

Incidental visual memory for objects in scenes

Monica S. Castelhana and John M. Henderson

Department of Psychology and Cognitive Science Program, Michigan State University, East Lansing, MI, USA

Three experiments were conducted to investigate the existence of incidentally acquired, long-term, detailed visual memory for objects embedded in previously viewed scenes. Participants performed intentional memorization and incidental visual search learning tasks while viewing photographs of real-world scenes. A visual memory test for previously viewed objects from these scenes then followed. Participants were not aware that they would be tested on the scenes following incidental learning in the visual search task. In two types of memory tests for visually specific object information (token discrimination and mirror-image discrimination), performance following both the memorization and visual search conditions was reliably above chance. These results indicate that recent demonstrations of good visual memory during scene viewing are not due to intentional scene memorization. Instead, long-term visual representations are incidentally generated as a natural product of scene perception.

What is the nature of the representation that is created during ongoing, natural scene perception? Intuitively, it seems that the visual system generates a complete and highly detailed model of the external world. The perceptual experience of a stable and detailed visual world has led many vision researchers in the past to conclude that the visual representation formed for a scene is veridical and complete (McConkie & Rayner, 1976; Neisser, 1967). Such a detailed visual

Please address all correspondence to: Monica S. Castelhana, Department of Psychology, 129 Psychology Research Building, Michigan State University, East Lansing, MI 48824, USA. Email: monica@eyelab.msu.edu

This research was submitted in partial fulfilment of the requirements for the Masters of Arts degree by Monica Castelhana, who was supported by the MSU graduate school on behalf of the NSF IGERT program in Cognitive Science. The authors wish to thank Jeremy Athy, for his help with data collection, and the members of the masters committee, Erik Altmann, Thomas Carr, and Fred Dyer, for their guidance and comments. This work was also supported by the National Science Foundation (BCS-0094433 and ECS-9873531) and the Army Research Office (DAAD19-00-1-0519; the opinions expressed in this paper are those of the authors and do not necessarily represent the views of the Department of the Army or any other governmental organization. Reference to or citations of trade or corporate names do not constitute explicit or implied endorsement of those entities or their products by the author or the Department of the Army).

representation could serve as the basis for computations involving motor interaction with the environment, could underlie perceptual learning and priming, and might also serve as the basis for our visual phenomenology of a panoramic, full-colour, and visually detailed experience of the world. Such a representation, if it exists, would need to be created across multiple eye fixations so that the fine detail and colour information provided by the fovea during each fixation could be included in the composite global scene representation. Many of these theories were based on research in scene memory, which showed that observers could recognize a vast number of previously viewed photographs with high accuracy (Nickerson, 1965; Shepard, 1967; Standing, 1973; Standing, Conezio, & Haber, 1970).

Despite its intuitive appeal, the *composite scene representation* hypothesis has been challenged over the years by a number of investigators coming from a variety of empirical and theoretical perspectives (e.g., Ballard, 1992; Brooks, 1987; Churchland, Ramachandran, & Sejnowski, 1994; O'Regan, 1992; Wolfe, 1999). Studies of transsaccadic memory have shown that visual information is not fused across saccades (Bridgeman, Hendry, & Stark, 1975; Irwin, 1991; Irwin, Yantis, & Jonides, 1983; McConkie & Zola, 1979; Rayner & Pollatsek, 1983). The failure to fuse images across saccades has been replicated with simple patterns (Irwin et al., 1983), visual features of words (O'Regan & Levy-Schoen, 1983), contours of objects (Henderson, 1997), and features of scenes (Henderson & Hollingworth, 2003a). Without the ability to fuse visual information from one fixation to the next, it is improbable that the visual system creates and uses a composite point-by-point visual representation across fixations.

Studies on change blindness have also contested the existence of a complete, veridical visual representation of the outside world. Research on change detection has found that large discrepancies in a scene can go unnoticed, if changes are made during a saccade (Bridgeman et al., 1975; Grimes, 1996; Henderson & Hollingworth, 1999; Hollingworth & Henderson, 2002) or some other disruption, like a mudsplash (O'Regan, Rensink, & Clark, 1999), insertion of a blank screen (Rensink, O'Regan, & Clark, 1997), movie cut (Levin & Simons, 1997), or a temporary physical occlusion (Simons & Levin, 1998). In other words, participants show change blindness when the change is not accompanied by the usual transient motion cues. In many of these studies, participants showed change blindness even when they were actively looking for a change (e.g., Grimes, 1996; Henderson & Hollingworth, 1999; Hollingworth & Henderson, 2002; Rensink et al., 1997). Change blindness has thus been taken as evidence for the general absence of a composite scene representation (O'Regan, 1992; O'Regan & Noë, 2001; Rensink, 2000a, 2000b, 2002), and instead for a *localist-minimalist* approach to scene representation.

As a class, localist-minimalist theories assume that the visual representation is temporally and spatially limited to the object that is currently the focus of

attention (Rensink, 2000a, 2000b, 2002; Wolfe, 1999), with additional representation in visual short-term memory (VSTM) of up to three or four recently attended objects (Irwin, 1992a, 1992b; Irwin & Andrews, 1996). In this view, once an object is no longer attended or preserved in visual short-term memory, visual information about that object is no longer retained; instead, information about the object is stored as an abstract, semantically-based, nonvisual representation in long-term memory (LTM).

