Journal of Experimental Psychology:
Human Perception and Performance
2004, Vol. 30, No. 6, 1132–1146

0096-1523/04/$12.00 DOI: 10.1037/0096-1523.30.6.1132

Panoramic Search: The Interaction of Memory and Vision in Search Through a Familiar Scene

Aude Oliva
Massachusetts Institute of Technology

Jeremy M. Wolfe
Harvard Medical School and Brigham and Women's Hospital

Helga C. Arsenio
Brigham and Women's Hospital

How do observers search through familiar scenes? A novel *panoramic search* method is used to study the interaction of memory and vision in natural search behavior. In panoramic search, observers see part of an unchanging scene larger than their current field of view. A target object can be *visible*, present in the display but *hidden* from view, or *absent*. Visual search efficiency does not change after hundreds of trials through an unchanging scene (Experiment 1). Memory search, in contrast, begins inefficiently but becomes efficient with practice. Given a choice between vision and memory, observers choose vision (Experiments 2 and 3). However, if forced to use their memory on some trials, they learn to use memory on all trials, even when reliable visual information remains available (Experiment 4). The results suggest that observers make a pragmatic choice between vision and memory, with a strong bias toward visual search even for memorized stimuli.

The human visual environment is relatively stable and predictable: Objects do not float randomly in the ether, your computer screen is unlikely to vanish between fixations, and your cup of coffee should still be where you left it. If you are looking for an object in this stable and predictable world, it seems likely that object search mechanisms will take advantage of your knowledge of the scene. This knowledge will serve as a source of *top-down guidance* (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989), directing your attention to likely locations of the desired target. Although visual context is likely to benefit object processing, recent studies have shown that visual search efficiency does not necessarily improve with hundreds of repetitions of a stable visual scene (Wolfe, Klempen, & Dahlen, 2000; Wolfe, Oliva, Butcher, & Arsenio, 2002), a result that challenges the influence of familiar environments on search mechanisms. In this article, we investigate this issue by means of a novel experimental task, termed *panoramic search*. Panorama stimuli make it possible for observers to choose between vision and memory while searching for items.

A large body of evidence shows the effects of context on subsequent object processing. Seeing a familiar context has been shown to automatically activate the representations of consistent objects within the scene as well as their locations (Biederman, Mezzanotte, & Rabinowitz, 1982; de Graef, Christiaens, & d'Ydewalle, 1990; Henderson & Hollingworth, 1999; Palmer, 1975). When observers search for an object that is semantically consistent with its environment (e.g., a pen on a desk), their eyes fall on the object faster than when they are searching for a semantically inconsistent object (Henderson, Weeks, & Hollingworth, 1999). In a similar vein, while observers explore a scene, details of information about previously fixated objects are retained in long-term memory and used to plan further exploration of the image (Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001).

Other results show that one does not need detailed object information to recognize the semantic category of a scene (Biederman, 1987; Schyns & Oliva, 1994). A coarse spatial configuration of regions, automatically computed in a feed-forward manner, can be used to select regions in the image and guide the early deployment of the eyes toward locations likely to contain a specific target object (Oliva & Torralba, 2001; Oliva, Torralba, Castelhano, & Henderson, 2003; Torralba, 2003). Similar evidence has been seen in the performance of familiar tasks such as preparing food. Observers tend to look directly at task-relevant locations within the scene without being influenced by the visual salience of the objects per se (Land & Furneaux, 1997; Land & Hayhoe, 2001; Shinoda, Hayhoe, & Shrivastava, 2001).

Aude Oliva, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; Jeremy M. Wolfe, Department of Ophthalmology, Harvard Medical School, and Visual Attention Laboratory, Brigham and Women's Hospital, Boston, Massachusetts; Helga C. Arsenio, Visual Attention Laboratory, Brigham and Women's Hospital.

This research was supported by National Institute of Mental Health Grant MH56020 and Air Force Office of Scientific Research Grant F49620-01-1-0071 to Jeremy M. Wolfe. We thank Todd Horowitz and Melina Kunar, whose comments greatly improved the manuscript, as well as Jennifer DiMase and Naomi Kenner for help with drafts of this article.

Correspondence concerning this article should be addressed to Aude Oliva, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, NE20-463, 77 Massachusetts Avenue, Cambridge, MA 02139, or to Jeremy M. Wolfe, Visual Attention Laboratory, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115. E-mail: oliva@mit.edu or wolfe@search.bwh.harvard.edu

It is interesting to note that human observers need not be explicitly aware of the scene's spatial-configuration context. Chun and his colleagues have shown that repeated exposure to the same arrangement of random elements produces a form of learning that they call *contextual cuing* (Chun & Jiang, 1998, 1999; Chun & Nakayama, 2000). If configurations of distractor items in the display are repeated during an experiment, and if specific configurations predict target locations, then observers will have their attention guided to a target location even if they do not realize that they have viewed a configuration multiple times. There are similar, implicit effects of knowledge of target identity (e.g., priming of pop-out; Maljkovic & Nakayama, 1994, 1996).

In light of these studies, one might expect object search mechanisms to benefit from context familiarity whenever a robust association is found between an object and its location. However, there are limits on the ability of knowledge about the scene to guide the deployment of attention. In a series of experiments, Wolfe and colleagues (Wolfe, 2003; Wolfe et al., 2000, 2002) described *repeated search* tasks in which familiarity with a search display produced little or no improvement in search efficiency. In typical search tasks, observers look for a target amidst some number of distractors (the *set size*). The reaction time (RT) is the time required to find the target or to determine that the target is not present. The change in RT and/or accuracy with set size is a measure of the efficiency of search. In highly efficient searches, attention is deployed directly to the target location and, thus, the number of distractors does not influence RT. In inefficient search through items that can be identified without fixation, each additional distractor typically adds 20–40 ms to the time required to find a present target and 40–80 ms to the time required to determine that a target is not present (Wolfe, 1998).

In these standard search tasks, observers look for the same target throughout a block of trials. In Figure 1, the target is identified by the lowercase letter in the center of the display, and observers search through the uppercase letters to determine if that target is present or absent. Wolfe et al. (2000) compared standard visual search with repeated search conditions. The lower row of Figure 1 illustrates a repeated search condition. In repeated search, the search display remains unchanged. Even though the figure seems to show discrete trials, the letters {D G P Z C} would remain visible continuously during repeated search. Only the target identity would change from trial to trial. After a few trials, observers would be familiar with the display and would have committed it to memory. Nevertheless, Wolfe et al.'s (2000) surprising finding was that repeated search was no more efficient than standard search. In fact, repeated search did not differ in efficiency from a search task in which both the target identity and the search array changed on every trial. For letter search of the sort shown in Figure 1, target-present slopes were approximately 35 ms per item in repeated and unrepeated search conditions. Search slopes remained at this inefficient level even after 350 repetitions of search through a single, unchanging display.

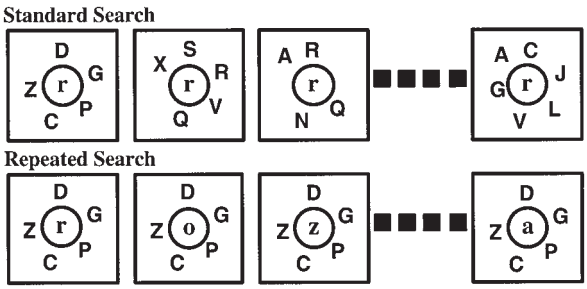

Figure 1. In these search tasks, the central, lowercase item identifies the target in a search through the surrounding uppercase items. In standard search, the search items change, but target identity remains the same. In repeated search, the search items remain unchanged, but the target identity changes from trial to trial.

This pattern of results has been reproduced with a wide variety of search tasks using many different search stimuli: letters, meaningless figures, objects, and so forth. The same pattern was observed when the target was identified by a direct physical match between target cue and search stimuli (e.g., both uppercase) and when the target was identified by an auditory cue (Wolfe, 2003).

This continuously inefficient visual search behavior was also found in search arrays for which observers could associate a meaningful scene context with objects. For example, Wolfe et al. (2002) had observers look at an essentially stable display (see Figure 2), searching for the target object that would "disintegrate" in front of their eyes. The observer would be viewing the scene on the left in the figure. The context, here defined as the configuration of the set of objects together with the scene background, remained identical for a few hundred trials. At the start of a trial, in this version, a cue appeared at the center of the display identifying the item that might be scrambled on that trial. In this example, the parrot has, indeed, been scrambled. The observer would respond affirmatively in this case, and the scrambled object would revert to its normal configuration, as on the left in Figure 2. The cue was always reliable: Observers simply had to check whether the target object was scrambled or not. Note that there is a small viewpoint shift between the two versions of the scene. This was inserted to mask the low-level visual transients that would otherwise cue the location of the scrambled object (Rensink, O'Regan, & Clark, 1997). Even though observers knew that the parrot was always on the floor at the bottom of the scene, they behaved as though they were searching inefficiently (17 ms/object) through the familiar scene on each trial. There was a modest but significant advantage for repeated over unrepeated search; however, RTs remained clearly dependent on object set size. Had attention gone immediately to the remembered location of the object, slopes should have dropped to zero.

