

Incidental visual memory for targets and distractors in visual search

CARRICK C. WILLIAMS, JOHN M. HENDERSON, and ROSE T. ZACKS
Michigan State University, East Lansing, Michigan

We explored incidental retention of visual details of encountered objects during search. Participants searched for conjunction targets in 32 arrays of 12 pictures of real-world objects and then performed a token discrimination task that examined their memory for visual details of the targets and distractors from the search task. The results indicate that even though participants had not been instructed to memorize the objects, the visual details of search targets and distractor objects related to the targets were retained after the search. Distractor objects unrelated to the search target were remembered more poorly. Eye-movement measures indicated that the objects that were remembered were looked at more frequently during search than those that were not remembered. These results provide support that detailed visual information is included incidentally in the visual representation of an object after the object is no longer in view.

Normal viewing of scenes and other complex visual stimuli involves a succession of fixations in which high-quality visual information is received from the region of the visual field falling on or near the fovea, an area subtending about 2° of visual angle. The deployment of fixations across the visual field thus provides a series of high-fidelity but small-extent “snapshots” of the input. Given this situation, a central question about scene perception has been whether memory for visual details encoded during successive fixations is essential to building and maintaining a representation of the scene. Loftus and colleagues (Loftus & Bell, 1975; Loftus & Kallman, 1979) suggested that picture recognition is based on general and detailed information, and they pointed out the usefulness of visual details in recognition. Nevertheless, some theories of visual processing claim that little or no visual memory is necessary for the processing of scenes (O’Regan, 1992; O’Regan & Noë, 2001; Rensink, 2000; Rensink, O’Regan, & Clark, 2000). In contrast to these theories, there is good evidence that relatively detailed scene representations are in fact generated and retained both in short- and long-term memory (Angelone, Levin, & Simons, 2003; Castelano & Henderson, in press; Hen-

derson & Hollingworth, 2003a, 2003b; Hollingworth, 2003a, 2004; Hollingworth & Henderson, 2002; Hollingworth, Schrock, & Henderson, 2001; Hollingworth, Williams, & Henderson, 2001; Mitroff, Simons, & Levin, 2004; Simons, Chabris, Schnur, & Levin, 2002; Standing, Conezio, & Haber, 1970). For example, Hollingworth and Henderson presented participants with full-color three-dimensional rendered scenes with the instruction to memorize them. Participants were then asked to determine which of two object tokens (e.g., which of two clocks in a living room scene) had been presented. This type of discrimination requires memory for the visual details of the objects seen earlier, because the knowledge that a clock was present does not provide sufficient information for a correct response. Although the scenes had been out of sight for, in some cases, several minutes, participants were able to successfully choose the token seen earlier at better than chance levels, indicating that they retained relatively detailed visual information about the object.

When asked to memorize scenes, participants were successful in retaining the visual details of objects in them (Hollingworth & Henderson, 2002). However, in tasks that do not require memorization, memory representations could be relatively sparse. For example, it has been suggested that little or no memory is retained for what has been attended and where those items are during visual search (e.g., Horowitz & Wolfe, 1998, 2001). More recently, however, there has been evidence that participants do indeed retain some information about search objects and use this information online in the selection of objects to be examined (e.g., McCarley, Wang, Kramer, Irwin, & Peterson, 2003; Peterson, Kramer, Wang, Irwin, & McCarley, 2001). Research on the memory retained during visual search has focused primarily on the use of that memory in the deployment of attention to items during

This research was supported by National Institute of Aging Grant AGO 4306 to Lynn Hasher and R.T.Z., National Science Foundation Grant BCS 0094433 to J.M.H., and Army Research Office Grant DAAD 19-00-1-0519 to J.M.H. This research derives from C.C.W.’s doctoral dissertation at Michigan State University. Aspects of the data were reported at the third annual meeting of the Vision Science Society, Sarasota, FL, April 2003, and the 44th Annual Meeting of the Psychonomic Society, Vancouver, BC, November 2003. We thank dissertation committee members Tom Carr and Zach Hambrick for their helpful discussion and comments on the dissertation. Correspondence concerning this article should be addressed to C. C. Williams, who is now at the Department of Psychology, P.O. Box 6161, Mississippi State University, Mississippi State, MS 39762 (e-mail: cw499@msstate.edu).

the search task, but little research has focused on the information retained after the search. In other words, what visual information is retained of attended targets and distractors following search?

Although Hollingworth and Henderson (2002) demonstrated inclusion of visual details in the representations of scenes, the memorization instructions they used may have led to additional memory encoding that is not typical of viewing. The question arises as to whether the visual information encoded incidentally as a part of normal viewing is of a similar level of detail as that from intentional memorization. Castelhana and Henderson (in press) investigated this question by having participants view scenes under memorization or visual search instructions. Using a token discrimination task, they found little difference in memory performance based on the encoding task performed. Similar to Castelhana and Henderson, the visual search task used in the present study did not encourage participants to actively engage a memorization strategy while viewing the objects in a display. Because participants were attempting to find the search target rather than trying to memorize the details of the various objects, any encoding of the visual details of the objects in the display must be purely incidental.

In addition, we used an extremely difficult visual memory test that eliminated the contribution of semantic and contextual information to memory performance. We used a two-alternative token discrimination task that required a choice between two tokens of the same basic type and color. This discrimination can be made only on the basis of the visual details of the objects rather than a knowledge of the identity or semantic category of the particular object that had been presented. Unlike the Hollingworth and Henderson (2002) and Castelhana and Henderson (in press) test procedures, which involved presenting test items in the original viewing context or in a patch of the original context surrounding the object, each trial of the current memory test presented the previously seen object and the foil on an otherwise blank background. Given the task and memory test used, this study provided a strong test of the retention of relatively detailed visual information after viewing.

The second goal of the present study was to explore the relationship between allocation of attention and memory for different objects encountered in a search. When one searches an environment, different objects are more or less likely to be attended on the basis of their similarity to the search target. For example, looking for a red pen on a somewhat messy desktop is a type of conjunction search task where the searcher is trying to find the combination of two features (red and pen shaped). Some objects on the desktop, like a red coffee cup and blue pen, share features that define the sought-for object, whereas a brown stapler shares none of the features that define the search target. Although all of these objects are distractors in the search, it would seem that differences in their relationships to the search target would influence the allocation of attention to the different classes of dis-

tractors. This is precisely what was found by Kim and Cave (1995), who demonstrated that for detecting a probe dot on items in a visual search task, the relationship of the item to the search target was critical. Attentional effects were greatest (as indicated by faster and more accurate probe detection) for the targets, followed by the distractors that shared the target color, distractors that shared the target's shape, and lastly the objects that had none of the target features.