Although localist-minimalist theories of scene representation provided an initial account of the basic change blindness phenomenon, it makes several predictions about the preservation of visual detail that have not been supported by direct empirical testing (see Henderson & Hollingworth, 2003c). First, this class of theory predicts that changes to objects in scenes should always be very difficult to detect across saccades. Contrary to this prediction, good change detection is observed when care is taken to ensure that the changed scene region is fixated (and therefore attended) both before and after the change (Henderson & Hollingworth, 1999; Hollingworth & Henderson, 2002). Second, this view predicts that other behavioural measures of visual representation beyond explicit change detection should also show no evidence for the preservation of visual representations. Counter to this prediction, fixation duration, a covert behavioural measure of change detection, is elevated following unreported scene changes, suggesting that the information needed to detect the change has been generated and preserved (Henderson & Hollingworth, 2003c; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001). Third, localist-minimalist theories predict the absence of memory for the visual attributes of an object that is no longer held in VSTM during scene perception. In an online memory test for objects that had been fixated (and therefore attended) in scenes several seconds and multiple fixations before the test, Hollingworth and Henderson (2002) showed that participants were able to discriminate between a previously fixated object and a visual distractor from the same basic-level category (e.g., two kinds of radios). Exceptional memory performance was also observed when participants had to discriminate between the original view of the critical object and a 90° rotated view of that same object. Hollingworth and Henderson also tested long-term visual memory with type and token discrimination memory tests administered following presentation of the entire set of scenes. Participants were able to discriminate a viewed token from a distractor token differing only in visual detail with high accuracy despite multiple intervening scenes between study and test.

Contrary to localist-minimalist theories, then, recent change detection and visual memory results suggest that visual information acquired from an object during scene perception is represented in a relatively stable form in memory once attention has been withdrawn from that object and deployed elsewhere (Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002; see also Henderson & Hollingworth, 2003c; Hollingworth et al., 2001). To account for

change blindness under some conditions and not others, as well as the overwhelming evidence that iconic representations are not fused across saccades (Irwin, 1992a, 1992b; Irwin & Andrews, 1996; Pollatsek & Rayner, 1992), Hollingworth and Henderson (2002) proposed a *visual memory theory* of scene representation (see also Henderson & Hollingworth, 2003c). Like the localist-minimalist theories described above, visual memory theory assumes that attention to (or fixation on) an object is necessary to create and encode into VSTM a local visual representation. Unlike localist-minimalist theories, however, this theory posits that these visual representations leave a lingering long-term visual trace in memory. In this view, change blindness is caused not by the lack of visual representation, but rather by a failure of one (or more) of three memory-related processes: Initial encoding of a prechange representation, encoding of a postchange representation, and retrieval of the prechange representation following the change. When care is taken to ensure that a viewer has fixated and therefore has had an opportunity to encode a scene region prior to and after a change, as well as to retrieve the encoded prechange information following the change, then change detection (i.e., absence of change blindness) is observed (Henderson & Hollingworth, 1999; Hollingworth & Henderson, 2002; see also Henderson & Hollingworth, 2003b; Hollingworth, 2003; Hollingworth et al., 2001).

Results from the scene memory literature support the notion that visual details of objects are encoded in memory (Bahrick & Boucher, 1968; Mandler & Parker, 1976; Mandler & Ritchey, 1977). Although relatively simpler scenes were used in these earlier studies (line drawings of between six to nine objects in each), participants were able to distinguish between the target object and similar distractors of the same basic-level category (Bahrick & Boucher, 1968) and were able to recall different types of visual details (Mandler & Ritchey, 1977). Several studies also demonstrated that participants were able to recognize object type (Friedman, 1979; Goodman, 1980; Hock, Romanski, Galie, & Williams, 1978), visual details (Friedman, 1979; Mandler & Parker, 1976; Pezdek, Maki, Valencia-Laver, Whetstone, Stoeckert, & Dougherty, 1988; Pezdek, Whetstone, Reynolds, Askari, & Dougherty, 1989), and verbally recall and recognize object descriptions (Brewer & Treyns, 1981). More recent studies have found that the context of a target object is not only encoded incidentally over the course of training, but that this information is retained in LTM and used to improve performance in subsequent visual searches (Chun & Jiang, 1998, 2003).

The studies demonstrating visual memory appear to provide compelling evidence against localist-minimalist theories and for detailed long-term visual scene representation. These data, however, have been largely ignored or dismissed in the scene perception literature. A lingering concern is whether these results are due to memory processes normally recruited in the service of scene perception, or instead are the result of aberrant scene processing brought about by the use of viewing instructions that stress scene memorization. In prior

studies demonstrating visual scene memory, it is possible that participants employed strategies that called into play processes and mechanisms not typical of normal scene viewing (e.g., storing visual information as verbal descriptions). It could be that these long-term representations can be created when viewers engage in intentional memory encoding, but that these representations are not typically generated during natural scene perception tasks that do not require intentional memorization. For example, viewers are unlikely to engage in intentional scene memorization when locomoting through an environment, or searching through a scene. In this view, although natural dynamic scene perception does not involve the retention of detailed visual representations, atypical viewing strategies involving memorization can be engaged during scene perception when necessary. If this view is correct, then the evidence for good visual memory performance obtained in prior studies can be dismissed as artifactual and irrelevant to normal visual processing. Instead, the conclusion might be that localist-minimalist theories are correct and visual representations are in fact minimal, local, and transient during natural scene perception.