The failure of efficiency to improve in repeated visual search is more surprising in light of the fact that memory search does become more efficient with practice. After several hundred trials, the ability to determine if a letter is in a memory set becomes independent of memory set size (e.g., Logan, 1992). In repeated searches of the sort shown in Figure 1, slopes dropped to zero when the letters were committed to memory but were not visible (Wolfe et al., 2000). This can be seen as a form of a *consistent mapping* memory search experiment in which one set of letters is mapped to the *target-present* response, and another is mapped to the *target-absent* response (Schneider & Shiffrin, 1977).

Given that observers in repeated visual search tasks know the locations of the objects in a display, why do they fail to use their memory to guide attention efficiently to the target? At least three hypotheses can be entertained:

Figure 2. Example of stimuli used in Wolfe, Oliva, Butcher, and Arsenio (2002, Experiment 7), with a known set size of six objects (radio, violin, hat, car, parrot, laptop). The target item is the scrambled parrot object on the floor in the right-side panel. From "An Unbinding Problem? The Disintegration of Visible, Previously Attended Objects Does Not Attract Attention," by J. M. Wolfe, A. Oliva, S. J. Butcher, and H. C. Arsenio, 2002, *Journal of Vision, 2,* p. 266. Copyright 2002 by the Association for Research in Vision and Ophthalmology. Reprinted with permission.

1. *Vision first:* Because repeated search tasks ask about the visual stimulus, it is not unreasonable for observers to rely on the visual display. This might be an implicit strategic decision, immune to observers' knowledge that the contents of the display will stay the same.

2. *Inefficient memory:* Memory search might, in itself, be inefficient even after hundreds of trials. In that case, repeated search would be inefficient whether it was repeated visual search or repeated memory search.

3. *Pragmatic choice:* Repeated memory search might be efficient (as the automatization literature suggests; e.g., Logan, 1992). Nevertheless, access to memory might be slow relative to visual search even when the display is well-learned. That is, it might be faster, overall, to perform an inefficient visual search through an unchanging display than to search efficiently through one's memory for the contents of that display. This might be especially true if observers feel the need to check the visual stimulus to confirm the results of a memory search (e.g., "I remember that my chair is next to the desk, but perhaps I should look there before I sit down").

In this article, we examine the relationship between visual search and memory search through stable, realistic scenes. For this purpose, we introduce panoramic search, which permits the concurrent assessment of visual search and memory search strategies. In panoramic search, the full stimulus is a wide-angle view of a visual display containing a background scene context (e.g., kitchen or living room) and a set of target objects, identified to the observer. We call the total number of objects in the full panoramic scene the *context* set size (e.g., five or eight objects). In panoramic search, we need to call this the context set size to distinguish it from the *visible* set size and the *hidden* set size. At any given moment, the observer sees only a subset of the full panoramic scene. The view shifts from one part of the scene to another, as if by a head movement (see Figure 4, presented later). As a consequence of the shift, different objects become hidden and visible. If the full scene contains a context search set of eight objects, a specific view might have three objects visible while five others are left hidden. The context set size is the full set of items that could be stored in memory. On each trial, observers are asked about the presence of an item. That item might be in the visible set, the hidden set, or it might be absent altogether. In a well-learned scene, hidden items can be found by memory search. Visible items can be correctly identified by either a visual search or a memory search.

This article describes four experiments. Experiment 1 used standard, static search stimuli. Observers confirmed the presence of an object in a scene with which they became familiar. The configuration of objects within a particular scene never changed. This experiment replicated the previous finding that repeated and unrepeated search tasks are similar in their efficiency. This served as a baseline for Experiments 2, 3, and 4, which used the new panoramic search paradigm.

In Experiment 2, observers responded *yes* if the target item was present in the visible portion of the display. They responded *no* if the item was absent from the scene altogether. Observers were never asked about hidden items in this experiment. Thus, observers could base responses on completely reliable visual or memory information.

In Experiment 3, observers responded *yes* if the target item was present in the visible portion of the display. They responded *no* if the item was absent from the scene altogether or if it was part of the currently hidden set. In this condition, observers were forced to use visual information, with the possible assistance of memory.

In Experiment 4, observers responded *yes* if the target item was present in the visible or hidden sets (i.e., "Is the target anywhere in the panoramic scene?"). They responded *no* if the item was absent. In this condition, observers were forced to use memory, with the potential for assistance of visual information.

General Method of Panoramic Search

The four experiments used the 16 stimulus objects shown in Figure 3. To avoid issues related to how a specific object is constrained by a scene context, we used objects that, for the most part, represented toys or items that could be located anywhere within a room (e.g., a cat). The objects were designed using a 3-D graphics software (Home Designer [Version 5.0];

Figure 3. The 16 objects used in Experiments 1–4.

Data Becker, Newton, MA). Objects subtended a visual angle between 3° and 4°, and the whole scene subtended an angle of 25° × 19° (screen size: 21 in. [53.34 cm], 1,024 × 768 pixels) at the 50-cm viewing distance.

The basic structure of panoramic search is illustrated in Figures 4 and 5. We simulated the normal situation of a visual world that extends beyond the current field of view. The 3-D scene depicted a large, human-scale room, with a kitchen area and a living room area. Only a portion of the room was shown to the observer at any given moment. The window of visibility, indicated by the dashed square in Figure 4, slid back and forth, mimicking the visual effects of a head movement. It moved by steps of 2.5° between trials and remained stationary during the actual search. The scene was composed of a set of objects that could be targets. We refer to the remainder of the display as the *background scene*. Figure 4 illustrates a trial in which 5 of these objects are visible targets (bear, cat, parrot, books, pot), while 3 others are hidden targets (car, violin, laptop). If, on the next trial, the viewing window shifted one step to the right, the bear and book would join the hidden set and the laptop at the far right would become part of the visible set. In this manner, the number of visible and hidden objects changed from trial to trial, but the scene and the configuration of objects in the whole panorama scene, known to the observer, did not.

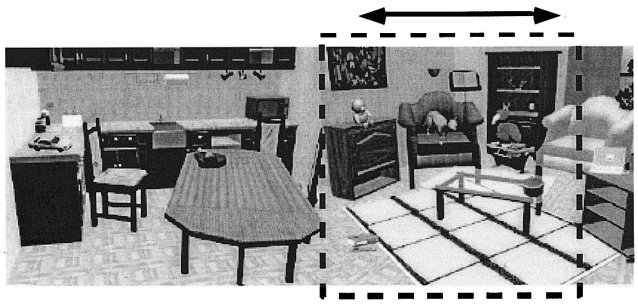

Figure 4. Illustration of the panorama procedure. The window of visibility (dashed square) slides back and forth, mimicking a head movement. It moves by steps of 2.5° between trials and remains stationary during a search.

In the panorama procedure, a virtual viewer was centered in the room, with eyes at a height of 5 ft, 8 in. From this position, we rendered 25 views of the scene. Each one was a 2.5° of rotation to the left or right of its neighbor. The views were played together as a motion picture. The movie was made of 48 frames, sequenced as if a virtual observer were looking from the right side of the panorama (e.g., the kitchen) to left side (e.g., the living room), then back to the right, and so forth. A loop of the panorama was composed of five forward steps (from kitchen to living room) and three backward steps (from living room to kitchen). Each of the eight steps in the sequence was used as a static search display. The view shifted smoothly by steps of 2.5°, for a total of 15° of difference, between two static search displays (see Figure 5). To furnish the room with target objects, we selected C objects randomly from among the 16 possible objects shown in Figure 3. These were pasted randomly into plausible locations in the room (see Experiment 2 for a more detailed description), with no overlap or occlusion among objects. These C objects constituted the context set size, corresponding to the total number of objects present in the panorama. When the movie stopped, the view of the scene displayed a visible set size of V objects. The remaining items constituted the hidden (or *memory*) set size of H objects. The relation between the C, V, and H set sizes was as follows:

$$C \text{ context objects} = V \text{ visible objects} + H \text{ hidden objects}.$$

Each pause in the sequence was considered a trial. On each trial, a tone was followed, after a 100-ms pause, by a visual probe, presented in the center of the screen. The probe was a black-and-white picture of 1 of the 16 objects. The status of a trial as a visible, hidden, or absent trial depended on whether the probe matched a currently visible object (e.g., the violin in Figure 5A), a hidden object (e.g., the car in Figure 5B), or an object that was not present in the room (e.g., the dinosaur). Observers were given auditory feedback. All of the experiments were run on a G4 Macintosh, using MATLAB (Version 5.1; The MathWorks, Needham, MA) with the Psychophysics Toolbox and the Video Toolbox extensions (Brainard, 1997; Pelli, 1997).

Once an observer became familiar with the entire scene, different rules could be devised to govern his or her responses to target probes:

1. In Experiment 1, the visual display consisted of a background scene, with repeated and unrepeated search arrays. The goal of

Figure 5. The five views in panoramic search at which the picture motion pauses: ABCDE (forward) and DCB (backward). This sequence shows eight objects. Each view, from A to E, is composed of a different number of visible and hidden objects.

this experiment was to replicate the pattern of performance observed with previous repeated search displays (Wolfe et al., 2002). In the repeated condition, the task was to respond to the presence of a visible object in an unchanging scene. In the unrepeated condition, observers searched for a new object in a novel scene on each trial.

