was followed by the next target label, and the cycle repeated until all 12 objects had been found.

After completing the first block, there was a second block of search though the same 12 experimental scenes for the same target objects.

The scene items appearing in the preview session were counterbalanced across participants. The order of experimental scenes in the preview and search blocks was determined randomly. Within each trial, the order of the 12 target objects was determined randomly.

Data Analysis

Eye movement measures provided the principal data. A rectangular region was defined around each target object, 0.82° larger on each dimension than the object itself, with a minimum region size of 1.37° × 1.37°. Saccades were classified using a velocity criterion (>30° per second). The principal eye movement measure was the elapsed time from label onset to the first fixation in the target region for the entry that preceded the manual response. In a few cases, the eyes entered the target region, exited the target region, and the returned to the target region before the response; the elapsed time measure reflected the second entry. Note that our measures of search were timed from label onset—rather than offset, as in Vö and Wolfe (2012b)—accounting for absolute differences in the reported values. The elapsed time to target fixation measure included the time required to read the label. To ensure that effects of multiple searches were not caused by changes in the efficiency of reading the label, we also calculated the elapsed time measure from the onset of the first saccade that took the eyes away from the center. The two approaches produced the same numerical pattern of results and precisely the same pattern of statistical significance.

Two additional measures are reported. The first is *elapsed number of fixations* from the fixation that left the center of the screen to the first fixation in the target region for the entry that preceded the response. The measure included the first fixation following the saccade that took the eyes away from the center, the first fixation within the target region, and any fixations occurring between these two events. The minimum value was 1 (i.e., the first saccade that left the center landed in the target region). The second additional measure was *reaction time (RT)* from label onset to manual response.

Searches were eliminated from the analysis if the eyes did not enter the target region during the 1000 ms preceding the response (14.7% of searches) or if response time was <500 ms or >10 s (an additional 0.4% of the data). This did not alter the pattern of RT results. For the elapsed number of fixations measure, searches were also eliminated if the participant did not fixate the center of the screen prior to the initiation of the search (3.8% of the remaining data).

Results

The 12 searches within a trial were grouped into four epochs. The main results are displayed in Figure 2, and numerical values are reported in Table 1. Analyses are reported for the elapsed time

until the first fixation on the target region. Analyses over elapsed number of fixations and manual RT produced the same pattern of data.

Repeated-Search Effect

There was a reliable effect of block, indicating that the second searches for the objects were more efficient than the first searches, $F(1, 46) = 480, p < .001, \eta_p^2 = .91$. This replicates repeated search effects in Vö and Wolfe (2012b).

Multiple-Searches Effect

Second, there was an effect of multiple searches within a scene. In Block 1, the epoch variable produced a reliable linear effect, $F(1, 46) = 6.98, p = .01, \eta_p^2 = .13$, indicating that search times decreased as searches progressed through the set of 12. This replicates the effects of multiple searches reported by Vö & Wolfe (2012a, 2012b). In Block 2, there was no linear effect of epoch, $F(1, 46) = 1.56, p = .22, \eta_p^2 = .03$, but this is unsurprising, because search times were close to floor in Block 2.

To calculate the numerical magnitude of the multiple search effect in Block 1, we fit a regression line to the elapsed time to fixation means for Searches 1–12 and then compared points corresponding to the first and last searches. The difference was 87 ms, suggesting that the last search through a particular scene was approximately 87 ms faster than the first search. We conducted the same analysis over the equivalent data from Experiments 1, 3, and 4 of Vö and Wolfe (2012b).² The difference was 91 ms over the course of 15 searches. Thus, the two studies produced effects of similar numerical magnitude.