If the strategic memory encoding explanation for good visual memory is true, then evidence for long-term memory for the visual details of previously viewed objects should only be observed in intentional memorization tasks. Viewing tasks for which intentional memory encoding is unnecessary (i.e., an incidental memory task) should produce no lingering visual memory. Such a result would be consistent with localist-minimalist theories. On the other hand, if visual memory theory is correct and long-term visual representations are generated and stored as a natural product of visual scene perception, then evidence for the long-term preservation of visual detail in memory should be observed in both intentional and incidental memorization conditions. The present study was designed to test these competing predictions.

In the current study, the nature of the visual memory representation generated during scene perception was investigated by examining memory performance for visual information obtained from objects either intentionally or incidentally during scene viewing. In each of three experiments, participants viewed scenes while engaged in an intentional-learning memorization task or an incidental-learning visual search task. A memory test for a critical object in each scene was administered after both viewing tasks had been completed, although for the visual search task no memory test was anticipated by the participant during initial viewing. In the memorization task, participants were instructed to view the scenes in preparation for a difficult memory test that would require knowledge of details of specific objects. These instructions are similar to those given in other recent tests of visual memory (Hollingworth & Henderson, 2002). In the visual search task, participants were instructed to search for a specified target object in each scene, and were not told that they would receive a memory test for the objects in these scenes. In both cases, the memory test (given after all scenes had been viewed) always focused on the visual properties of a specific

critical object drawn from each scene. The critical objects for the search scenes were distractor objects in the scenes and were never the search targets. In Experiments 1 and 2, all participants took part in both the memorization and visual search tasks. In Experiment 1, the memory test involved discriminating between a previously seen critical object drawn from each scene and another foil object that was a different token of the same basic-level category type. In Experiment 2, participants had to discriminate between the previously viewed orientation of the critical object and a mirror-reversal distractor version of the same object. In Experiment 3, a between-participant version of Experiment 1 was conducted in which each participant was given only one of the two study conditions to ensure that there was no contamination from memorization to visual search. The main question in the three experiments was whether long-term visual memory would be observed for objects that were incidentally encoded during scene viewing.

EXPERIMENT 1

The purpose of Experiment 1 was to determine whether the visual properties of objects are stored incidentally in long-term memory as a natural consequence of scene viewing, or whether such representations are only stored when participants are intentionally memorizing a scene. Participants viewed digitized photographs of real-world scenes in each of two tasks: memorization and visual search. In the memorization task, participants were asked to examine each scene in preparation for a difficult memory test that would require memory for specific objects; in the visual search task, participants were asked to search for and locate a prespecified target object in each scene. After viewing all of the scenes in both tasks, a memory test was given. The memory test was expected for the memorization scenes but not for the search scenes. The memory test consisted of forced-choice token discrimination involving a critical test object taken from each viewed scene and a distractor object of the same basic-level category taken from a similar scene not viewed during the study session (see Figure 1). In the case of the search scenes, the critical test object was never the search target, but instead was another (distractor) object from the scene. Neither the scene from which the critical object was taken nor the task in which that scene had appeared were cued at the time of the memory test, although a small patch of scene appeared around the target and distractor objects. Thus, during the test the participant was presented with pairs of objects with minimal contextual aid.

If localist-minimalist theories of scene representation are correct in their claim that visual memory for scene information is an artifact of intentional memorization, then token discrimination should be above chance for critical objects taken from the memorization scenes and be at chance for critical objects taken from search scenes. If instead visual memory theory is correct and the generation of a visual representation is a natural consequence of scene

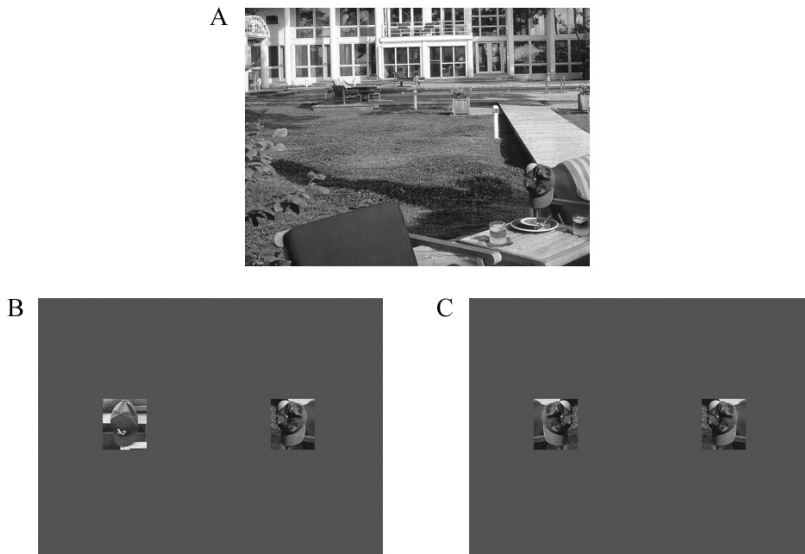


Figure 1. Example of the scene and test screens used in Experiments 1 and 2. (A) Example scene studied in one of two task instructions used in Experiments 1 and 2. For the visual search task, participants were asked to search for a swimsuit in this scene. (B) Example of test screen used in Experiment 1. The distractor was always a different object taken from the same basic-level category. (C) Test screen from Experiment 2 that depicts the target and the distractor, made by horizontally flipping the selected section of the scene.