2. In Experiment 2, using the panorama procedure, we had observers respond *yes* if the probe item was present. If it was present, it was always in the visible set. Observers responded *no* if the item was not present in the scene at all. Hidden items were never probed in this experiment. In this case, observers could perform either a visual search or a memory search, because both sources of information were perfectly reliable.

3. In Experiment 3, observers responded *yes* if the probe item was present in the visible set, *no* if it was in the hidden set, and *no* if it was not present in the scene at all. In this case, observers had to consult the visual stimulus, because a memory search would not necessarily produce the correct answer. Note that the same object would receive a positive response when it was visible and a negative response when it was hidden from view. Compared with Experiment 2, Experiment 3 could reveal interference of memory on visual search.

4. In Experiment 4, observers responded *yes* if the probe item was present in the visible set, *yes* if it was in the hidden set, and *no* if it was not present in the scene at all. In this case, observers had to consult memory because a visual search, by itself, would not necessarily have produced the correct answer.

Experiment 1: Repeated Search Versus Unrepeated Search

The goal of Experiment 1 was to confirm that the task we were using would produce the same pattern of results in standard and repeated visual search paradigms that has been found in previous research (Wolfe et al., 2000, 2002), Specifically, we wished to replicate the finding that repeated visual search does not become efficient with practice. The purpose was to compare the processes of searching for an object in a new configuration of objects to the process of confirming the presence of an object in an old and unchanged configuration. The repeated search task was similar to the letter-array task used by Wolfe et al. (2000; see Figure 1). The potential target items never vanished, changed form, or moved location during an experimental block. Therefore, this task could, in theory, be performed entirely in memory.

Method

Participants. Ten participants were paid $10 per hour to take part in the experiment. They all had at least 20/25 visual acuity (with correction, as needed), and all passed the Ishihara color screen test.

Procedure. Observers performed a visual search task (e.g., "Is the parrot in the scene?"). They were asked to answer as quickly and as accurately as possible whether a probe object was present in the picture by pressing the appropriate *yes* or *no* key on a keyboard. Each observer performed a repeated and an unrepeated version of the experiment. In both conditions, the background scene (either the kitchen or the living room) did not change for the entire block of trials. In the *repeated* condition, the objects and their locations also remained constant. Each trial was indicated by the appearance of a new probe image identifying the target at the center of the screen. In the *unrepeated* condition, a new set of objects was distributed at random but plausible positions in the scene on each trial. For *absent* trials, an object was selected at random from among the items that were not presented in the display (i.e., the remainder of the 16 total objects minus the set size). Set sizes were 2, 3, 5, or 8 objects. The set size variable was blocked because of the constraints of the repeated condition.

A tone indicated the start of a trial and was followed by a central probe image (see Figure 6). The probe was a black-and-white version of one of

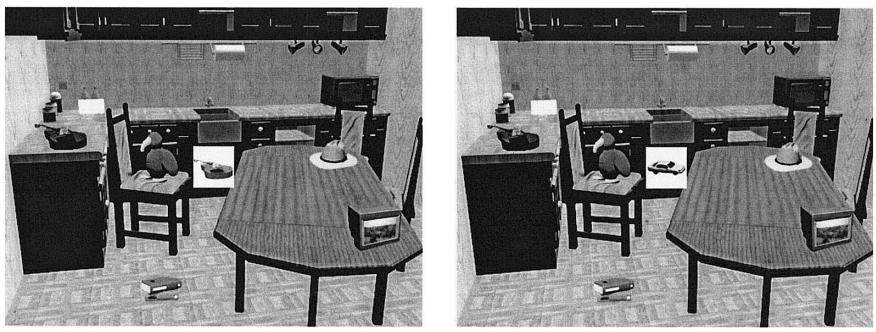

Figure 6. Two frames of sample stimuli for the repeated search condition, with a set size of 5 objects, in Experiment 1. Participants viewed a static display of 2, 3, 5, or 8 objects. From one trial to the other, the entire image remained static on the screen and a new probe appeared in the center (a violin in the left panel, a car in the right panel).

the 16 objects. It remained visible until the observer had pressed the *yes* or the *no* key. The probed item was present on 50% of trials. As an illustration of the *repeated* condition, for a set size of 5, one participant might see a parrot toy on the chair, a guitar on the counter, books on the floor, and a hat and a TV on the dining table (see Figure 6). Another participant would see a different set of 5 objects at different locations. The probe item never occluded a search item.

Each experimental block was composed of 200 experimental trials. Participants performed a total of 8 blocks (200 trials × 4 set sizes × 2 tasks = 1,600 experimental trials). Participants were randomly assigned to perform either the repeated or the unrepeated task first. For each task, order of blocks of fixed set size was randomized for each participant. Participants were allowed to pause between each block and were fully informed about the conditions. Thus, it was made clear that in the repeated condition, none of the stimuli would ever be removed, hidden, or changed during the course of a block. Participants performed 30 practice trials at the beginning of each block.

Results

RTs higher and lower than 3 standard deviations from the mean RT of all data were discarded from the analysis.[1] The error rate (including the discarded items) was significantly lower in the repeated condition (3.4%) than in the unrepeated condition (7.6%), $F(1, 9) = 8.3$, $p < .02$, and errors increased slightly with set size, $F(3, 27) = 4.6$, $p = .01$.

Main results are shown in Figure 7. A two-factor within-subject analysis of variance (ANOVA) performed on the correct mean RTs of present trials showed that observers responded faster in the repeated condition (626 ms) than in the unrepeated condition (834 ms), $F(1, 9) = 49.6$, $p < .0001$. We observed the usual set size effect, $F(3, 27) = 69.3$, $p < .0001$, but no significant interaction ($F \approx 1$), indicating that the slopes were not significantly different between the repeated and unrepeated conditions (29 and 33 ms/item, respectively). The slopes for target-absent trials were reliably different, with repeated slopes being significantly shallower than unrepeated slopes, $F(3, 27) = 5.1$, $p < .01$.

To evaluate the possible effect of memory, we looked at the results for epochs consisting of the first and second halves of the experiment (100 trials/epoch). Figure 8 clearly shows that practice with the search task did not alter the pattern of results in any sense.

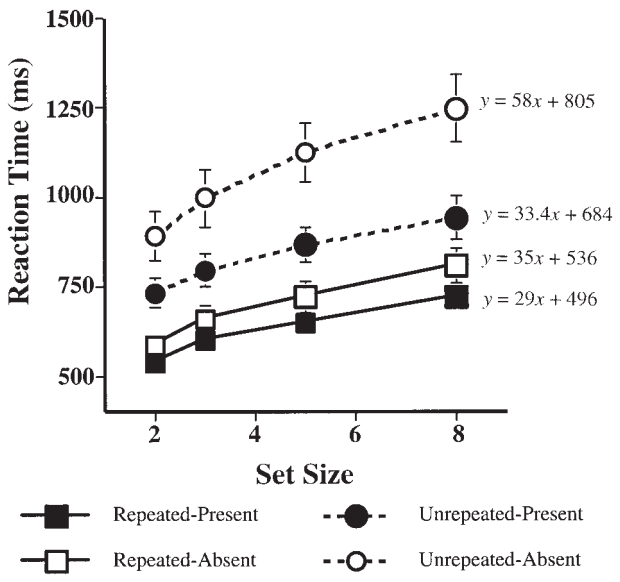

Figure 7. Mean reaction times as a function of set size for the repeated and unrepeated conditions of Experiment 1. Error bars represent plus or minus 1 standard error of the mean.

Discussion

The results of Experiment 1 replicate previous data of Wolfe et al. (2000, 2002). Although participants were faster overall in the repeated task, repeated search slopes remained quite inefficient (30 ms/item) even after 200 searches through the same, static scene. The RT difference between repeated and unrepeated search is not particularly mysterious in this experiment. The repeated condition was a *consistent mapping* task, in which the same objects were mapped to the same response keys throughout a block of trials. The unrepeated condition was an *inconsistent mapping* task in that the objects that were present were randomly chosen at each trial. Any specific probe object might have required a *yes* response on one trial and a *no* on the next. Inconsistent mapping tasks are known to be harder and slower than consistent mapping tasks (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Moreover, the unrepeated condition included some perceptual cost for loading a new scene on each trial.

The roughly equal slopes for target-present and target-absent trials in the repeated search condition are reminiscent of memory search of the sort studied by Sternberg (1966). Similar results have been seen in other visual search tasks in which observers had information about target location (Zelinsky, 1999).

The mystery in repeated search experiments is the failure of those searches to become efficient with practice (Wolfe et al., 2000, 2002). For example (see Figure 6), after 200 trials, observers certainly knew that the parrot was on the chair. Why, then, did they behave as if they were searching through the display de novo on each trial? To reiterate, the three hypotheses offered earlier are as follows:

1. *Vision first* argues that observers might always choose to perform a visual search, even if memory search is possible.

2. *Inefficient memory* suggests that memory search and visual search might be of similar efficiency in this task, giving no advantage to memory search.

3. *Pragmatic choice* holds that repeated memory search might be more efficient but might impose other costs that favor visual search.