Theories of visual search (e.g., feature-integration theory, Treisman & Gelade, 1980; guided search, Wolfe, Cave, & Franzel, 1989) claim that before attention is deployed, some information about the items in the environment is processed. This *preattentive* processing allows attention to be allocated to those items that are most likely to be the target, which invites the question, what features can be processed without attention? Rock, Linnett, Grant, and Mack (1992) demonstrated that color, location, and number (coarsely coded) could all be processed to some extent without attention. However, they found that shape could not be processed without attention. In addition, Mack, Tang, Tuma, Kahn, and Rock (1992) proposed that Gestalt grouping could not be processed without attention, but this hypothesis has been contradicted by the findings of Moore and Egeth (1997). The fact that some basic features are processed prior to attention allows the searcher to prioritize the objects to examine. However, there appears to be a limit to the complexity of the information that can be extracted without attention.

Given that the allocation of attention is not uniform across the search space, it was expected that the visual details retained from a search would vary depending on the relationship of a particular distractor to the target object. It has been demonstrated by Potter and Levy (1969) that as the amount of time that a picture is viewed in a rapid serial visual presentation stream increases, recognition memory improves. In other words, the more time devoted to processing, and presumably attending, the picture, the better memory is. Furthermore, Intraub (1984) found that selectively attending to a picture improved memory for that picture over other pictures in a stream, even if the presentation duration was shorter than for the other pictures. As these studies indicate, attended objects are remembered better than are unattended objects; extending this finding to a visual search environment, the differences in the likelihood that an object will be attended should lead to differences in postsearch memory.

The present study examined incidental visual memory for objects previously encountered in a visual search. Participants were presented with a modified conjunction search task in which they were required to count the number of targets present in an array of photographs of 12 real-world objects. The search target in each trial was defined by a color-category combination (e.g., the telephone that is white). In addition to targets, there were objects related to the search targets either in color (white nontelephones) or category (nonwhite telephones). These objects are referred to as *related distractors* because they

possess a feature that defines the target object. Lastly, the search arrays contained objects that were not related to the target (objects that were neither white nor telephones; *unrelated distractors*).

Most research on visual search uses multiple tokens of each distractor in each display, repeating targets and distractors across trials. Repeated presentation of the stimuli both within and across trials makes it impossible to examine the stored visual information about individual items from a search. To examine visual memory for specific objects in the search arrays, every object presented in the present study was a unique token and appeared in only one search array. By using unique object tokens, we could evaluate visual memory for specific targets and distractors. It is important to mention that in the present study, each search array was presented twice within the experiment to increase the likelihood that the critical objects would be attended. Because targets appeared with the same distractors in both presentations of an array, the condition within which each distractor appeared was maintained across repetitions.

In conjunction search tasks like those presented here, there are many potential objects to which attention may be deployed that may or may not share a feature of the target. One way to examine the deployment of attention is to use eye-movement measures, because attention and fixation tend to be relatively tightly linked. There have been several examinations of eye movements and visual search (e.g., Bertera & Rayner, 2000; Findlay, 1997; Findlay, Brown, & Gilchrist, 2001; Hooge & Erkelens, 1999; Motter & Belky, 1998; Scialfa & Joffe, 1998; Shen, Reingold, & Pomplun, 2003; Shen, Reingold, Pomplun, & Williams, 2003; D. E. Williams, Reingold, Moscovitch, & Behrmann, 1997; I. G. Williams, 1967; Zelinsky, 1996; Zelinsky & Sheinberg, 1997). Williams found a preference to fixate stimuli sharing the target color. Other target characteristics, such as size and shape, were not as preferentially selected. In addition, several recent studies (e.g., Hooge & Erkelens, 1999; Shen, Reingold, Pomplun, & Williams, 2003; Zelinsky, 1996) found that participants were more likely to fixate distractor objects possessing a target feature than objects that did not. However, this preference did not account for all of the variance in object selection, indicating the imperfection of eye-movement guidance during search (Zelinsky, 1996).

We expected that differences in eye-movement behavior would partially account for the visual memory differences found for the different types of objects. Specifically, we expected that objects that were related to the search target would be examined more frequently and for a longer time than those that were not. Targets should be viewed on almost every trial, because the task was to count these objects. Because the related distractors possessed a target feature, they should be the next most likely to be viewed, and unrelated distractors should be the least likely.

To preview the results, memory was reliably better for targets than for distractors. Of most interest for this study,

however, was memory for the distractor items. Related distractors were discriminated from foils at similar levels, and both were discriminated reliably better than unrelated distractors. The eye movements made during the search task demonstrated patterns similar to the memory results, in that the objects that were remembered better after search tended to be looked at more frequently and for a longer time during search. Overall, we conclude that visual search for real-world objects leaves a lingering memory trace of both targets and attended distractors. This trace is incidentally generated, is relatively long lasting, and contains enough visual detail to support discrimination of a presented object from a visually similar memory foil.

METHOD

Participants

Twenty-four participants were recruited from general psychology courses at Michigan State University.

Design

This experiment included three factors: number of targets, array presentation, and object type. During the search task, the participants were required to count the number of targets that were present (ranging from 0 to 3). Second, there were two presentations of each array (first and second presentation). Finally, each search array consisted of four types of objects: targets, category distractors, color distractors, and unrelated distractors. The dependent variables of primary interest were visual memory performance and the eye-movement measures for the four different types of objects. However, measures of search efficiency and accuracy were also examined.

Materials

Sixty-four search categories were divided into two groups of 32 categories each. A participant only searched for targets from one of the two groups of categories. Targets were defined by a specific category–color combination. For each category, three colors were selected in which the target could appear (e.g., white, yellow, or blue telephones). An object could be one of eight possible colors (black, brown, white, silver, red, blue, yellow, or green), each of which was used equally often as a target color (i.e., for four targets) across categories and participants. For each color within a category, five photographic exemplars were selected, primarily from an electronic database of pictures of real-world objects. Each object was resized so that its longest dimension (horizontal or vertical) was 3.9° in visual angle (from a viewing distance of 57 cm) while maintaining its original proportions.