perception, then above-chance memory performance should be observed for objects viewed in both the memorization and visual search tasks.

Method

Participants. Twenty Michigan State University students with normal vision participated in this experiment in exchange for credit in an introductory psychology course.

Stimuli. The critical stimuli were 30 photographed scenes of indoor and outdoor environments. Fifteen scenes were shown in each of the memorization and search tasks, with assignment of scene to task counterbalanced across participants. In the search task, the object that was the target of the search was never included in the scenes; instead, five noncritical filler scenes containing search targets were presented. These filler scenes allowed participants to successfully find a target on some trials, but were challenging enough (determined in a pilot study) to prevent them from guessing that most of the

search scenes contained no target. Each participant thus viewed 35 scenes in total, 15 memorization scenes and 20 search scenes (15 critical scenes that contained no target and 5 filler scenes that contained targets). The search target differed for each search scene. A word naming the search target was presented before the onset of each search scene. The word was printed in a 72-point Times New Roman font in black, also centred on a uniformly grey background. The words averaged $3.96^\circ \times 1.24^\circ$, with a range from 8.57° to 1.7° in width and 1.51° to 0.92° in height. In both study conditions, each trial began with a black fixation cross centred in a uniformly grey screen.

The memory test images were created by cutting out a square region containing the test object from each scene. Each test object was then matched with a foil object in its basic-level conceptual category (i.e., conceptual type), size, and orientation. Foils were taken from a separate picture set not viewed during the study session. The two objects (test object and foil) were placed on a uniformly grey background (see Figure 1B). Thirty memory test images were created this way in total, one for each critical scene viewed in the two tasks. The correct response was on the left for half the test images and on the right for the others. The test objects and foils were the same for the memorization and visual search conditions, although a given participant saw each test object and foil only once.

Apparatus. Participants' eye movements were monitored during scene viewing using a Generation 5.5 Stanford Research Institute Dual Purkinje Image Eyetracker (Crane, 1994; Crane & Steele, 1985), which has a resolution of $1'$ of arc and a linear output over the range of the visual display used. A bite-bar and forehead rests were used to maintain the participant's viewing position and distance. The position of the right eye was tracked, although viewing was binocular. Signals were sampled from the eyetracker using the polling mode of the Data Translations DT2802 analogue-to-digital converter, producing a sampling rate of better than 1000 Hz.

The monitor was placed 1.13 m from the participant. The screen subtended 15.20° horizontally and 11.93° vertically. The resolution was set at 800×600 pixels in 16-bit colour and the refresh rate was 100 Hz. The eyetracker and display monitor were interfaced with a microcomputer running a 90 MHz Pentium processor. The computer controlled the experiment and maintained a complete record of time and eye position values and participant responses over the course of each trial.

Procedure. Participants took part in two sessions, a learning session and a test session. In the learning session, the participants performed two tasks (memorization and visual search) while viewing the scenes. The tasks were presented in blocks and block order was counterbalanced across participants. In the test session, participants' memory for objects in scenes was tested using forced-choice recognition.

At the beginning of the experiment, the two tasks were described to participants as two separate experiments to discourage them from guessing that the visual search scenes would be included in the memory test. A bite-bar was prepared and the experimenter explained the eyetracker requirements. The first task was then explained in further detail and the eyetracker was calibrated. Each trial in both tasks began with an eyetracker calibration check; the eyetracker was recalibrated whenever the computer's estimate of fixation position (as shown on the screen by a white dot) was off by approximately ± 8 pixels. Participants began each trial in both tasks by fixating a central fixation cross; once the cross was fixated, the experimenter initiated the trial.

In the memorization task, participants were instructed to view each scene in preparation for a memory test, to be administered at the end of the session that would examine memory for specific objects in the scenes. Participants were given three practice trials before commencing the experimental trials. Each participant then viewed each of the 15 scenes for 10 s.

In the visual search task, participants were instructed to indicate when they had located the target object by holding fixation on the object and pressing the response button. At the beginning of a trial, a word naming the search target for the immediately following scene was presented in the centre of the computer screen for 2 s. The scene was then displayed for 10 s or until the participant pressed the response button, whichever came first. To encourage participants to exhaustively search the critical scenes for the entire 10 s presentation duration, the search target was never present in these scenes. The search target was present in five filler scenes. The upcoming memory test for the distractor objects in the search scenes was not mentioned. Each viewing task was explained separately just before it was to be performed, allowing for each participant to receive a 2–3 min break between the two tasks.