In the remaining three experiments, we used our panoramic search method to distinguish among these three hypotheses. All three experiments had similar methods and displays but differed in their instructions. In Experiment 2, observers were free to choose

[1] A similar pattern of results and statistical significance was found when only RTs greater than 3,000 ms were discarded. The 3,000-ms cut-off was the same as the cut-off used in Wolfe et al. (2002), and results do not vary with more or less severe cut-offs.

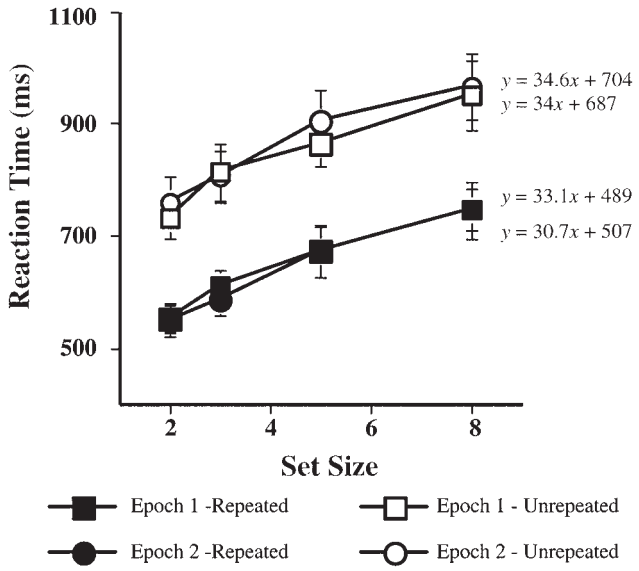

Figure 8. Reaction Time × Set Size (SS) functions for both 100-trial epochs, of object-present trials in Experiment 1. Error bars represent plus or minus 1 standard error of the mean.

between a visual search and a memory search. The results show that observers appeared to search the visual information. In Experiment 3, we forced a visual search and found that search remained fairly inefficient. In Experiment 4, we forced memory search on some trials, leaving observers free to choose between vision and memory on others. In this case, the memory search was efficient, and all searches became efficient with time. In addition, one can evaluate the repetition priming effect occurring between two consecutive trials with the same probe. Experiments 2, 3, and 4 allowed testing of the facilitation that occurs between successive probes that are both visible, both hidden, or both absent altogether from the panorama scene. We expected priming effects to be related to the search strategy used by participants. If observers are performing a visual search, they should benefit more when a previous visible target is repeated than when the object is absent from the scene (Experiments 2 and 3). However, if observers are searching in memory, one might expect facilitation effects to be higher for hidden than for visible targets (Experiment 4).

Taken together, the data seem to reject the inefficient memory hypothesis, because memory search was efficient in Experiment 4. The data also reject strong forms of vision first, because there were circumstances under which memory search appeared to be performed in the presence of the visual stimulus (again, in Experiment 4). Faced with visual stimuli, observers tend to prefer visual search even if memory search might be more efficient. However, in Experiment 4, it was possible to create situations under which observers could be induced to perform the more efficient memory search, even if the visual stimuli were present.

Experiment 2: Visual Task in Panoramic Search

Experiment 2 used the panoramic search procedure described in the General Method of Panoramic Search section. In this experiment, observers were asked, "Can you see the probe in the current view?" They responded *yes* to visible probes and *no* to absent probes. Hidden items were never probed, making this a consistent mapping task. For example, the parrot in Figure 5 might be the probe item if Views D or E were visible. It would not have been probed on trials showing Views A, B, or C. Note that this is a repeated search through the unchanging virtual scene. Observers were fully informed about the rules. Thus, they knew that hidden objects were never probed. Once they had learned the set, it would have been possible for the observers to base a response entirely on their memory for whether or not the probed item was in the memory set. It would not have been necessary to consult the scene at all.

The task was also a somewhat unusual version of an unrepeated search in the visible domain because the visible window moved to show a different part of the panoramic view for each trial. Thus, the task allowed us to determine whether observers prefer to use visual search or memory search. If, after observers had been familiarized with the scene, we observed that the visible set size was irrelevant (e.g., the number of objects in the window of visibility did not affect *target-present* and *target-absent* responses), this result would constitute evidence suggesting that observers adopted a memory search strategy, taking advantage of the consistent mapping of probe to response. In this case, the only set size of importance would be the size of the full context set (memory search through 5 or 8 objects). If RT increased with visible set size, this would suggest that observers adopted a visual search strategy. Context set size would be irrelevant because the hidden component of the set size would be irrelevant.

We can also examine these data for priming effects. Do RTs change when the probe is repeated on successive trials? In the present experiment, repeated probes would require the same answer. If an item was in the memory set, it would always be in the memory set. Note that this was different in later experiments, in which a probe on trial N might indicate a visible item, whereas the same probe might designate a now-hidden item on trial $N + 1$.

Method

This experiment used the panorama procedure described in the General Method of Panoramic Search section and the 16 objects shown in Figure 3. Sixteen participants were told to answer affirmatively if the probe object was currently visible in the scene. After their answer, the scene shifted toward the next window view, and a new probe image, selected at random from among the 16 possible objects, was presented on the screen. We used two context set sizes of 5 and 8 objects, in separate experimental blocks. The visible set size could be a subset of 2, 3, or 4 items for a context set size of 5 objects, and 2, 3, 4, 5, or 6 items for a context set size of 8 objects. Visible set size was a within-block factor. For each participant and each block, a new set of context objects was selected from among the 16 items and pasted at random among predetermined plausible locations (table, chair, counter and floor surfaces). The random matching of object and location for each new block controlled for any consistency effects that could have incidentally arisen between an object and its background.

The experiment was composed of three steps: a demonstration phase, a practice phase, and an experimental phase. During the demonstration, participants were familiarized with the panorama sequence and the context objects selected for that block (5 or 8 objects). As the window of the panorama moved, the experimenter pointed to the context objects on the screen to indicate which objects would be the subject of search and to inform the observer that other, background objects would not be part of the search task (e.g., table, armchair). The instructions made it clear that the 5

or 8 critical objects would remain the same and would remain in fixed locations for the entire experiment. Observers were told to answer affirmatively if the probe object indicated an object that they could see and negatively if the object was not in the room. They were explicitly told that they would never be asked about a hidden object that was present somewhere else in the room. Target-present and target-absent conditions were therefore consistently mapped. Observers performed a total of 128 practice trials, split into three epochs. After each practice epoch, observers' memory for the full context set of objects was tested. They were required to be 100% correct on the second and the third practice sets to ensure that they perfectly remembered the objects (this was not hard). The experiment was composed of 480 experimental trials (or 60 panorama loops). Each participant performed one experimental block per context set size (5, 8), counterbalanced. In total, they performed 960 experimental trials. Target-present trials and target-absent trials were determined in a random process (50% each).

Results

RTs higher and lower than 3 standard deviations from the mean RT of all data were removed from the analysis. Results and significance did not vary when a fixed RT cut-off was applied (e.g., RT > 3,000 ms). Altogether, error rates averaged 4.50%, and they did not differ across visible set size conditions (Fs < 1) and context set size (5 objects = 4.45%, 8 objects = 4.50%).

The results described below concern the experimental trials, run after observers had performed 128 practice trials, had shown that they were very familiar with the scene, and had memorized the objects. As noted earlier, the design of Experiment 2 permitted observers to search reliably through either memory or the visual display. Memory search can be assessed by examination of the effect of varying the context set size (the number of total objects in the room). Visual search can be assessed by examination of RT as a function of the visible set size (number of objects in the current view). The results suggest that vision wins.

To assess the relative impact of memory and visual factors, we performed an ANOVA on the correct mean RTs with context and visible set size as factors. To balance the design in the analysis and compare the same visible set sizes, we limited analysis of visible set size to those set sizes that were common to both contexts (visible set sizes 2, 3, and 4; context set size 5 lacked visible set sizes of 5 and 6). There was no significant effect of context set size ($F < 1$). Average RTs in panoramas of 5 or 8 objects were 563 ms and 567 ms, respectively, for target-present trials, and 598 ms and 588 ms, respectively, for target absent trials.

Figure 9 shows the effect of the visible set size, which was highly significant, $F(2, 30) = 21.7$, $p < .0001$. As observed in Experiment 1, search efficiency did not differ between target-present and target-absent sets. Participants searched at a rate of about 20 ms per visible object in both cases. This is interesting because target-absent trials typically have steeper slopes in visual search experiments. The result suggests again that, although repeated search may be dependent on the visible set size, the nature of that dependency is not identical to that found with standard visual search. There was only a modest effect of extended practice with the task and the specific objects. Observers were searching at a rate of 26 ms per visible object during the first epoch (120 trials), and they were still searching at a rate of 17 ms per visible object during the last epoch of the experiment.

The Visible Set Size × Context Set Size interaction was significant, $F(2, 30) = 13.28$, $p < .0001$, because, as shown in Figure 10, the RT × Visible SS functions differed substantially for panoramas of 5 and 8 objects. It is interesting to note that the visible slope was much steeper with 5 objects in the panorama (28.5 ms/visible object) than it was with 8 objects (11.7 ms/visible object). This also produced a significant effect of context set size on intercepts, $t(15) = 3.86$, $p < .01$, even though, as noted, the mean RTs did not differ.