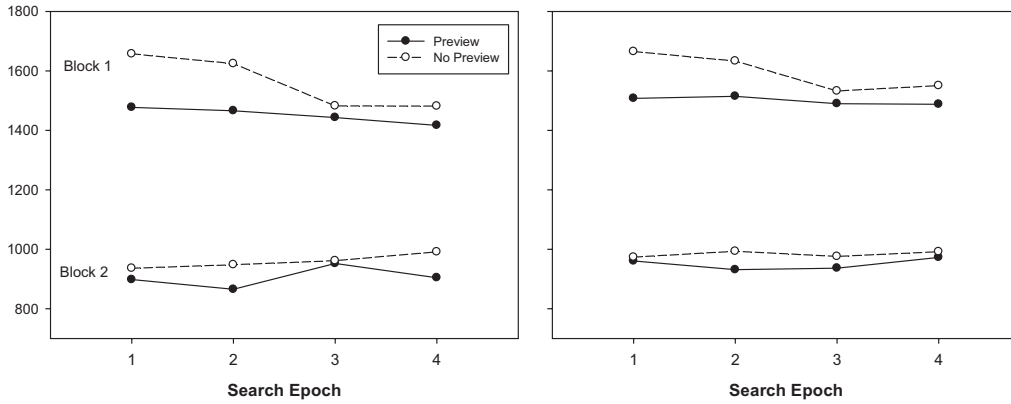
Vö and Wolfe (2012b) argued that the effect of multiple searches (~90 ms) was not substantial in relation to the effect of search repetition across blocks (~350 ms). However, this comparison is limited in two ways. As discussed in the introduction, the effect across blocks was generated not only by memory for the locations of objects but also by the ability to associate the visual properties of an object with its corresponding target label. Wolfe, Alvarez, et al. (2011) found that the latter type of learning accounts for approximately half of the repeated search effect in this type of paradigm. Such learning was not possible during multiple searches within a scene, because distractors were not associated with labels. The second issue is that search efficiency in the second block will be influenced by learning that occurred both when the object was a target and when it was a distractor. Because the effect across blocks includes the facilitative influence of distractor memory, it is hardly surprising that it is larger than the effect of distractor memory alone. In summary, the contribution of multiple sources of learning to the effect across blocks (including the influence of distractor memory itself) makes it very difficult to draw any clear comparison between the magnitudes of the two effects. Thus, there is no compelling reason to dismiss the effect of multiple searches as unsubstantial, especially because the effect has now been ob-

² These are the three experiments that implemented the conditions most similar to ours: Experiment 1 (no preview), Experiment 3 (“semantic assessment” preview), and Experiment 4 (“memorize objects and locations” preview). We thank Melissa Vö for providing access to the data from their study.

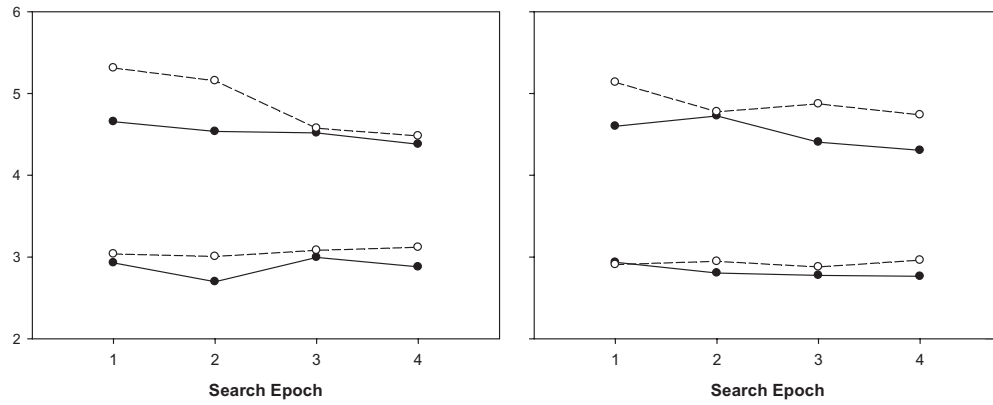
Related Preview Task
(memorize object locations)

Unrelated Preview Task
(semantic assessment)