After both learning tasks were completed, participants were given the memory test. Participants were told that they would be receiving a memory test for objects that had appeared in the scenes that they had studied in the memorization task. These instructions coincided with the previous description of the experiment and were given to prevent any last minute strategic encoding of the previously viewed information (especially when the visual search task was administered in the second block). Participants were instructed that they would see a series of trials in which two objects would appear together, one on either side of a grey screen. They were informed that one of these objects had been cut out from one of the scenes that they had studied previously, while the other was from a photograph that they had not seen. They were instructed to respond by pressing the button that corresponded to the side of the screen on which the previously seen object appeared (left or right). If they were not sure which object they had seen before, they were told to guess. Participants were informed that they could take as long as they desired to make their decisions. Instructions for the memory test required about 2–3 min, ensuring that there was some delay

between presentation of the last scene in the second block of the learning session, and the first trial in the memory test.

At the end of the experiment, participants were debriefed. They were told that the test had included objects from the scenes in the visual search task, and were asked whether they had tried to memorize the scenes while looking for the target objects in the visual search task.

Results

Responses to debriefing question at the end of the experiment. All participants responded that they had concentrated on finding the target objects during the search task. No participant reported having known, guessed, or suspected that their memory for the visual search scenes would be tested.

Percentage of critical objects fixated. Memory performance on both immediate and long-term memory tests is dependent on fixation of the critical object at the time of memory encoding (Henderson & Hollingworth, 1999; Hollingworth & Henderson, 2002; Nelson & Loftus, 1980). Therefore, in the present experiment, memory accuracy scores in the memory test were calculated only for trials in which the critical object was fixated at least once during the learning session. Fixation on the critical object was defined as follows: At least one fixation with a minimal duration of 90 ms within a region defined as the smallest rectangle encompassing the bounding contour of the critical object. On this definition, the critical objects were fixated on 84% of the memorization trials and 66% of the visual search trials, $t(19) = 4.54, p < .01$. The average total fixation time on the critical objects that were fixated (sum of the durations of all fixations) was 929 ms during memorization and 860 ms during visual search, $t(19) = 0.91, p > .1$. Figure 2 shows all fixations across all participants in each task on an example scene.

Memory accuracy by task. Table 1 shows the mean accuracy for both tasks. One sample *t*-tests revealed that memory accuracy in both task conditions was significantly above chance (50%): memorization task, $t(19) = 13.652, p < .05$; visual search task, $t(19) = 5.998, p < .05$. To determine if there was any difference in the means between tasks, a paired sample, two-tailed *t*-test was computed and showed that performance in the memorization task was marginally higher than in the visual search task, $t(19) = 1.973, p = .063$.

Memory accuracy by task and block. The learning tasks were administered in blocks, with task order counterbalanced across participants. It is possible that task order had an effect on memory accuracy. First, memory accuracy could be affected by a recency effect, with test items studied in the second block remembered better than those studied in the first. Second, and more critically,



Figure 2. All fixations across all participants in Experiment 1 are displayed on an example scene (A) for the memorization task and (B) for the visual search task. The participants were searching for an apple when performing the visual search task on this scene.

TABLE 1
Mean memory accuracy percentages as a function of block and task
for Experiment 1

<i>Task</i>	<i>Block 1</i>	<i>Block 2</i>	<i>Task means</i>
Memory	80.3%	81.6%	81.0%
Visual search	64.8%	78.9%	71.8%
Block means	72.6%	80.2%	

performing the memorization task first could have placed the participants in a memorization mind-set for the following visual search task. This possibility would undermine the use of the visual search task as an incidental learning task.

To investigate the possibility that memory accuracy in the visual search task was contaminated by exposure to the prior memorization task, memory performance was analysed as a function of task and block in a mixed-design repeated-measures ANOVA. If memory for objects in search scenes was a consequence of contamination, memory performance for the objects from the search scenes should be above chance only when search was completed in the second block. For this analysis, block was treated as a within-participants factor, while task order (memory followed by visual search versus visual search followed by memory) was treated as a between-participants factor (see Table 1 for memory accuracy means by block and task). The analysis revealed that there were no main effects of block, $F(1, 18) = 3.023$, $MSE = 0.014$, $p > .1$, or task order, $F(1, 18) = 2.977$, $MSE = 0.0193$, $p > .1$, although there was a marginally significant interaction, $F(1, 18) = 4.308$, $MSE = 0.0193$, $p = .053$. The trend toward an interaction is consistent with the possibility that the above-chance memory performance for the visual search task was an artifact of contamination by the prior memorization task. To test this hypothesis directly, we examined whether memory performance for the visual search task differed from chance when visual search was presented in the first block. The results showed that in both blocks, memory accuracy was significantly above chance: Block 1, $t(9) = 4.504$, $p < .01$; Block 2, $t(9) = 4.942$, $p < .01$. Thus, above-chance memory performance for objects viewed during the visual search task was not the result of having engaged in a memorization task first.