Search arrays were constructed for each color–category combination (see Figure 1). A search array consisted of 0–3 targets (e.g., a white telephone), 3–4 category distractors, 3–4 color distractors, and 3–4 unrelated distractors.¹ The category distractors were objects from the other two colors of the search category (e.g., blue and yellow telephones). The color distractors and unrelated distractors were selected from the nonpresented set of 32 categories for that participant. Thus, the color and unrelated distractors were from categories that were never searched for by a participant, preventing any confusion about the distractor's relationship to the target. Color distractors matched the target on the color that was being searched for, and within an array each color distractor came from a different category. Unrelated distractors were not related to the search target by either search category or search color. Each unrelated distractor was a different color and from a different category than the other distractor objects. For each search array, one target, one color distractor, one category distractor, and one unrelated distractor were selected

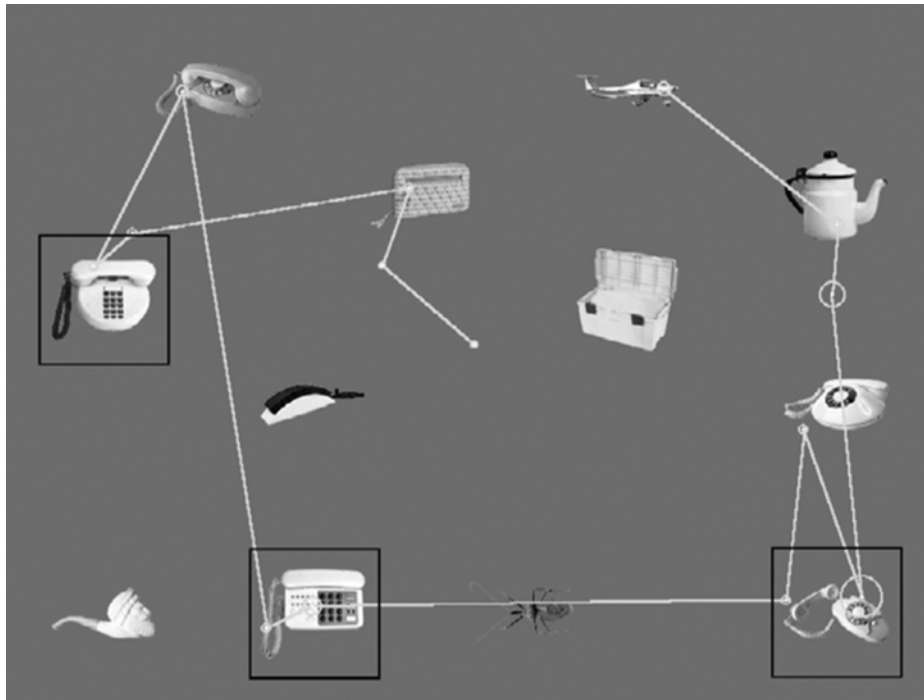


Figure 1. An example of a search array and scan pattern for a participant looking for white telephones. Images in this study were presented in color on a neutral gray background. Saccades are represented by lines and fixations by circles. The size of the circle is proportional to the fixation duration. The three white telephones are highlighted in the figure by the black boxes, which are the size of the scoring region used. Any fixation that fell within the region was counted as being on the object. The color distractor was the white pipe (lower left corner), the category distractor was the yellow telephone (middle right), and the unrelated distractor was the green lady's wallet (left center).

for the visual memory test. The objects were counterbalanced in that each memory-tested object served as a target, a color distractor, a category distractor, and an unrelated distractor across participants.

Visual Search Task

An array of 12 objects was created for each of the factorial combinations of the number of targets (4 levels), the 3 target colors, and 64 categories. The arrays were 800×600 pixels in size (approximately $35^\circ \times 26^\circ$) with a neutral gray background (RGB = 120). Each search category was assigned to one of eight object configurations based on a grid of 48 (8 horizontal \times 6 vertical) $3.9^\circ \times 3.9^\circ$ squares. The 12 objects for each array were placed into separate squares in the grid with the constraint that squares containing objects could not be closer than 3.9° from the nearest object square. The objects were never closer to the edge of the display than 40 pixels in the horizontal dimension and 30 pixels in the vertical dimension.

As mentioned earlier, each array was presented twice in the search task. The arrays were presented in two blocks, and each of the 32 arrays appeared in each block once. The number of targets and the object configuration were changed in each block (e.g., if there were two targets in the first viewing, there could be one or three targets in the second viewing). These changes required the participant to search the array on the second presentation rather than relying on memory for the number of targets or for the location of targets and distractors. To ensure that a target was processed for a search array, a target was always present in the second block. All tested distractors were present in both presentations of an array. The participants were not told that in the second block, all of the arrays had at least one target. Each participant was presented with a total of 384 unique objects in the visual search task.

Visual Memory Test

The memory test was a two-alternative forced-choice token discrimination task. Two tokens were presented on each side of a central fixation cross on an otherwise empty gray field. One target, one color distractor, one category distractor, and one unrelated distractor were tested from each search array. The participants were asked to determine which object was identical to one presented in the earlier search task. The nonpresented foil was selected so that it matched the semantic label of the presented object. In other words, if the presented object was a yellow car, the foil was also a yellow car that had not been seen by that participant. The memory foils for half of the participants were the presented objects for the other half. The objects were presented at the same size (3.9°) as in the search displays.

Procedure

Upon arrival, the participant was provided with a brief description of the experiment along with the informed consent form and a general demographic questionnaire. After these forms were completed, the experimenter positioned the participant approximately 57 cm from the computer screen, and the participant was presented with the following instructions for the search task:

In this task, you will be given an item to search for in an object array (e.g., "the saw that is yellow"). Your job is to count the number of objects in the array that fit the description given. Input the number of search targets (0, 1, 2, or 3 objects) on the button box in front of you. We want you to be as accurate as possible in your responding, so take as much time as you need to be certain of your response.

The experimenter then gave a sample array to familiarize the participant with the types of judgments that he or she would have to

make. The participant was then informed that while he or she was looking for the objects described, his or her eye movements would be monitored. A chin- and forehead rest was used to maintain relative stability during eye tracking. Once the participant was comfortable in the chinrest, the experimenter adjusted the eye-tracking camera to obtain a trackable image, and the participant was asked to fixate nine points on an otherwise-blank screen for calibration. A trial began with a display of the search target description (e.g., "the saw that is yellow"). The participant read the description and looked back to a central box. When the experimenter determined that the participant was looking at the central box, the search array was presented. After each trial, the calibration screen was presented, allowing the experimenter to determine whether the system was still calibrated. The stimuli were presented using E-prime Version 1.0 software.

Following the search task, the participants worked on unrelated tasks (e.g., a vocabulary test) for approximately 10 min, and then were given the visual memory test. The instructions for this test informed the participants that they would see pairs of pictures and that they should do their best to decide which picture (left or right) was presented during the visual search task. In addition, the participants were informed that the test would be very difficult and to take their best guess if they were uncertain. Half of the targets appeared on the left, and half appeared on the right.

Equipment

Eye movements were monitored using an ISCAN RK-726PCI pupil-tracking system sampling at 240 Hz. The calibrated position of this system is accurate to within 1° of visual angle horizontally and vertically. Because of this limitation, the scoring region of each object (the area in which a fixation must fall to be counted as being on the object) was slightly larger than the object itself (0.4° larger than the object). The ISCAN equipment can accommodate minor head movements, and thus no bite bar was required. However, a chin- and forehead rest was used to limit larger head movements.