A. Elapsed Time to First Fixation on Target (ms)



B. Elapsed Fixations to First Fixation on Target



C. Reaction Time (ms)

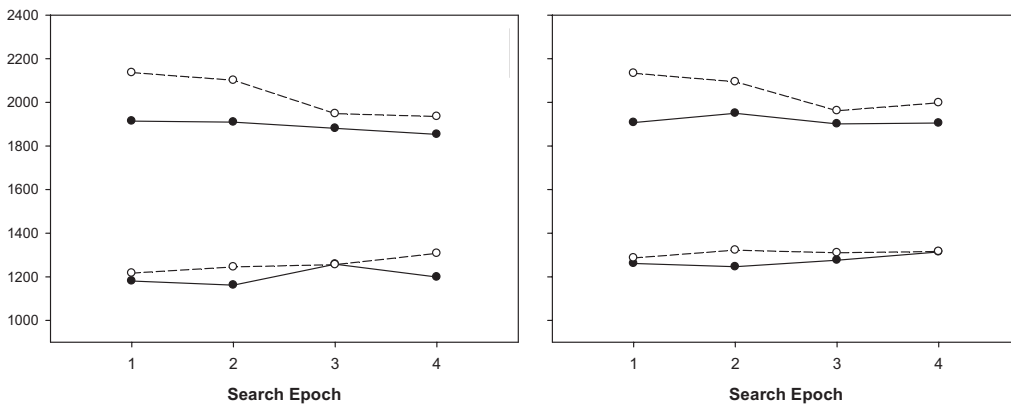


Figure 2. Visual search data as a function of preview task, preview availability, search block, and search epoch. Three measures of search efficiency are reported: (A) elapsed time from the onset of the search label to the first fixation on the target object immediately preceding the response, (B) elapsed number of fixations from the first saccade that left the screen center to the first fixation on the target object immediately preceding the response, and (C) reaction time from the onset of the search label to the manual button response.

Table 1
Mean Values (Standard Errors of the Means) for the Three Measures of Search Efficiency

Preview instructions	Block 1				Block 2			
	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 1	Epoch 2	Epoch 3	Epoch 4
Memorize objects and locations								
Elapsed time to target fixation (ms)								
Preview	1477 (73)	1466 (69)	1443 (80)	1417 (85)	898 (32)	866 (36)	952 (41)	904 (35)
No preview	1657 (84)	1624 (74)	1482 (68)	1481 (63)	936 (37)	948 (32)	962 (43)	992 (49)
Elapsed fixations to target fixation								
Preview	4.66 (.27)	4.54 (.22)	4.52 (.25)	4.38 (.21)	2.93 (.12)	2.70 (.11)	3.00 (.13)	2.88 (.14)
No preview	5.31 (.23)	5.15 (.24)	4.57 (.24)	4.48 (.16)	3.04 (.15)	3.01 (.12)	3.08 (.12)	3.12 (.17)
Reaction time (ms)								
Preview	1914 (87)	1909 (87)	1881 (90)	1853 (102)	1181 (48)	1162 (49)	1258 (54)	1199 (45)
No preview	2136 (97)	2101 (87)	1948 (80)	1935 (74)	1217 (43)	1245 (40)	1256 (58)	1308 (62)
Semantic assessment								
Elapsed time to target fixation (ms)								
Preview	1507 (108)	1515 (94)	1490 (95)	1488 (85)	961 (42)	931 (48)	937 (39)	973 (49)
No preview	1665 (81)	1633 (85)	1532 (85)	1550 (91)	974 (33)	993 (61)	976 (66)	992 (64)
Elapsed fixations to target fixation								
Preview	4.60 (.33)	4.73 (.30)	4.40 (.26)	4.30 (.24)	2.94 (.13)	2.80 (.12)	2.78 (.11)	2.76 (.13)
No preview	5.14 (.20)	4.78 (.22)	4.87 (.27)	4.74 (.28)	2.91 (.12)	2.95 (.12)	2.88 (.18)	2.96 (.18)
Reaction time (ms)								
Preview	1907 (127)	1950 (116)	1901 (117)	1905 (114)	1261 (57)	1246 (69)	1276 (63)	1314 (71)
No preview	2133 (101)	2094 (110)	1961 (109)	1998 (114)	1286 (45)	1322 (80)	1310 (81)	1316 (81)

served in three different studies, each using different scene materials (this study; Vö & Wolfe, 2012a, 2012b).

Preview Effect

Finally, there were reliable effects of preview. In Block 1, elapsed time to target fixation was reliably shorter when the scene had been previewed (1475 ms) than when it had not (1578 ms), $F(1, 46) = 12.0$, $p < .001$, $\eta_p^2 = .21$. The type of preview instructions did not modulate the preview effect ($F < 1$), and there was no interaction between preview and epoch, $F(3, 138) = 1.04$, $p = .38$, $\eta_p^2 = .02$. In Block 2, mean elapsed time was again shorter in the preview condition (928 ms) than in the no-preview condition (972 ms), $F(1, 46) = 9.02$, $p = .004$, $\eta_p^2 = .16$.