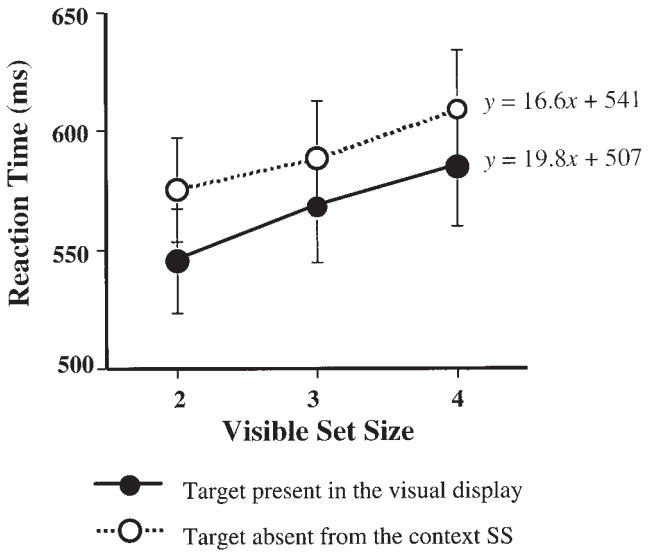

Figure 9. Mean reaction times as a function of visible set size (SS) for the target-present and target-absent conditions of Experiment 2. Error bars represent plus or minus 1 standard error of the mean.

The steeper slope in the context set size 5 appears to be a priming effect, as illustrated in Table 1. As a consequence of the panorama design, the probability of a probe being repeated in two consecutive trials (trial $N - 1$ and trial N) was greater for context set size 5 than for context set size 8. The probability of repetition was greatest when the visible set size was 2 and the context set size was 5: 27% of those trials were preceded by the same present probe versus 9–13% in all other conditions. If there was substantial repetition priming in this task, then the greater rates of repetition could have sped responses to set size 2 in context set size 5 (see Figure 10).

Table 1 shows that there was a substantial repetition priming effect in this experiment. Priming was computed as the difference between the response given to a trial, N (e.g., parrot), when it was preceded at trial $N - 1$ by the same probe (e.g., parrot) or by a different probe (e.g., laptop). Priming rates can be computed for target-present and target-absent sets.

The benefit of responding twice in a row to an object absent from the panorama was similar for the two context set sizes (34 and 40 ms; see Table 1). This benefit represents the baseline perceptual priming benefit related to the advantage of processing the same visual probe on two successive trials. In comparison, priming on target-present trials was significantly greater for the smaller context set size (5 objects = 100 ms, 8 objects = 74 ms), $F(1, 15) = 6.72$, $p < .05$. Thus, it seems unlikely that the slope difference between the two context set sizes reflects a memory search effect in this experiment. It is more likely that it represents a priming effect.

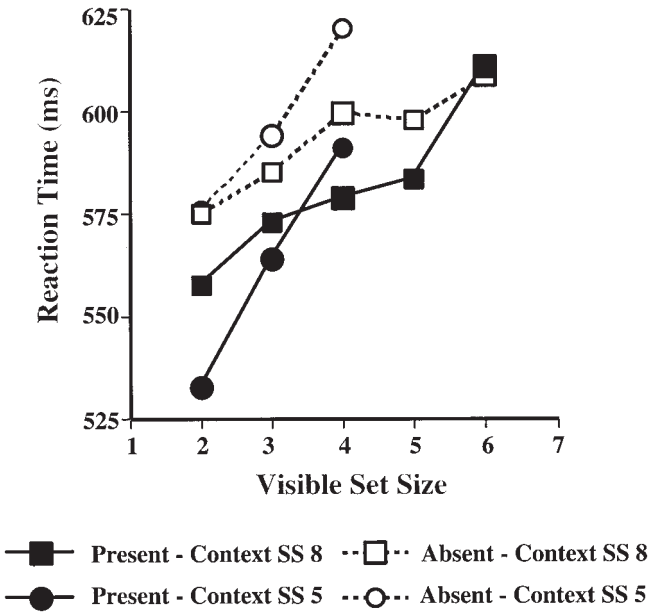

Figure 10. Reaction Time × Visible Set Size functions for context set sizes (SSs) of 5 and 8 in Experiment 2.

Discussion

The results of Experiment 2 suggest that, given the choice, observers rely more heavily on visual search than on memory search. Search performance was slightly more efficient in the panorama procedure (slopes of about 20 ms/item) than in the repeated search task of Experiment 1 (slopes of about 30 ms/item). The apparent benefit might be explained by the fact that observers were trained more heavily in panoramic search. The relatively steep panorama slopes indicate that participants did not make use of an automatized visual memory in this task. Had memory search been automatized and that memory used, then search time ought to have been independent of context set size, and there should have been no effect of the visible set size. Instead, there was a strong effect of visible set size. Faced with a well-learned task that permitted either visual or memory search, observers seem to have searched the set of visible objects rather than based their responses on memory.

Experiment 3: Forcing Visual Search

From the results of Experiment 2, one might argue that in the presence of a visible stimulus, participants neglect to use memory, either because visual search is still more efficient or because access to memory is slow. In any case, vision dominated the search strategy, suggesting that memory did not play any significant role. Experiment 3 tested the effect of memory on vision by using a task that put the two search strategies into competition. Here, we added a small but crucial change in the rule governing observers' responses: Observers responded *yes* if the probe item was in the visible set, as in Experiment 2. They responded *no* if the probe item was not visible, either because it was in the hidden set or because it was completely absent from the panoramic scene. In Experiment 3, therefore, memory for the context set was not a reliable guide to the correct answer as it was in Experiment 2. This rule forced the supremacy of visual information.

Method

We used the same panorama sequence and 16 objects as in Experiment 2. Fourteen observers (the same observers who would participate in Experiment 4) performed the task. Panoramas were designed with 5 or 8 total context objects. These were arranged in such a way as to have possible visible set sizes of 2, 3, 4, or 5 objects and possible hidden set sizes of 2, 3, 4, or 5 objects. Objects were assigned to random locations at the beginning of each block. The probe conditions (visible, hidden, or absent) were selected at random on each trial, with a rule that produced an average of 50% target-present trials.

This experiment was composed of four test blocks. Observers performed two blocks with 5 context objects and two blocks with 8 context objects. Each block was composed of a new selection of objects from among the 16 possible items, located at different positions within the panorama. After a demonstration phase, as in Experiment 2, participants completed a three-step learning process as follows: For the first step, they practiced for 40 trials with a single, sample panorama and were then presented with all 16 objects at once on the screen (as shown in Figure 1). They were asked to indicate the 5 or 8 context objects that belonged to the panorama. They then performed this task a second and a third time. They had to perfectly identify the context objects before entering the experimental phase and performing four sets of 480 trials each with a different panorama (1,920 total experimental trials). Experiments 3 and 4 were run together on the same set of observers. Each experiment took about 2 hours, and observers were tested in four sessions of 1 hour, spread over 2 days. Half of the observers performed Experiment 3 first, and the other half began with Experiment 4.

Results

Outlier RTs (±3 standard deviations from the overall mean RTs) were removed from the analysis. One participant did not finish the experiment. For the remaining 13 observers, errors averaged 4.45% for visible trials, 6.30% for hidden trials, and 1.20% for the absent set (see Table 2 for details).

Figure 11 shows RT as a function of visible set size for the three types of probe items: present in the visible set (*yes*), present in the hidden set (*no*), and absent from the entire scene (*no*). Context set sizes 5 and 8 are plotted separately. Table 2 gives the slopes, intercepts, and errors for these six functions. Figure 12 shows the same data plotted as a function of the context set size (5 or 8).

Table 1
Mean Reaction Times (in Milliseconds) for Trials N as a Function of Target (Same or Different) on Trials N − 1 in Experiment 2

Target and context SS	Reaction time		Priming rate (*SEM*)
	N same	*N* different	
Present			
5	477	577	100 (13)
8	513	587	74 (15)
Absent			
5	564	598	34 (24)
8	554	594	40 (17)

Note. SS = set size.

Table 2
Performance in All Conditions of Experiment 2

Target/response and context SS	Slope	Intercept	Error (%)
Visible/*yes*			
5	19.1	564	4.7
8	15.8	583	4.2
Hidden/*no*			
5	40.4	603	6.2
8	25.7	650	6.4
Absent/*no*			
5	14.4	587	1.2
8	12.4	580	1.3

Note. SS = set size.

Recall that the goal in this experiment was to force the observers to perform a visual search—they were told to answer positively to visual objects and negatively to hidden objects. The results suggest that observers did perform a visual search as expected, because all classes of responses were dependent on the visible set size. The dependence of the *yes* responses on visible set size is not surprising. Observers were asked to determine if a probe was in view. They searched and determined that it was. When the probe item might have been visible but happened to be hidden, observers behaved as observers typically behave on target-absent trials, producing slopes that were about twice as steep as the target-present slopes.