Discussion

Localist-minimalist theories of scene perception posit that once attention has been withdrawn from a viewed object, visual information about that object is not retained in memory. Therefore, this class of theory predicts that memory performance for the visual attributes of previously viewed objects should be at chance. In this view, prior demonstrations of good visual memory for scene

information are dismissed as the product of unusual strategic memory encoding strategies. In the present study, visual search was used as an incidental memory task. The results clearly demonstrated that the visual detail needed to discriminate between two objects of the same basic-level category was stored in memory during scene viewing under these incidental-learning conditions. This visual memory was established in less than 1 s of total fixation time on the tested objects, suggesting that even a relatively short direct glimpse of an object is sufficient to incidentally establish a visual representation that endures over a relatively long period of time.

Experiment 1 demonstrated that participants were able to distinguish between an object incidentally encoded into memory during a visual search task and another perceptual token of the same basic-level category. Experiment 2 was designed to provide a further test of incidental long-term scene memory for visual information. In Experiment 2, the foils in the forced-choice memory test were created by mirror reflecting the critical objects seen in the learning session; in this way, they contained the identical visual and semantic information as the critical objects. Accurate memory performance in this case requires that object orientation be encoded in memory.

EXPERIMENT 2

Experiment 2 used an orientation discrimination task to explore further the nature of the visual information stored in long-term memory during extended scene viewing. Hollingworth and Henderson (2002) have shown in both immediate and long-term memory tests that object orientation is acquired and stored in memory during scene viewing when memory encoding is required for the task. If object orientation is encoded into memory as a natural consequence of scene perception, then accuracy in an orientation-discrimination task should be above chance following an incidental-learning scene viewing task. If instead the generation of detailed visual representations requires strategic encoding, then accuracy for orientation discrimination should be above chance for scenes viewed under memorization instructions but at chance for scenes viewed under visual search instructions. The orientation discrimination also controls for the scene patch surrounding the object. For instance, in Figure 1, the blue background behind the green baseball cap may have provided additional cues for the token discrimination test in Experiment 1. These cues are not available in the orientation discrimination test of Experiment 2.

Method

Participants. Twenty Michigan State University undergraduate students with normal vision received credit in an introductory psychology course or received \$7 for their participation.

Stimuli. The stimuli were the same as those used in Experiment 1, with the exception of the test images. Again, a square region containing the test object was taken from each critical scene. The region was copied and mirror-reversed using commercial graphics software. This manipulation maintained all visual features of the critical object (as well as the surrounding scene patch) except for orientation. The two alternatives were placed to the right and left of the centre of the screen, as was done in Experiment 1, with position of the target and foil counterbalanced across scenes. Figure 1C shows an example of a test image used in Experiment 2. Again, the same target objects and foils were used in the memorization and visual search conditions, although a given participant saw each only once.

Procedure. With the exception of the foils used in the memory test, the procedure for Experiment 2 was identical to Experiment 1.

Results

Responses to debriefing question at the end of the experiment. As in Experiment 1, no participant reported having known, guessed, or suspected at the time of the search task that their memory for the visual search scenes would be tested.

Percentage of critical objects fixated. As in Experiment 1, the memory accuracy analyses were based on trials in which the critical objects were fixated. The critical objects were fixated (as defined in Experiment 1) on 76% of the trials during memorization and 69% of the trials during visual search, $t(19) = 1.421$, $p > .1$. The average total fixation time on the critical objects was 896 ms in the memorization task and 764 ms in the visual search task, $t(19) = 1.625$, $p > .1$.

Memory accuracy by task. Memory accuracy as a function of task is shown in Table 2. Memory performance following both the memorization and visual search conditions was significantly above chance, $t(19) = 3.242$, $p < .05$ and $t(19) = 4.376$, $p < .05$, respectively. Performance following the memorization task was not significantly better than following the visual search task, $t(19) = 0.141$, $p > .1$, although as found in Experiment 1, there seemed to be a tendency for performance to be slightly higher for the memorization task.

Memory accuracy by task and block. Memory accuracy as a function of block and task is shown in Table 2. There were no significant main effects of task order, $F(1, 18) = 3.434$, $MSE = 0.0345$, $p = .08$, or block, $F(1, 18) = 3.049$, $MSE = 0.0189$, $p = .098$, and no interaction, $F(1, 18) = 0.784$, $MSE = 0.0189$, $p > .1$. Thus, in contrast to Experiment 1, there was no evidence in the current

TABLE 2
 Mean memory accuracy percentages as a function of block and task
 for Experiment 2

<i>Task</i>	<i>Block 1</i>	<i>Block 2</i>	<i>Task means</i>
Memory	55.6%	72.2%	63.9%
Visual search	60.8%	59.4%	60.1%
Block means	58.2%	65.8%	

experiment that memory accuracy for objects viewed during visual search was better when the visual search task followed the memorization task. As in Experiment 1, memory performance for objects viewed during visual search was significantly above chance in both blocks: Block 1, $t(9) = 6.877, p < .01$; Block 2, $t(9) = 2.551, p < .05$. Thus, above-chance memory performance for the visual search scenes was not the result of having engaged in a memorization task first.