Eye-Tracking Analyses

The eye-tracking data consist of x, y coordinates of the calibrated eye position for each sample of the eye-tracking system. The ISCAN system used in this experiment samples at 240 Hz or roughly one sample every 4 msec. To stabilize the eye-tracking record, each sample was averaged with the two preceding and the two following samples in the postexperimental analysis. Fixations were determined by grouping consecutive samples that were no more than eight pixels in Euclidian distance from the previous sample. Any fixation less than 90 msec or greater than 4,000 msec was discarded.

All analyses were performed on object regions, defined as the $4.7^\circ \times 4.7^\circ$ box ($3.9^\circ + 0.4^\circ$ in each dimension) centered on the object that had been used to place it in the array. Fixations were determined to be on a particular object if it fell within this region. An example of the scoring region can be seen in Figure 1.

RESULTS

For all statistical tests, an alpha level of .05 was used to determine statistical reliability.

Visual Search

The two measures of interest for the visual search task were search time and accuracy in counting the number of targets. Because each array was presented twice, the search data were divided into first and second presentations (Table 1). However, because the zero target condition was limited to the first presentation, the arrays for which there were zero targets in the first presentation were eliminated from these analyses.

Correct search times were faster for the second presentation of the search arrays than for the first presentation [$F(1,23) = 12.27, MS_e = 628,516, p = .002$]. There was no effect of the number of targets in the search array [$F(2,46) = 2.07, MS_e = 740,134, p = .138$] nor was there an interaction between presentation and the number of targets [$F(2,46) = 1.76, MS_e = 442,635, p = .183$].

In contrast, for search accuracy, there was no improvement in accuracy on the second presentation [$F(1,23) = 1.94, MS_e = 0.010, p = .18$]. Also, there was an effect of the number of targets [$F(2,46) = 4.80, MS_e = 0.014, p = .013$], with accuracy decreasing as the number of targets in the search array increased. There was an interaction of presentation and number of targets [$F(2,46) = 4.34, MS_e = 0.018, p = .019$], with accuracy improving in the three-target condition across the two presentations, but not in the other conditions.

Visual Memory

The more important results are the visual memory performance results. Because one-quarter of the target

Table 1A
Mean Correct Search Times (in Milliseconds), With Standard Errors

Presentation	Number of Targets							
	0		1		2		3	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
1	4,985	326	4,463	238	5,005	363	4,924	311
2			4,287	276	4,455	256	4,261	317

Table 1B
Mean Search Accuracies (Proportion Correct), With Standard Errors

Presentation	Number of Targets							
	0		1		2		3	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
1	.91	.026	.89	.023	.84	.029	.75	.025
2			.87	.021	.80	.025	.87	.028

objects for each participant were only viewed once, all of the analyses reported here, unless otherwise stated, eliminated these singly viewed target objects from consideration. This ensured that target objects and distractor objects had equal opportunities to be viewed during the search task.

Because a two-alternative forced-choice design was used, chance level for the memory test was 50%. Two tests addressed the primary questions of this study. The first test compared the different object conditions with chance level using one-sample t tests. If objects in any of the conditions are discriminated at better than chance levels, some visual information of the objects has been retained from the incidental encoding during the visual search task. The second test compared the different types of objects to determine whether the relationship to the search target affected visual memory.

The visual memory results are shown in Figure 2. All classes of objects were remembered better than chance [proportion correct $\geq .55$; all $t_s(23) \geq 3.36$, all $p_s \leq .003$], indicating that some visual details were retained after search. Turning to the comparison of differences between the object conditions, there was an overall main effect of condition [$F(3,69) = 80.43$, $MS_e = 0.0057$, $p < .001$]. Targets were remembered better than were category distractors [$t(23) = 9.98$, $p < .001$] and color distractors [$t(23) = 11.70$, $p < .001$]. Objects in the two related conditions were remembered at similar levels [$t(23) = 0.14$, $p > .20$], and both types of related distractors (though marginally in the case of color distractors) were remembered better than unrelated distractors [category distractors, $t(23) = 2.28$, $p = .032$; color distractors, $t(23) = 2.02$, $p = .055$].

Because a subset of the tested target objects was viewed only in the second presentation block with the remaining

target objects being viewed in both presentation blocks, it is possible to provide some indication of the role of additional processing for visual memory. The participants had better memory for targets viewed twice (.86) than for those viewed once (.78) [$F(1,23) = 7.02$, $MS_e = 0.009$, $p = .01$]. The influence of number of viewings will be further examined using the eye movement data.

Eye Movement

The eye-tracking analyses we report focus on the data from memory-tested objects. Thus, eye-tracking results for one target, one category distractor, one color distractor, and one unrelated distractor from each array were examined. However, because the participants did not know which object was going to be tested while they were viewing the array (or, for that matter, that there was a test at all), the tested objects are equivalent to the other objects in each condition. The three measures of specific interest were the proportion of arrays in which an object was viewed (proportion viewed), the number of separate looks at an object (number of entries), and the total time an object was examined during search. An example scan pattern can be seen in Figure 1.

Proportion viewed. The proportion-viewed measure reflects the proportion of arrays in which a particular object was viewed in either (or both) of the trials in which an array was presented. "Viewed" in this case means that at least one fixation landed in the object region. Because the targets were viewed on almost every trial (94% and 96% of trials for the first and second presentations, respectively), only the distractor conditions were analyzed in the proportion-viewed analysis. However, all conditions are presented in Figure 3 (top panel). Most memory-tested objects were viewed during the search phase, although they may not have been viewed on both pre-

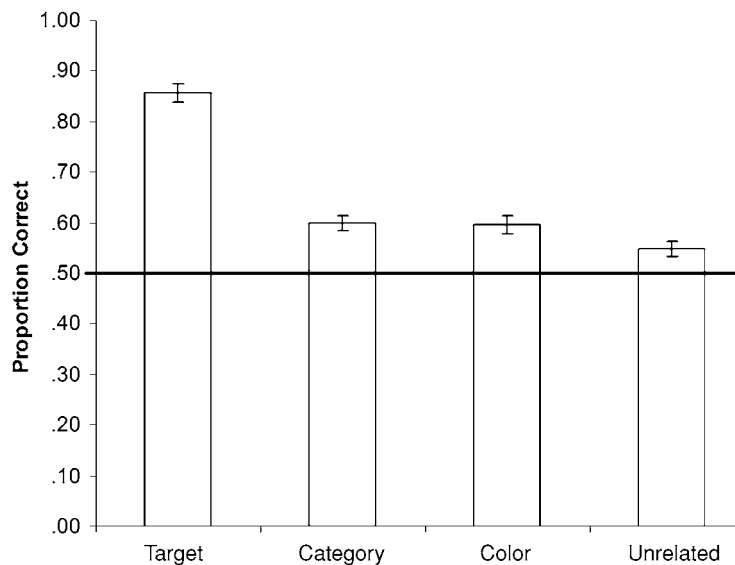


Figure 2. Mean proportions correct for the visual memory test. Error bars represent standard errors. The solid line represents chance performance.