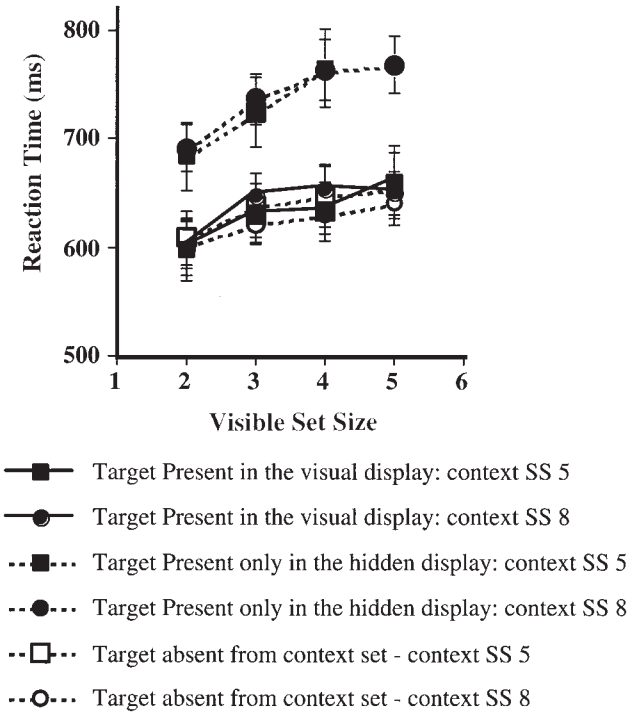

Figure 11. Mean reaction times as a function of visible set size (SS) for all conditions of Experiment 3. Error bars represent plus or minus 1 standard error of the mean.

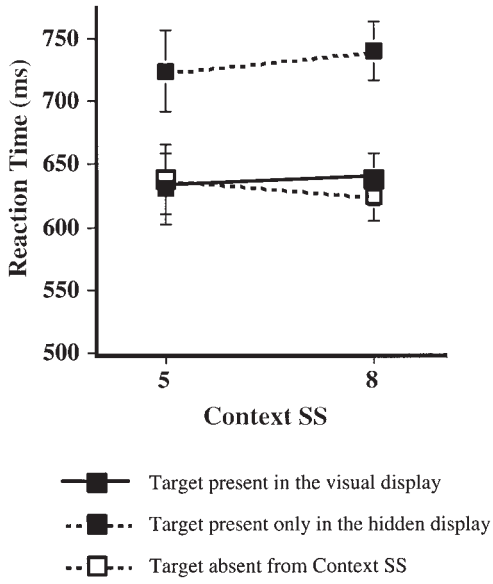

Figure 12. Mean reaction times as a function of context set size (SS) in Experiment 3. Error bars represent plus or minus 1 standard error of the mean.

The most interesting trials are those on which the probe item was absent—not in the scene at all. Because objects absent from the panorama were always the same objects, observers could determine that the probe item was not present without performing a visual search. Nevertheless, RTs on target-absent trials remained clearly dependent on the number of objects present in the current view. Table 2 shows that search was inefficient in all conditions of Experiment 3.

These observations are borne out statistically. An ANOVA performed on the *hit* trials (target-present trials on which the participant's response was correct), with the context set size and the visible set size as factors, showed a main effect of visible set size, $F(3, 36) = 34.7$, $p < .0001$, but no main effect of context set size ($F < 1$). The Context Set Size × Visible Set Size interaction was not significant, $F(3, 36) = 2.06$, $p = .12$. The same pattern of results was observed for the negative set answers (hidden and absent trials). Responses to hidden stimuli were slower and steeper than responses to visible or totally absent items (see Table 2), $F(2, 24) = 17.39$, $p < .0001$.

Although the results show that search was dependent on the visible set size even for the completely absent stimuli, this was not a simple, standard visual search. If observers had been relying entirely on visual search, one might have expected the absent RTs and slopes to be similar to the hidden RTs and slopes. This was not the case. Instead, the absent RTs and slopes were similar to the RTs for target-present trials in this experiment.

Table 3 summarizes the priming results in the visible, hidden, and absent cases: Priming rates in the three conditions were significantly different, $F(2, 24) = 3.45$, $p < .05$. Results are similar to those observed in Experiment 2: Visible present targets benefited more from repetition (87 ms in Experiment 2; 85 ms in Experiment 3) than did absent objects (30–40 ms). Priming for hidden targets (61 ms) fell between the absent and visible priming but did not differ significantly from either.

Table 3
Repetition Priming Effects in Experiment 3: Mean Reaction Times (in Milliseconds) for Trials N as a Function of Target (Same or Different) on Trials N − 1

Target	Response	Reaction time		Priming effect (*SEM*)
		N same	*N* different	
Visible	*Yes*	560	645	85 (6)
Hidden	*No*	676	737	61 (18)
Absent	*No*	598	634	36 (14)

Discussion

The goal of Experiment 3 was to force the supremacy of the visual information. The RT data suggest that observers were, indeed, searching inefficiently through the visual scene. This conclusion is bolstered by the finding of a stronger repetition priming benefit for visible objects.

The target-present and hidden RTs produced data consistent with a fairly typical visual search. Observers appear to have searched through the visible stimuli in some capacity-limited fashion. They terminated search either when they found the target or when they assured themselves that the target was not present. The cost of each added distractor was about twice as great for the hidden trials as it was for the target-present trials.

When the target was completely absent from the panorama, observers were faster and more accurate in responding *no* than they were when the target was merely hidden from view. There are two factors that may have made the completely absent trials more efficient than the hidden trials. First, responses to the completely absent trials were consistently mapped. If the cat was not in the room (not in the context set), it could never be a target in that block of trials. In contrast, a hidden item on one trial could become a visible item on the next, demanding a different response. This inconsistent mapping is known to make search slower and less efficient (Czerwinski, Lightfoot, & Shiffrin, 1992). Not only were the fully absent trials consistently mapped, they could also be done as memory searches, unlike the visible and hidden trials. If the cat was not in the context set, then all that the observer had to do was to remember that *cat* was consistently mapped to the *no* response. Extensive trials with consistently mapped memory search produce efficient memory search (Czerwinski et al., 1992).

Why was there any slope on the consistently absent trials? Why are these trials not dealt with through an automatized memory search? One possibility is that the relative inefficiency of the fully absent slopes represents the dark side of Chun and Jiang's (1998) contextual cuing effect. In a contextual cuing experiment, observers search for a target amid randomly placed distractors. The arrangement of stimuli changes from trial to trial. Unbeknownst to the observer, some types of displays are repeated many times during the experiment. Every time a specific "random" display appears, the target item is in the same place. Without awareness of the repetition, observers nevertheless learn to deploy attention to the target location more quickly.

In the present experiment, observers may have learned the 5 or 8 target locations that were present within the panorama. On each trial, they may have been compelled by the nature of the task to begin deploying attention from known target location to known target location in a visual search even on trials for which, in principle, they could have used memory. Because the number of these known target locations changed as the view of the panorama changed, the time required to search the known target locations was dependent on the visible set size. Because the background scene changed together with the visible objects (e.g., kitchen, dining table, living room area), the scene context itself may have cued the number of potential target locations for each view of the panorama. Thus, the results for the trials on which the probe item was not in the context set might represent a mixture of relatively inefficient absent-trial visual searches through the known target locations and relatively efficient, automatized memory search trials. If the slope for the target-present, visible trials is called *X*, then the hidden trials would be typical target-absent trials with a slope of about 2*X*. The fully absent trials might be a mixture of 2*X* (visual) search and 0-ms-per-item (memory) search—leading to an average slope of about 14 ms in this experiment. For present purposes, the most important point is that there was a strong visual component to the completely absent trials, even though memory alone would have sufficed for those trials.

Experiment 4: Encouraging Memory Search

Experiment 3 showed that observers could be coerced into performing a visual search even when a presumably more efficient memory search would have been adequate. In Experiment 4, search through the same panoramic displays was structured so as to encourage memory search. The experiment differed from the previous experiment only in the rule for responding to targets in the hidden set. In Experiment 3, those were considered to be target-absent trials. In Experiment 4, they were target-present trials. Observers were told to respond *yes* to indicate that the probe was present if it was present anywhere in the well-learned panorama—whether or not it was currently visible. They responded *no* only if the target probe was not present in the context set at all (target-absent trials). Once observers became familiar with the objects in the scene, this task could be done as a pure memory search. The visual status of the items was irrelevant to the response. Nevertheless, a strong version of the vision first hypothesis would predict that observers would search the visible scene even if memory information was reliable and might be accessed faster.

Method

As in Experiment 3, observers were tested with panorama displays containing 5 or 8 items. Each observer (the same group as in Experiment 3) performed two blocks of 480 trials at each context set size. Each block was composed of 40% target-absent trials, 30% visible target-present trials, and 30% hidden target-present trials. Observers were familiarized with each panoramic display prior to the experimental trials using the method described for Experiment 3. All other aspects of the experiment were the same as those for the previous experiment.