Discussion

Although the accuracy scores for Experiment 2 were lower overall than in Experiment 1, the main results were replicated. Participants were able to perform a difficult visual memory test, in this case discriminating the orientation of a viewed object from its mirror reflection, both when they had intentionally tried to memorize the objects while viewing scenes and when they had searched through the scenes for a different (nonpresent) target object. As found in Experiment 1, no participant reported suspecting that they would be given a memory test for the visual search scenes. Furthermore, memory accuracy for objects that had appeared as distractors in the visual search task was equivalent whether this task was performed in the first or second block. Therefore, the possibility that memory performance in the visual search task was due to strategic encoding of visual information is remote.

EXPERIMENT 3

The purpose of Experiment 3 was to provide a final test of the strategic encoding hypothesis. Specifically, the experiment was designed to rule out the possibility that the memory results observed following visual search in the first two experiments was due to contamination of the visual search task by the memorization task. In Experiment 3, each participant performed only one of the two viewing tasks; the unassigned task for a given participant was never mentioned. Following presentation of all 30 critical scenes in the assigned memorization or search task, each participant took part in the token discrimination memory test used in Experiment 1. If the above-chance memory performance observed for

the search task in Experiments 1 and 2 was due to contamination of the search task by the memorization task, then in Experiment 3, visual object memory should be above chance under intentional encoding (memorization) instructions and should be at chance under the incidental encoding (visual search) instructions. If, however, detailed visual memory for objects results from both intentional and incidental memory encoding, then memory for the critical test objects should be above chance following both viewing conditions.

Method

Participants. Twenty Michigan State University undergraduate students with normal vision received \$7 for their participation.

Stimuli. The stimuli were identical to those used in Experiment 1.

Procedure. The procedure for Experiment 3 was identical to Experiment 1 with the following exceptions. Each participant took part in only the memorization or the visual search task. Participants assigned to the memorization task were informed that they would be memorizing the scenes, and participants in the visual search task were informed that they would be searching for target objects in the scenes. The specific instructions for each task were the same as those given in Experiments 1 and 2. Unlike Experiments 1 and 2, each participant viewed all 30 critical scenes in the assigned task (participants in the search task also saw the 5 filler scenes that contained a target). They also received the exact same set of targets and foils in the memory and search conditions. The order of scene presentation was randomized for each participant. For those participants assigned to the visual search task, the memory test was not mentioned until all search scenes had been viewed. The memory test instructions were the same for both groups and were those given in Experiment 1.

Results

Responses to debriefing question at the end of the experiment. As found in Experiments 1 and 2, no participant reported having known, guessed, or suspected at the time of the search task that their memory for the visual search scenes would be tested.

Percentage of critical objects fixated. The memory accuracy analyses were again based only on trials in which the critical object was fixated during initial scene viewing. The critical objects were fixated on 75% of the memorization trials and 76% of the visual search trials, $t(18) = -0.73, p > .1$. The average total fixation time on the critical objects was 939 ms during memorization and 796 ms during visual search, $t(18) = 1.718, p > .1$.

Memory accuracy by task. Memory accuracy following both the memorization task (78%) and the visual search task (75%) was significantly above chance, $t(9) = 10.582, p < .05$ and $t(9) = 9.643, p < .05$, respectively. Memory performance following the two viewing tasks did not differ reliably from each other, $t(18) = 0.818, p > .1$.

Discussion

The purpose of Experiment 3 was to determine whether the visual memory performance observed in Experiments 1 and 2 for distractor objects fixated during visual search was due to contamination from exposure to the memorization task. In Experiment 3, viewing task in the learning session was manipulated between participants so that those individuals assigned to the visual search task would have no knowledge that the study included a memory component. Nevertheless, participants assigned to the visual search task clearly demonstrated memory for the visual properties of the distractor objects they had viewed.

GENERAL DISCUSSION

The purpose of the present study was to test competing predictions generated by two theoretical perspectives on the nature of visual memory and scene representation. According to localist-minimalist theories, sustained visual scene representation is nonexistent (O'Regan & Noë, 2001; Rensink, 2000a, 2000b, 2002), or at best, is limited to high-level semantic representations of scene gist, spatial layout, and object identity (Irwin, 1992a, 1992b; Irwin & Andrews, 1996). In contrast, according to visual memory theory (Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002), a relatively detailed visual scene representation for attended scene information accumulates over time in long-term memory as a natural consequence of scene viewing. Recent investigations of scene memory have suggested that the visual attributes of objects are retained over both the short and long term, and that these preserved visual attributes can support both overt and covert change detection as well as direct memory test (Hollingworth & Henderson, 2002; Hollingworth et al., 2001), supporting visual memory theory. A possible criticism of these studies, however, is that participants may have engaged in intentional memory encoding strategies that are not characteristic of typical scene perception in order to complete the experimental task.