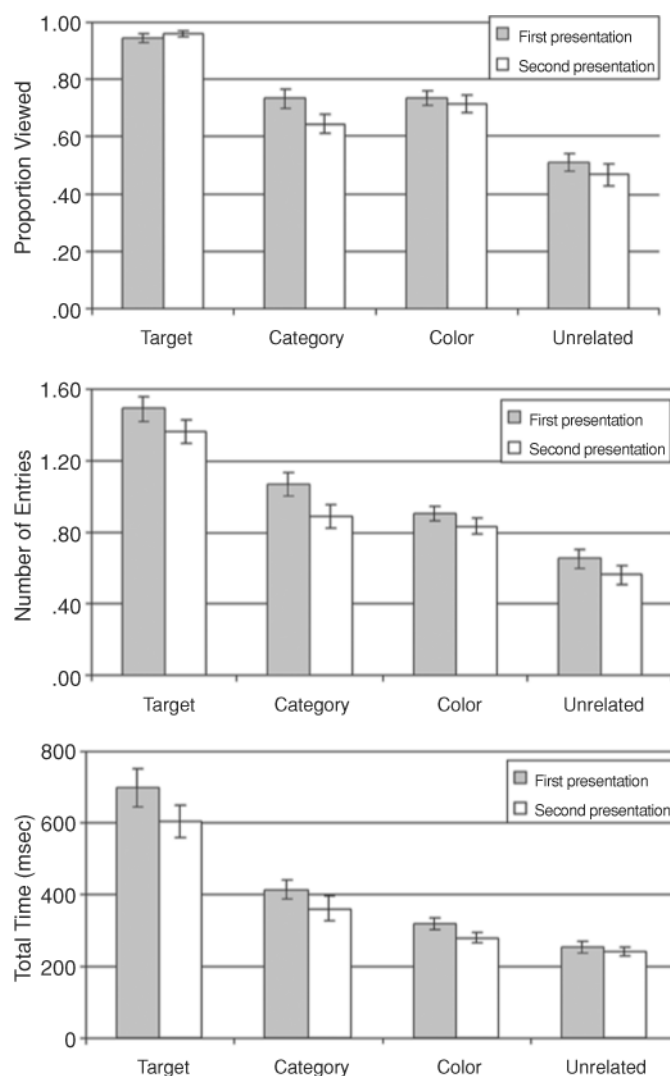


Figure 3. Eye-movement measures for memory-tested objects in visual search task for each presentation of the array: proportions of objects viewed, mean numbers of entries into the object regions, and mean total fixation times within the object regions. Error bars represent standard errors.

presentations. There was a reliable effect of array presentation [$F(1,23) = 10.32$, $MS_e = 0.009$, $p = .004$], with more objects being viewed during the first presentation than during the second presentation. There was also an effect of object condition [$F(2,46) = 59.94$, $MS_e = 0.013$, $p < .001$], but there was no interaction of presentation and distractor condition [$F(2,46) = 2.14$, $MS_e = 0.007$, $p > .10$]. Because there was no interaction with presentation, we collapsed across the two presentations to compare the proportion viewed with the distractor conditions.² Both types of related distractors were viewed equally often [$t(23) = 0.33$, $p > .20$], and both were more frequently viewed than the unrelated distractors [category distractors, $t(23) = 7.77$, $p < .001$; color distractors, $t(23) = 6.86$, $p < .001$]. The similarity of these results and the memory results supports the im-

portance of viewing an object for finding memory of the visual details.

Number-of-entries analysis. The visual memory retained after search could also depend on the number of separate times an object is viewed during search, referred to as the *number of entries*. For example, to count the number of entries to the white telephone in the lower right-hand corner of Figure 1, we see that the participant initially fixated the white telephone, left the object region to fixate a yellow telephone (directly above the white telephone), and then returned to the white telephone, yielding an entry count of 2. An entry count of 0 indicates that an object was not viewed on a particular trial.

As in the proportion-viewed analyses, the effect of presentation was reliable [$F(1,23) = 34.44$, $MS_e = 0.019$, $p < .001$] with objects entered more frequently during

the first than during the second presentation. In addition, the different types of objects were entered at different rates [$F(3,69) = 115.07$, $MS_e = 0.049$, $p < .001$], but, as in the proportion-entered data, presentation and object condition did not interact [$F(3,69) = 1.84$, $MS_e = 0.016$, $p > .10$]. Again, we collapsed across the two viewings (by summing the data from both presentations) to examine the object condition effects more closely. As can be seen in Figure 3 (middle panel), targets were entered more frequently than were category distractors [$t(23) = 8.64$, $p < .001$]. Category distractors, in turn, were entered more frequently than were color distractors [$t(23) = 2.54$, $p = .018$], and finally, color distractors were entered more frequently than were unrelated distractors [$t(23) = 7.03$, $p < .001$]. This pattern is slightly different from the proportion-viewed or visual memory results because a difference was found between the related distractors. Category distractors were entered more frequently than were color distractors, which could indicate that category distractors required more processing to determine that they were not targets.

Total-time analysis. Total time is simply the sum of all fixation time a participant spent on an object during a presentation. Only objects viewed at least once were included in this analysis. The total time could be potentially more informative than number of entries or proportion viewed, since it can be thought of as a direct measure of processing time.

The total time analysis yielded results similar to the entry analysis with a main effect of presentation [$F(1,23) = 10.78$, $MS_e = 10,834$, $p = .003$; see Figure 3 (bottom panel)], and object condition [$F(3,69) = 89.72$, $MS_e = 17,245$, $p < .001$], but no interaction [$F(3,69) = 1.62$, $MS_e = 7,941$, $p > .10$]. We collapsed across the two viewings by summing across the two presentations to examine the object condition effect. Target objects were viewed longer than were category distractors [$t(23) = 9.27$, $p < .001$]. Category distractors were in turn viewed longer than were color distractors [$t(23) = 4.08$, $p < .001$], which in turn were viewed longer than were unrelated distractors [$t(23) = 8.09$, $p < .001$].

Regression Analyses

The next section examines the relationship of two different eye-movement measures to the visual memory test accuracy. We tested the hypothesis that the more an object was processed (as indexed by fixation), the greater memory test accuracy would be. The analyses regressed visual memory performance against the collapsed data for number of entries and total time.³ To simplify the graphical presentation (Figure 4), rather than present all of the data points analyzed, we plot number of entries (top panel) and total time (bottom panel), divided into quartile bins for each condition against the mean visual memory accuracy for each quartile.