To probe the relationship between eye movement behavior during the preview session and later search efficiency, we examined the correlation between the total fixation time on an object during the preview session and the elapsed time to fixation of that object as a target in search (see Figure 3). Each search event was treated as an observation. Specifically, the data consisted of (1) the elapsed time to target fixation for each search in the preview condition of Block 1 paired with (2) the total time that the target object had been fixated by the participant during the preview. Mean elapsed time to fixation during search was 1887 ms for objects that were not fixated in the preview. When the target had been fixated in the preview, there was a reliable negative correlation ($r = -.12$) between these variables, $t(2151) = -6.11$, $p < .001$. Objects fixated longer during the preview session were found faster during search.

The preceding analysis might have been compromised by differences between target objects. For example, large objects might have been preferentially fixated in the preview and more conspicuous during search, producing a correlation unrelated to memory. Thus, an additional analysis was conducted treating target object as a random effect. For each of the 144 target objects, the 24 searches for the object in Block 1 were split evenly by preview fixation

duration on that object. Mean elapsed time to fixation during search was 1429 ms for the searches with relatively long preview fixation times and 1606 ms for the searches with relatively short preview fixation times, $t(143) = 2.95$, $p < .001$, $\eta_p^2 = .11$. Thus, not only did we observe an overall effect of preview on search, there was a direct relationship between the fixation of a particular object during the preview and the formation of a memory representation that facilitated later search.

Discussion

We replicated two findings from Vö and Wolfe (2012b). First, there was a repeated search effect: The second search for a particular object was more rapid than the first. Second, search times decreased over the course of multiple searches through the scene, again demonstrating that distractor memory transfers to later searches. Thus, we can be confident that the present study was sensitive to the same mechanisms probed by Vö and Wolfe. The key evidence came from the preview manipulation. The availability of a scene preview reliably facilitated later search through the scene, and eye movement behavior during the preview was directly related to later search performance. Thus, visual memory representations formed in the context of a nonsearch task reliably transferred to visual search. The most plausible explanation for the null effect of preview observed by Vö and Wolfe is that comparisons between preview and no-preview conditions were hampered by group differences in baseline search efficiency. In addition, it is unlikely, given Vö and Wolfe's discussion of power, that their design could have detected a preview effect of the magnitude observed here (103 ms in Block 1).

Strikingly, the preview effect in the present study was independent of the instructions during the preview session. Instructions to remember the locations of the objects, which one might expect to have maximized transfer to visual search, did not generate a larger preview benefit than instructions to judge the objects' semantic

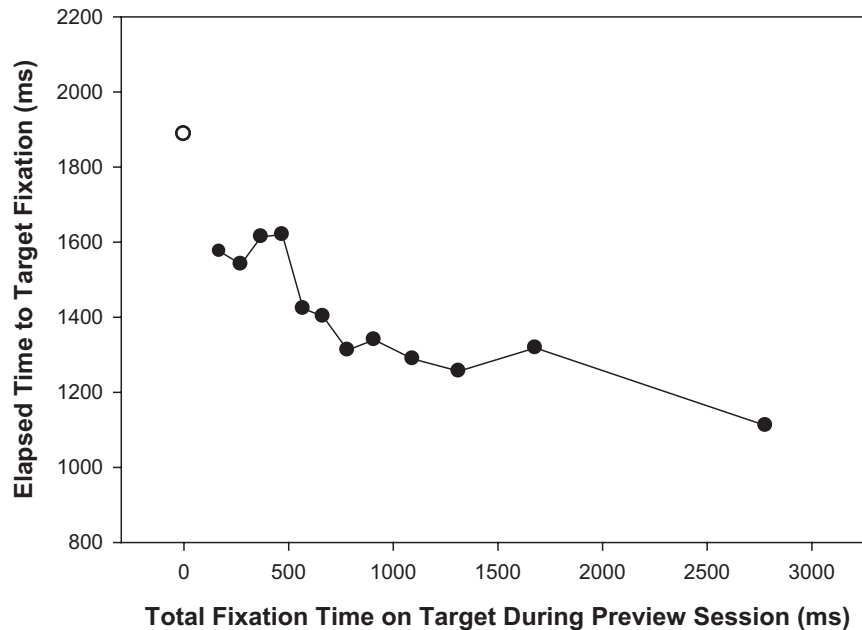


Figure 3. Elapsed time to target fixation during search as a function of the total time that the target object was fixated during the preview session. Mean elapsed time for targets that were not fixated during the preview session is indicated by the open circle. Searches on which the target was fixated in the preview (solid circles) were divided into 12 equal bins by preview fixation time, with mean elapsed time in each bin plotted against mean preview fixation time in that bin.