Results

Figure 13 shows the RT × Visible Set Size functions for visible, hidden, and absent items. Note that the *y*-axis is the same for this figure and for the comparable data from Experiment 3 shown in Figure 11. Errors averaged 2.8% for visible trials, 3.6% for hidden

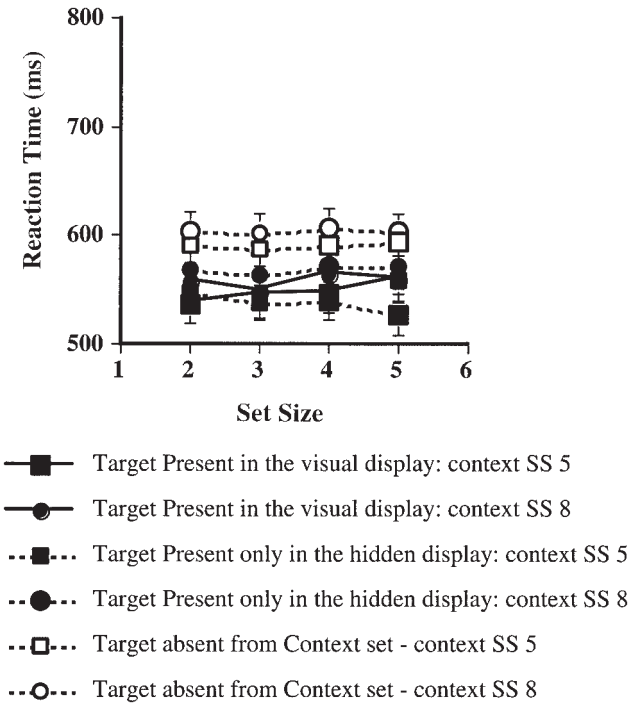

Figure 13. Mean reaction times as a function of set size (SS) for all conditions of Experiment 4. Error bars represent plus or minus 1 standard error of the mean.

trials, and 5.9% for absent trials. Mean RT as a function of context set size is shown in Figure 14, again with the same *y*-axis scale as for the comparable results for Experiment 3. RTs outside 3 standard deviations from the mean RT were discarded from analyses.

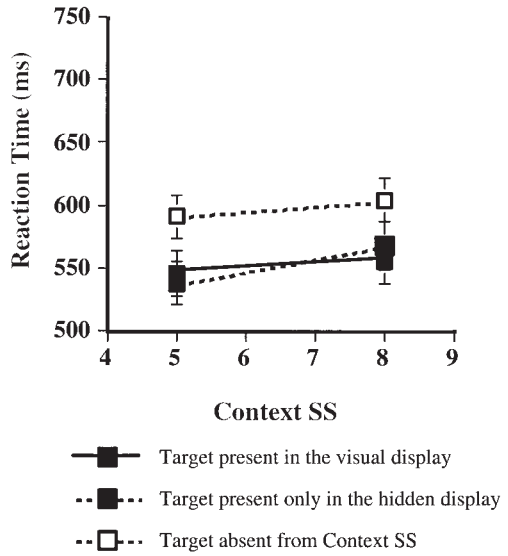

Figure 14. Mean reaction times as a function of context set size (SS) in Experiment 4. Error bars represent plus or minus 1 standard error of the mean.

The most salient aspect of these data is the great reduction in the effect of the visible set size on RT in comparison with Experiment 3. The Visible Set Size × Experiment (3, 4) interaction was significant, $F(3, 36) = 10.27$, $p < .0001$. On average, responses in the Experiment 4 task were faster than those in Experiment 3 (550 ms vs. 634 ms, respectively), for the visible set, $F(1, 12) = 33.70$, $p < .0001$ (see Figures 11 and 13).

An ANOVA performed on the positive responses (visible and hidden trials) with context set size as a factor (see Figure 14) indicated that there was a significant effect of the number of context objects, $F(1, 12) = 9.10$, $p < .05$, and a significant Context Set Size × Target Type interaction, $F(1, 12) = 9.3$, $p < .01$. As Figure 14 shows, context set size had a very modest effect on responses to visible targets (slope = 3.3 ms/object) and a significantly greater effect on hidden target-present responses (slope = 10.0 ms/object). Context set size also influenced RT in the target-absent set to a small but reliable extent (slope = 4.3 ms/item), $t(12) = 2.40$, $p < .05$. These results suggest some differences in strategy when observers responded to a visibly present object as compared with a hidden object.

The results of Experiment 4 are broadly consistent with an automatized memory search. It may be possible to see something of the transition from a visual search to a memory search by examining the changes in slope and intercept over the course of the experimental trials. Table 4 shows slopes and intercepts for Experiments 3 and 4 in the first and second halves of the 480-trial block. In Experiment 4, the visible target-present trials had a significant slope in the first half that vanished in the second. This was the only significant change, and it suggests that responses were dominated by the visual information in the earlier trials but not the later ones. Note that no such change was seen for essentially the same trials in Experiment 3.

Figure 15 shows a similar pattern of results in the slope of the RT × Visible Set Size functions when the block is divided into three, 160-trial epochs (for finer temporal resolution at a cost in statistical power). Absent and hidden trials had minimal slopes throughout the experiment. Visible trials show a reduction in slope from about a statistically reliable 10 ms per item at the beginning of the experiment, $t(12) = 2.85$, $p < .02$, to about 3 ms per item—not statistically different from zero, $t(12) < 1$.

One might imagine that the decrease in slope with practice reflects a change in the outcome of a "horse race" between visual

Table 4
Results of Visible, Hidden, and Absent Trials for Each 240-Trial Epoch in Experiment 3 and Experiment 4

Experiment and target	Response	Slope (Intercept) Epoch 1	Slope (Intercept) Epoch 2	Significant change in
3				
Visible	*Yes*	15.4 (567)	18.0 (577)	Neither
Hidden	*No*	27.0 (637)	30.0 (636)	Neither
Absent	*No*	15.0 (578)	8.0 (600)	Neither
4				
Visible	*Yes*	9.1 (517)	2.2[a] (544)	Slope
Hidden	*Yes*	−1.5[a] (560)	3.2[a] (544)	Neither
Absent	*No*	0.2[a] (601)	4.4[a] (579)	Neither

[a] Slope is not different from zero.

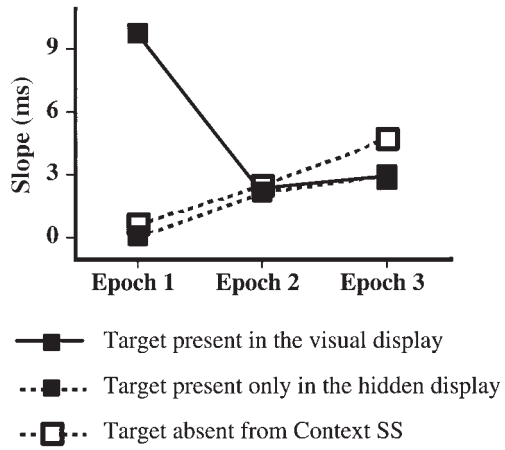

Figure 15. Slopes of the Reaction Time × Visible Set Size (SS) Function for each 160-trial epoch in a block of 480 trials in Experiment 4.

search and memory search. Perhaps the early trials were dominated by a relatively inefficient visual search because memory search was even slower. With practice, the memory search might have become faster and come to dominate the later trials. The data do not seem to support this account. Figure 16 shows RT × Visible Set Size functions for Epochs 1 and 3. In Epoch 1, when the visible target RTs were still dependent on the visible set size, RT at the smallest set size was actually shorter than it was in Epoch 3, when the RTs have become independent of the visible set size. Although there may have been a switch from dependence on visual information to dependence on memory, that switch was not driven by an accurate assessment of the relative speeds of these two types of search.

Table 5 shows priming results for Experiment 4. This experiment allowed us to compare the benefit of repetition when the two objects were visible, when they were both hidden, and in the mixed case (visible → hidden, hidden → visible). Priming effects in both mixed cases had a roughly comparable size. Overall, memory priming (hidden and mixed cases) was significantly different from pure visual priming, $t(12) = 3.80$, $p < .01$. In Experiment 4, repetition had a greater effect on memory search than on visual search.

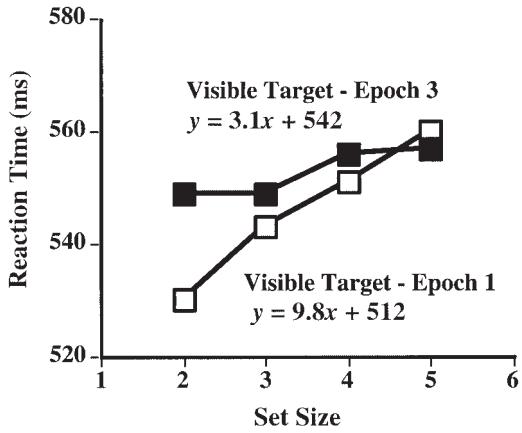

Figure 16. Reaction Time × Visible Set Size functions for Epochs 1 and 3 in Experiment 4.

Table 5
Repetition Priming Effects in Experiment 4: Mean Reaction Times (in Milliseconds) for Trials N as a Function of Target (Same or Different) on Trials N − 1

Prime–target trial	Response	Reaction time		Priming effect (*SEM*)
		N same	*N* different	
Visible	*Yes*	489	549	60 (16)
Hidden	*Yes*	457	569	112 (11)
Mixed				
H→V	*Yes*	463	563	100 (13)
V→H	*Yes*	474	562	88 (11)
Absent	*No*	536	602	66 (20)

Note. H = hidden; V = visible.