Number of entries. Regressions were performed separately for each condition. There was a reliable relationship between the number of entries and memory test accuracy for the category distractors [$r_{pb} = .16$; $F_{\text{change}}(1,743) =$

18.39 , $p < .001$] and color distractors [$r_{pb} = .10$; $F_{\text{change}}(1,743) = 7.55$, $p = .006$]. There was not a reliable relationship between memory test accuracy and the entries in the target condition [$r_{pb} = .04$; $F_{\text{change}}(1,551) = 10.2$, $p > .20$] or the unrelated distractor condition [$r_{pb} = .05$; $F_{\text{change}}(1,743) = 1.51$, $p > .20$].⁴

One interesting aspect of Figure 4 is that target objects appear to have been remembered better than distractors, regardless of the number of separate looks at the objects. Even in the quartile with the most entries for each of the distractor conditions (all of which have more entries than the smallest number-of-entries quartile for the target objects), memory performance did not reach the level of the lowest quartile for the target objects.

Total time. The same type of regression analysis related total viewing time and memory test accuracy within each object condition. As in the number-of-entries analysis, a reliable relationship between total time and memory test accuracy (Figure 4, bottom panel) was found for category distractors [$r_{pb} = .19$; $F_{\text{change}}(1,656) = 25.50$, $p < .001$] and color distractors [$r_{pb} = .13$; $F_{\text{change}}(1,656) = 11.58$, $p = .001$]. The relationship between visual memory and total time spent on targets was marginally reliable [$r_{pb} = .08$; $F_{\text{change}}(1,548) = 3.45$, $p = .06$]. As in the number-of-entries analysis, there was no reliable relationship between total time spent on an object in the unrelated distractor condition and memory test accuracy ($r_{pb} = .03$; $F_{\text{change}} < 1$).⁵

As can be seen in Figure 4, total time yielded a pattern similar to number of entries, with target objects remembered better than were distractor objects, regardless of the total time spent examining the objects. Again, target objects were remembered better than were distractors at all levels of total time, and memory performance in the distractor conditions did not reach the level of the lowest quartile for target objects.

DISCUSSION

The goals of this study were to provide a strong test of the retention of relatively detailed visual information incidentally acquired from real-world objects during visual search and to examine the nature of the visual memory resulting from search as a function of the viewing behavior that occurred during search. Visual memory should be tied to viewing behavior, which in turn should be tied to the relation of objects to the search target. Certain objects should be viewed more frequently or longer than others on the basis of their relationship to the search target. Our data clearly support the role of viewing behavior in visual memory for different objects.

The visual memory results demonstrate two important points. First, participants remembered all classes of objects better than chance. The present study provided a more stringent test of the maintenance of visual details in memory than has been done previously. Because participants did not expect a memory test, they had no reason to attempt to memorize the objects in the arrays, which provides evidence that detailed visual information

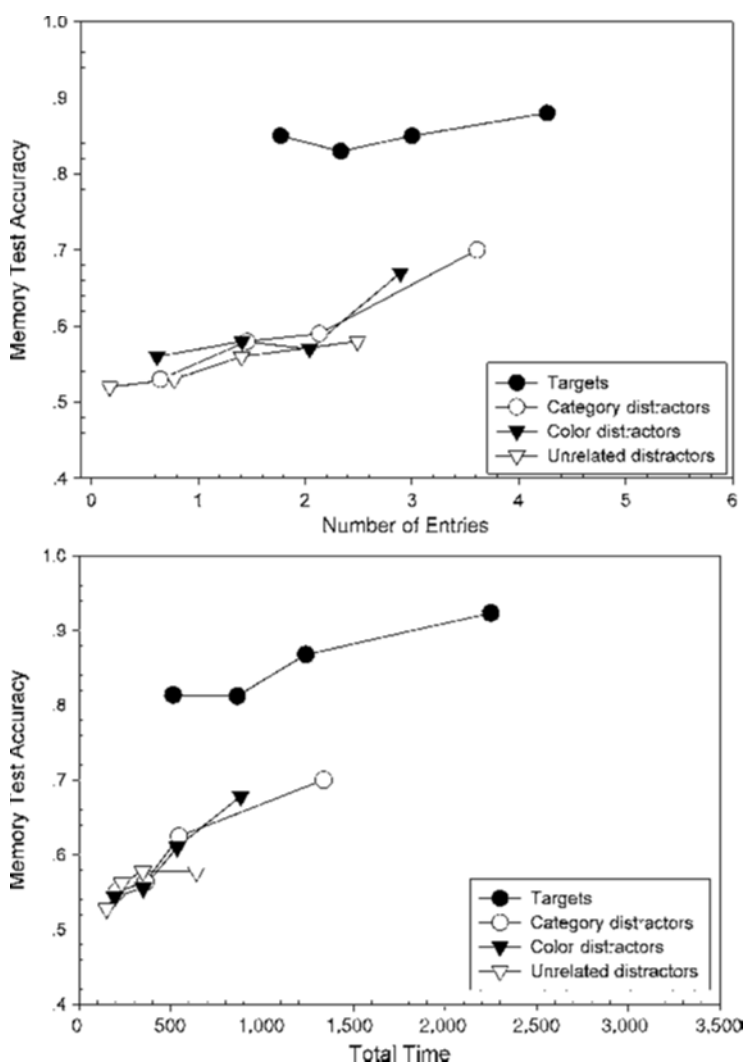


Figure 4. Memory test accuracy as a function of number of entries and total time. The data points represent the average for each quartile of the measure.

is included in the representation of objects during normal viewing. In addition, objects were learned in an array that did not form a coherent scene (see Figure 1), limiting the role of organization strategies in memory. Previous demonstrations of retained visual memory (Castelhano & Henderson, in press; Hollingworth & Henderson, 2002) have relied on full scenes as stimuli that may provide numerous potential cues for memory. The present study did not provide these internal cues, and thus any strategy that relied on the scene context was eliminated. Finally, the visual memory test required participants to decide which of two object tokens was presented. The test was devoid of context from the search phase, and thus the decision could be based on only the visual details attached to the object in memory. Given the demanding nature of the memory test, the results of the present study strongly argue against theories of visual memory claiming that no visual information is retained (O'Re-

gan, 1992; O'Regan & Noë, 2001; Rensink, 2000; Rensink et al., 2000).

The second finding of interest was the difference in the visual memory rates for the different object conditions. Unlike Castelhano and Henderson (in press), the present study explicitly manipulated the relationships of distracting objects to the target. This manipulation allowed for the examination of the role played by attentional allocation in both the search for and the memory of objects. The finding that targets were remembered better than were distractor objects was not a surprise, but the finding of reliable memory for distracting objects was. Specifically, objects related to search targets were remembered better than were objects that were not, supporting our expectation that memory is tied to the likelihood of an object being processed (see Kim & Cave, 1995). Furthermore, the finding that target objects that were viewed on both presentations were remembered better than those

that were viewed once demonstrates that the more an object is viewed, the better memory tends to be.