properties. Thus, memory representations generated under substantially different task conditions (neither of which involved visual search) were applied in a functionally similar manner to guide later search. This is not to say that the visual memory representations formed while viewing a scene are independent of task. The encoding of object information is strongly dependent on attention and gaze (Hollingworth & Henderson, 2002; Irwin & Gordon, 1998; Loftus, 1972; Schmidt, Vogel, Woodman, & Luck, 2002), and the allocation of attention and gaze is strongly tied to task demands (Castelhamo, Mack, & Henderson, 2009; Hayhoe, 2000; Land, Mennie, & Rusted, 1999; Yarbush, 1967). Thus, the memory representation of a scene will be constrained by task (Ballard et al., 1995; Droll et al., 2005). However, the present data indicate that once encoded, the application of visual memory to future behavior is not strongly constrained by task. Visual memory representations appear to have a substantially “general purpose” character, efficiently applied to tasks significantly different from those in which they were encoded.

References

- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*, 66–80. doi:10.1162/jocn.1995.7.1.66
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 14325–14329. doi:10.1073/pnas.0803390105
- Brockmole, J. R., Castelhamo, 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. doi: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. doi:10.1080/13506280500165188
- Castelhamo, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 753–763. doi:10.1037/0096-1523.33.4.753
- Castelhamo, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*, 6, 1–15. doi:10.1167/9.3.6
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71. doi:10.1006/cogp.1998.0681
- Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1416–1438. doi:10.1037/0096-1523.31.6.1416
- Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition*, *7*, 43–64. doi:10.1080/135062800394676
- Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228. doi:10.1037/0096-1523.25.1.210
- Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 519–537. doi:10.1037/0096-1523.30.3.519
- Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1416–1438. doi:10.1037/0096-1523.31.6.1416

- Psychology: Learning, Memory, and Cognition*, 31, 396–411. doi:10.1037/0278-7393.31.3.396
- Hollingworth, A. (2006). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*, 14, 781–807. doi:10.1080/13506280500193818
- Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*, 17, 273–291. doi:10.1080/13506280802193367
- Hollingworth, A. (2012). Guidance of visual search by memory and knowledge. In M. D. Dodd & J. H. Flowers (Eds.), *The influence of attention, learning, and motivation on visual search, Nebraska symposium on motivation* (pp. 65–91). New York: Springer. doi:10.1007/978-1-4614-4794-8_4
- Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 113–136. doi:10.1037/0096-1523.28.1.113
- Howard, C. J., Pharaon, R. G., Körner, C., Smith, A. D., & Gilchrist, I. D. (2011). Visual search in the real world: Evidence for the formation of distractor representations. *Perception*, 40, 1143–1153. doi:10.1068/p7088
- Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention, and trans-saccadic memory. *Visual Cognition*, 5, 127–155. doi:10.1080/713756783
- Körner, C., & Gilchrist, I. D. (2007). Finding a new target in an old display: Evidence for a memory recency effect in visual search. *Psychonomic Bulletin & Review*, 14, 846–851. doi:10.3758/BF03194110
- Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, 28, 1311–1328. doi:10.1068/p2935
- Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 3, 525–551. doi:10.1016/0010-0285(72)90021-7
- Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research*, 46, 614–621. doi:10.1016/j.visres.2005.08.025
- Neider, M. B., & Zelinsky, G. J. (2008). Exploring set size effects in scenes: Identifying the objects of search. *Visual Cognition*, 16, 1–10. doi:10.1080/13506280701381691
- Pomplun, M. (2006). Saccadic selectivity in complex visual search displays. *Vision Research*, 46, 1886–1900. doi:10.1016/j.visres.2005.12.003
- 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. doi:10.3758/BF03194742
- 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. doi:10.1037/0033-295X.113.4.766
- Võ, M. L. H., & Wolfe, J. M. (2012a). *The interplay of episodic and semantic memory in guiding repeated search in scenes*. Manuscript submitted for publication.
- Võ, M. L. H., & Wolfe, J. M. (2012b). When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 23–41. doi:10.1037/a0024147
- Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception, & Psychophysics*, 73, 1650–1671. doi:10.3758/s13414-011-0153-3
- 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. doi:10.1016/j.tics.2010.12.001
- Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Received June 19, 2012

Revision received August 10, 2012

Accepted August 13, 2012 ■