Discussion

The results of Experiment 4 are consistent with the hypothesis that observers made a pragmatic choice (probably an implicit choice) to use memory rather than vision in this task. Initially, the responses to visible targets showed evidence of a dependence on the visible set size, as if observers were searching through the items in front of them in the manner of observers performing other repeated search tasks (e.g., the present Experiment 1). Over time, however, that dependence vanished. It is very unlikely that this reflects the development of efficient visual search through the display. We have done many versions of repeated visual search, and we have never seen visual search slopes become as shallow as those shown in Figure 15. For comparison, in the last epoch of Experiment 3, the visible target slope was 18 ms per item. The comparable visible slope became flat in Experiment 4.

Search was efficient for all three types of target probe (visible, hidden, absent) by the last third of Experiment 4. This finding rules out two of the three hypotheses that were proposed at the start of this article. The inefficient memory hypothesis is falsified by the efficient search when the probe was either a hidden or an absent item. The search became "automatic," mirroring the results of other well-learned, consistently mapped memory searches. The vision first hypothesis is falsified by efficient search for visible-target probes. Given that visual searches of this sort do not become efficient after a few hundred trials, we conclude that the response to visible probes became independent of the visible set size because observers chose to rely on a reliable, consistently mapped memory search rather than a variably mapped visual search. This strategy is reflected in the priming benefit. When observers turned to memory in Experiment 4, hidden trials benefited much more from identical twins of trials than did visible objects (see Table 5).

One could propose that observers are searching a mental image of the entire panorama. Such an account could be considered to be a hybrid of visual search and memory search. At least for the hidden items, one could imagine that observers might scan an internal visual representation of the living room. For this account to serve as an explanation of the results of Experiment 4, it would be necessary to assume that imagery search was efficient under conditions in which comparable visual search was not (i.e., re-

peated search). That could be the case; however, the bulk of the data suggest that imagery and perception are isomorphic (Farah, 1989; Kosslyn, 1973). If a Task A takes longer than Task B when the stimuli are visual, the same relationship will hold when the stimuli are imagined. Thus, even if observers had a sustained image of the panorama and did not have to engage in the time-consuming act of generating a new image on each trial, it is still unlikely that they searched an internal image of the scene to perform this task.

General Discussion

How do observers search through familiar scenes? The present set of experiments investigated the interaction of memory and vision in observers who were performing a realistic search task: confirming the presence or absence of an object in a familiar environment. The results of the four experiments argue that observers make a pragmatic choice between vision and memory, with a strong bias toward performing visual search even in the presence of well-known visual stimuli. Repeated search experiments, including the present Experiment 1, have shown that observers will continue to perform something that looks like a visual search over hundreds of searches through the same, static display (Wolfe et al., 2000, 2002). Experiments 2–4 of the present article provided many opportunities to perform memory search. More often than not, observers declined to take advantage of those opportunities. As an illustration, consider the first epoch of Experiment 4. When the target object was present in the panorama but hidden from the current view, observers were forced to access memory, and they showed that they could perform an efficient memory search (slope near zero; see Table 4). However, when the same object was visible, search appeared to be dependent on the visible set size. Why is there this reluctance to use a reliable, efficient memory search? In the real world, the choice of memory search over visual search is not an idle one. Even if visual search entails some cost, it is probably wise to check the visual stimulus rather than to rely on memory alone. Although it is true that the world is a relatively stable place, it is not a perfectly stable place. That coffee cup might not be exactly where you remember it, so if it is present, it might be wise to look at it before reaching out to pick it up.

In this regard, the results of our panorama experiments are consistent with work by Hayhoe and collaborators (Hayhoe et al., 2002; Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Land & Hayhoe, 2001). These authors have studied visual behavior during natural behaviors like making sandwiches. They observed that search for objects in real scenes involves a great deal of visual checking of information, even for well-trained tasks. The Hayhoe experiments involved behaviors that take an extended period of time (seconds). Our panorama procedure showed evidence of reliance on visual stimuli even when observers made a quick decision (e.g., within 500 ms).

If it is wise to look at the coffee cup, would it not also be wise to use memory to guide attention and the eyes to the cup? Under some circumstances, this must be what happens (Henderson & Hollingworth, 1999). In the real world, when you search the well-learned contents of your pantry for the oregano, it seems very likely that your memory of its likely location guides your visual search. In the lab, the contextual cuing work of Chun and his colleagues makes a similar point (Chun, 2000; Chun & Jiang, 1998, 1999; Chun & Nakayama, 2000; Olson & Chun, 2002; see also Melcher 2001; Melcher & Kowler, 2001). As noted earlier, in contextual cuing, observers implicitly learn to direct attention to the target if it appears in a specific contextual display. Why do observers not do this reliably in repeated search experiments?

The answer may lie in the cost of coordinating memory search and visual search. To use memory to guide visual search, the observer must perform a memory search and then perform a visual search, guided by the memory information. With relatively small set sizes, such as those used here, performing this pair of searches with the added delay of transferring information from memory to vision may take longer than a relatively inefficient visual search, uninformed by memory. The relatively shallow slopes of an experiment like the present Experiment 3 might reflect a mixture of simple visual searches and memory-guided visual searches. The mixture could take place within a single trial if observers make use of a medium-term memory for the locations of a few objects within a specific context (kitchen or living room; see Melcher, 2001; Melcher & Kowler, 2001). Further research would be needed to determine the time course of memory guidance in these tasks.

O'Regan (1992) has argued that observers do not memorize the visual world because the world serves as an outside memory available for consultation. The experiments presented here put this idea to test. Unlike other search tasks, the panorama procedure allowed us to compare the relative contributions to visual search of that outside memory—the visual stimulus—and an observer's "inside memory." In keeping with O'Regan's clever observation, our observers proved to be inclined to use the outside memory in preference to the inside memory until strongly persuaded (in Experiment 4) that inside memory could be relied upon.

References

Biederman, I. (1987). Recognition by components: A theory of human image interpretation. *Psychological Review, 94,* 115–148.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143–177.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 443–446.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4,* 170–178.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10,* 360–365.

Chun, M. M., & Nakayma, K. (2000). On the functional role of implicit memory for the adaptive deployment of attention across scenes. *Visual Cognition, 7,* 65–81.

Czerwinski, M., Lightfoot, N., & Shiffrin, R. M. (1992). Automatization and training in visual search. *American Journal of Psychology, 105,* 271–315.

de Graef, P., Christiaens, D., d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research/Psychologische Forschung, 52,* 317–329.

Farah, M. J. (1989). Mechanisms of imagery–perception interaction. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 203–211.

Hayhoe, M. M., Ballard, D., Triesch, J., Shinoda, H., Aivar, P., & Sullivan,

B. (2002). Vision in natural and virtual environments. In *Proceedings of the Symposium on Eye Tracking Research & Applications* (pp. 7–13). (Available from the Association for Computing Machinery Digital Library Web site: http://portal.acm.org/dl.cfm?coll=portal&dl=ACM&CFID=24564967&CFTOKEN=24635279)

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision, 3,* 49–63.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50,* 243–271.

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,* 210–218.

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.

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.

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception & Psychophysics, 14,* 90–94.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London, Series B, 352,* 1231–1239.

Land, M. F., & Hayhoe, M. M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41,* 3559–3565.

Logan, G. (1992). Attention and preattention in theories of automaticity. *American Journal of Psychology, 105,* 317–339.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22,* 657–672.

Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: II. The role of position. *Perception & Psychophysics, 22,* 657–672.

Melcher, D. (2001, July 26). Persistence of visual memory for scenes. *Nature, 412,* 401.

Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research, 41,* 3597–3611.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42,* 145–175.

Oliva, A., Torralba, A., Castelhano, M. S., & Henderson, J. M. (2003). Top-down control of visual attention in object detection. In *Proceedings of the 2003 IEEE International Conference on Image Processing* (Vol. 1, pp. 253–256). New York: Wiley–IEEE Press.

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9,* 273–302.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as outside memory. *Canadian Journal of Psychology, 46,* 461–488.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3,* 519–526.

Pelli, D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5,* 195–200.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84,* 127–190.

Shinoda, H., Hayhoe, M. M., & Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research, 41,* 3535–3545.

Sternberg, S. (1966, August 5). High-speed scanning in human memory. *Science, 153,* 652–654.

Torralba, A. (2003). Modeling global scene factors in attention. *Journal of Optical Society of America A, 20,* 1407–1418.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). Hove, England: Psychology Press.

Wolfe, J. M. (2003). The level of attention: Mediating between the stimulus and perception. In L. Harris & M. Jenkin (Eds.), *Levels of perception* (pp.169–192). New York: Springer-Verlag.

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

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 693–716.

Wolfe, J. M., Oliva, A., Butcher, S. J., & Arsenio, H. C. (2002). An unbinding problem? The disintegration of visible, previously attended objects does not attract attention. *Journal of Vision, 2,* 256–271.

Zelinsky, G. J. (1999). Precuing target location in a variable set size "nonsearch" task: Dissociating search-based and interference-based explanations for set size effects. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 875–903.

Received January 24, 2003
Revision received May 26, 2004
Accepted June 10, 2004 ■