Turning now to the eye-movement data, the present study demonstrated that the eye-movement data generally parallel the memory data. The proportion-viewed analysis mirrored the visual memory results, with the targets being viewed in almost every array and showing the highest memory rates. Related distractors were viewed in fewer arrays than was the target, but category and color distractors were viewed on an equal proportion of the arrays. This similarity in the proportion viewed mirrors the similarity in the visual memory performance. Unrelated distractors were viewed on the smallest proportion of trials and had the worst memory performance.

The number-of-entries and total-time analyses showed a pattern similar to the memory results, each demonstrating a relationship to accuracy on the visual memory test. These eye movement measures are related in that an increase in entries for an object increases the likelihood that it will be viewed longer, but, as was mentioned earlier, they are not necessarily perfectly correlated. In the end, the two measures yielded similar patterns of results. Target objects were entered most frequently and were viewed for the longest time, and unrelated distractors were viewed least frequently and for the shortest time. Category distractors received the second most frequent entries and the second longest viewing times. However, in contrast to the proportion-viewed and visual memory results, color distractors were entered fewer times and for less overall time than were category distractors.

The difference between the two types of related distractors in the number-of-entries and total-time analyses indicates that while the objects were viewed on an equal proportion of trials, they were not viewed with equal frequency or for an equal length of time. One possible reason for this difference is that it may be easier to determine the category membership of an object (i.e., shape) in this experiment than its predominant color. This statement may appear to contradict the finding that category distractors were entered more frequently and viewed for a longer period of time than were color distractors. However, it is important to remember that the labels *category* and *color distractor* refer to the property an object shares with the target. Thus, a category distractor for the target *white telephone* could be a *yellow telephone*, whereas a color distractor could be a *white pipe*. Given that there were fewer entries and less total viewing of the color distractors, it appears that participants were better able to identify a white pipe as a distractor than a yellow telephone. This indicates that distractor shape was easier to determine than distractor colors, which may have led to a difference in the eye-movement measures. This difference could be the result of the stimuli used in the present experiment. Because the stimuli were photographs of real-world objects, there was considerable color variation among the objects representing each color category. Take, for example, a green drill, which could vary from a yellowish-lime green to a more bluish-green while maintaining a label *green*. In addition, real-world objects

are rarely just one color. A yellow car, for example, could contain several other colors. The tires, grille, and bumpers could all be (and most likely are) a color other than yellow, but because the predominant color is yellow, it is labeled a yellow car. It seems reasonable that if the predominant color of some objects was relatively difficult to determine, a participant would require more time or looks at an object to accurately judge its color.

The regression analyses for the number-of-entries and total-time data indicate that for the related distractor conditions, the more an object is viewed, the better visual memory tends to be, but targets and unrelated distractors did not demonstrate this relationship. Although the results were mixed, there does appear to be some relationship between viewing behavior and memory test accuracy. By demonstrating this relationship, the present study provides evidence of the role of viewing behavior in the formation of visual memory. The lack of a relationship in the unrelated distractor conditions is likely the result of the fact that memory is barely better than chance. The lack of systematicity in the memory data in this condition would make it difficult to try to find a relationship with another variable. A restriction of range, however, does not appear to be an issue in the target condition, making the lack of a correlation puzzling. In general, the regression analyses support the notion that visual memory depends at least partially on how the object was viewed during the search task.

Although the regression analyses demonstrate some relationship between viewing behavior and the visual memory results, target objects were remembered better than distractor objects, regardless of the amount of viewing that a distractor object received. Target objects were looked at more frequently and longer, but an examination of Figure 4 shows that memory for target objects is privileged beyond what would be expected on the basis of viewing behavior. One possible explanation for this privileged memory is that because the search task involved counting the number of targets in an array and because targets only differed from each other in visual details, participants performed a type of recognition test when counting the targets. They had to decide whether or not a target was one that they had seen before, thus performing an implicit recognition test while viewing the targets. The additional "testing" of the target objects may have improved memory for their visual details. It may also have led to an encoding strategy that emphasized the objects' visual details. This difference may have carried through to the visual memory test.

Finally, the effect of presentation in the eye-movement data indicates that when the search was repeated, participants were able to restrict the number of objects viewed. While it is possible that they were becoming more accurate in targeting their eye movements with practice, the lack of an interaction with object type indicates that eye movements in both presentations were similarly distributed. Thus, this effect was more likely due to participants understanding better the task demands.

Overall, the present study provides strong evidence for the inclusion of visual details in the representation of objects and demonstrates the importance of selective processes for the memory that remains after viewing. This is not to say that we believe that an exact, sensory-level representation is stored during viewing. Rather, we believe that the representation of the objects in memory is a higher level representation that contains some critical visual details, allowing viewers to make a distinction on the token level. When an object is viewed, a *long-term object file* is formed (Hollingworth & Henderson, 2002) that contains various visual details (e.g., orientation and structural relationship between parts) that are tied to specific locations in the visual environment (see Hollingworth 2003a, 2004, and Hollingworth & Henderson, 2002, for a further discussion of visual long-term memory). These long-term object files are built up through successive fixations on the objects and give the viewer access to their visual details. Although the list of features that are contained in object files has yet to be completely determined, the concept of the long-term object file provides a strong framework for understanding what is retained from viewing.

The present data fit well with the long-term object file conceptualization of visual memory in that we found a link between viewing behavior and memory. The more objects were looked at, the better memory tended to be. The limited nature of this relationship in the present data could result from the fact that our testing method removed the objects from the spatial context in which they were learned, limiting access to the spatially indexed object files. Hollingworth (2003b) demonstrated the importance of spatial context, and specifically spatial relations between objects, in visual memory. The fact that our visual memory test removed objects from the spatial relations they had with other objects in the array may have limited the ability of participants to access the long-term object file and thus reduced the overall relationship between viewing behavior and visual memory.

In summary, the present study demonstrates that even when people do not intend to remember the visual details of objects during search, they establish relatively good visual memory. The role of an object in search was critical for this memory, with memory being found primarily for objects related to the search target. In addition, viewing behavior appears to contribute to the memory of these related objects. Returning to the desktop search example presented in the introduction, it appears that people will look at and remember a red coffee mug or a blue pen when they look for a red pen. However, they are less likely to attend to, and so have a much worse memory for, the brown stapler on the desk.

REFERENCES

- ANGELONE, B. L., LEVIN, D. T., & SIMONS, D. J. (2003). The relationship between change detection and recognition of centrally attended objects in motion pictures. *Perception*, **32**, 947-962.
- BERTERA, J. H., & RAYNER, K. (2000). Eye movements and the span of the effective stimulus in visual search. *Perception & Psychophysics*, **62**, 576-585.
- CASTELHANO, M. S., & HENDERSON, J. M. (in press). Incidental memory for objects in scenes. *Visual Cognition*.
- FINDLAY, J. M. (1997). Saccade target selection during visual search. *Vision Research*, **37**, 617-631.
- FINDLAY, J. M., BROWN, V., & GILCHRIST, I. D. (2001). Saccade target selection in visual search: The effect of information from the previous fixation. *Vision Research*, **41**, 87-95.
- HENDERSON, J. M., & HOLLINGWORTH, A. (2003a). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics*, **65**, 58-71.
- HENDERSON, J. M., & HOLLINGWORTH, A. (2003b). Global transsaccadic change blindness during scene perception. *Psychological Science*, **14**, 493-497.
- HOLLINGWORTH, A. (2003a). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 388-403.
- HOLLINGWORTH, A. (2003b, November). *The structure of scene representations*. Paper presented at the 44th Annual Meeting of the Psychonomic Society, Vancouver, BC.
- HOLLINGWORTH, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 519-537.
- HOLLINGWORTH, A., & HENDERSON, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 113-136.
- HOLLINGWORTH, A., SCHROCK, G., & HENDERSON, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*, **29**, 296-304.
- HOLLINGWORTH, A., WILLIAMS, C. C., & HENDERSON, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, **8**, 761-768.
- HOOGE, I. T. C., & ERKELENS, C. J. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research*, **39**, 1567-1575.
- HOROWITZ, T. S., & WOLFE, J. M. (1998). Visual search has no memory. *Nature*, **394**, 575-577.
- HOROWITZ, T. S., & WOLFE, J. M. (2001). Search for multiple targets: Remember the targets, forget the search. *Perception & Psychophysics*, **63**, 272-285.
- INTRAUB, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 115-125.
- KIM, M., & CAVE, K. R. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychological Science*, **6**, 376-380.
- LOFTUS, G. R., & BELL, S. M. (1975). Two types of information in picture memory. *Journal of Experimental Psychology: Human Learning & Memory*, **1**, 103-113.
- LOFTUS, G. R., & KALLMAN, H. J. (1979). Encoding and use of detail information in picture recognition. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 197-211.
- MACK, A., TANG, B., TUMA, R., KAHN, S., & ROCK, I. (1992). Perceptual organization and attention. *Cognitive Psychology*, **24**, 475-501.
- MCCARLEY, J. S., WANG, R. F., KRAMER, A. F., IRWIN, D. E., & PETERSON, M. S. (2003). How much memory does oculomotor search have? *Psychological Science*, **14**, 422-426.
- MITROFF, S. R., SIMONS, D. J., & LEVIN, D. T. (2004). Nothing compares 2 views: Change blindness can occur despite preserved access to the changed information. *Perception & Psychophysics*, **66**, 1268-1281.
- MOORE, C. M., & EGETH, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 339-352.
- MOTTER, B. C., & BELKY, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research*, **38**, 1805-1815.
- O'REGAN, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, **46**, 461-488.

- O'REGAN, J. K., & NOË, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral & Brain Sciences*, **24**, 939-1011.
- PETERSON, M. S., KRAMER, A. F., WANG, R. F., IRWIN, D. E., & MCCARLEY, J. S. (2001). Visual search has memory. *Psychological Science*, **12**, 287-292.
- POTTER, M. C., & LEVY, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, **81**, 10-15.
- RENSINK, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, **7**, 17-42.
- RENSINK, R. A., O'REGAN, J. K., & CLARK, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition*, **7**, 127-145.
- ROCK, I., LINNETT, C. M., GRANT, P., & MACK, A. (1992). Perception without attention: Results of a new method. *Cognitive Psychology*, **24**, 502-534.
- SCIALFA, C. T., & JOFFE, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception & Psychophysics*, **60**, 1067-1082.
- SHEN, J., REINGOLD, E. M., & POMPLUN, M. (2003). Distractor ratio influences pattern of eye movements during visual search. *Perception*, **29**, 241-250.
- SHEN, J., REINGOLD, E. M., POMPLUN, M., & WILLIAMS, D. E. (2003). Saccadic selectivity during visual search: The influence of central processing difficulty. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 65-88). Amsterdam: Elsevier.
- SIMONS, D. J., CHABRIS, C. F., SCHNUR, T., & LEVIN, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness & Cognition*, **11**, 78-97.
- STANDING, L., CONEZIO, J., & HABER, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2,500 visual stimuli. *Psychonomic Science*, **19**, 73-74.
- TREISMAN, A. M., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- WILLIAMS, D. E., REINGOLD, E. M., MOSCOVITCH, M., & BEHRMANN, M. (1997). Patterns of eye movements during parallel and serial visual search tasks. *Canadian Journal of Experimental Psychology*, **51**, 151-164.
- WILLIAMS, L. G. (1967). The effects of target specification on objects fixated during visual search. *Acta Psychologica*, **27**, 355-360.
- WOLFE, J. M., CAVE, K. R., & FRANZEL, S. L. (1989). Guided search: An alternative to the feature integration model of visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 419-433.
- ZELINSKY, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, **36**, 2177-2187.
- ZELINSKY, G. J., & SHEINBERG, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 244-262.

NOTES

1. The numbers of color, category, and unrelated distractors depended on the number of targets present. For search arrays in which there were no targets, four color, four category, and four unrelated distractors were presented. For arrays in which one target was present, one unrelated distractor was removed. For arrays in which two and three targets were present, one category and one color distractor, respectively, were also removed.

2. In collapsing the proportion-viewed data, if an object was viewed on either presentation of an array, it was identified as viewed. Using this method, target objects were viewed on 99% of the trials, category and color distractors were viewed on 88% of the trials, and unrelated distractors were viewed on 70% of the trials.

3. Because each memory test trial counted as an independent observation, each participant contributed as many as 120 data points to the regression analyses. To remove potentially spurious effects due to individual differences among participants, we included participant as a variable in the regression model (Hollingworth & Henderson, 2002). We report F values for changes in R^2 , attributable to the number-of-entries and total-time variables. Because visual memory accuracy is a dichotomous variable, the resulting correlation coefficients are point-biserial correlations.

4. As in most of the analyses described in this study, the target condition regressions were performed on the data for targets presented twice. However, it is possible to examine the relationship of memory and the entries for all targets. If number of entries is related to visual memory, the number of opportunities one has to view the target is not as critical as the number of separate looks at the object. Therefore, an additional regression between the number of entries and visual memory was computed with all targets included. When all target data were considered, a reliable correlation of memory and number of entries was found [$r_{pb} = .08$; $F_{change}(1,743) = 4.69, p = .03$].

5. An additional regression was computed for all target objects for total time and visual memory accuracy. Again, there was a reliable relationship between total time and memory test accuracy for target objects in this analysis [$r_{pb} = .10$; $F_{change}(1,734) = 7.74, p = .006$].

(Manuscript received February 16, 2004;
revision accepted for publication October 7, 2004.)