The present study contrasted the localist-minimalist and visual memory perspectives by investigating the degree to which detailed visual information is incidentally stored in memory during a scene viewing task that does not require intentional encoding. In three experiments, participants' memory for object detail was compared following an intentional scene memorization task and an incidental-learning visual search task. In the memorization task, participants

intentionally attempted to encode object detail. In the search task, participants searched for target objects and were unaware that their memory for distractor objects in the search scenes would later be tested. After completing the tasks, participants were given a difficult two-alternative forced-choice memory test involving previously viewed objects and foil objects distinguished from the viewed objects only by visual detail. Localist-minimalist theories predict either (1) that memory performance for objects appearing in both viewing tasks should be at chance because visual representations simply cannot be formed during scene viewing, or (2) that memory performance for objects appearing in the memorization task should be above chance due to strategic memory encoding, but that visual memory performance for objects appearing as distractors in the visual search task should be at chance because memorization strategies will not have been engaged. Visual memory theory, in contrast, predicts above-chance performance for objects viewed in both intentional and incidental learning tasks, because this theory holds that incidental storage of visual information from attended scene elements is a natural consequence of scene viewing.

In Experiment 1, two objects from the same basic-level category were presented in the memory test, one of which had been viewed in a scene during the study session. Accuracy was above chance for objects whether they had appeared in the visual search task or the memorization task. In Experiment 2, the same paradigm was used, except that the memory test involved distinguishing between a previously viewed object and its mirror reflection. Again, above chance performance was found whether the objects were viewed during the visual search task or the memorization task. In Experiment 3, viewing task was manipulated between participants to ensure that memory in the search task was not contaminated by exposure to the memorization instructions or task. The memory test was the same as in Experiment 1. Performance was again found to be above chance for both the intentional memorization and the incidental visual search conditions. The results from all three experiments are inconsistent with the hypothesis that visual memory for scene information is only established during intentional memorization of scene details. The evidence presented in this study thus demonstrates that visual information is encoded and stored in memory during scene viewing whether or not the viewer is intentionally trying to memorize that information. The results also suggest that prior demonstrations of good visual memory for scene information (Henderson & Hollingworth, 1999, 2003b; Hollingworth & Henderson, 2002; Hollingworth et al., 2001) were not the result of intentional memory encoding strategies.

The present study provides a relatively stringent test of visual memory for objects encoded during scene viewing, and hence should place a lower boundary on the degree of visual memory generated from viewed scenes. Some of the factors likely to contribute to low memory performance include the following. First, memory performance for the critical objects was based on a total of only 10 s of viewing time per scene and an average of less than 1 s of foveal fixation

time per tested object. Thus, the amount of time available for memory encoding was quite limited. Second, for both the memorization and search tasks, test objects were presented in the memory session without the contextual support that would be provided by the entire scene. Thus, the memory cues available for aiding retrieval were severely limited during test. Third, the retention interval in this study was relatively long (approximately 4–20 min) and the number of intervening objects relatively large between initial scene viewing and the object memory test, ruling out the use of short-term memory as the system supporting memory performance. All of these factors should have conspired to depress memory performance.

Finally, the total number of objects likely to have been encoded across the scenes was relatively large. A conservative estimate of the number of objects encoded per scene can be generated by considering the presentation durations of the scenes (10 s) and the average total fixation time per object, which can be estimated from the average fixation time on the critical objects. Assuming that only fixated objects were encoded (Hollingworth & Henderson, 2002; Nelson & Loftus, 1980), and that all objects were fixated for the same amount of time on average as the critical objects (929, 896, 939 ms in the memorization task of Experiments 1–3, respectively), we can estimate that about 10.7 objects were fixated per scene on average across experiments in the memorization task. In the search task, implicit encoding would have taken place for even more objects (12.4 per scene), because average fixation time per object was lower (860, 764, and 796 ms in Experiments 1–3, respectively). Using average fixation times for each experiment, and the fact that 15 scenes were presented in the memorization task and 20 (15 critical and 5 filler) in the search task, we estimate that 394 objects were fixated in Experiment 1 and 429 objects were fixated in Experiment 2. In Experiment 3, because participants saw all 30 scenes in only one condition or the other, we get two estimates: 373 objects for participants who memorized and 440 objects for participants who searched (including the five filler scenes that were not tested later). Given that our choice of critical test objects constituted a random sampling of all the objects in the scenes, memory performance for the test objects can be taken as an indication of the level of memory performance that untested objects would have shown if tested. We would therefore estimate that memory performance for token-level detail and left–right orientation for any of the hundreds of objects any given participant fixated in the scenes would have been at the levels we observed. Another way to think about this is that participants remembered $x\%$ of the objects viewed, with x estimated by the observed forced-choice performance y corrected for guessing ($x = [y - .5]/.5$). In terms of number of objects, in Experiment 1, we estimate that participants remembered the details of 244 objects (~16 objects per scene) in the memory condition and 173 in the search condition (~9 objects per scene). In Experiment 2, they remembered the orientation of 120 objects (~8 objects per scene) in the memory condition and 87 objects (~4 objects per scene) in the

search condition. Finally, in Experiment 3, they remembered the details of 178 objects (~6 objects per scene) in the memory condition and 219 (~6 objects per scene) in the search condition.

Given these considerations, the demonstrated existence of visual object memory poses a serious challenge to localist-minimalist theories, which assume that no visual representation is stored in memory from a viewed object once attention is allocated elsewhere in the scene.

Intentional versus incidental memory for objects in scenes

It is tempting to compare directly the memory results from the intentional and incidental learning conditions. In all three experiments, memory performance was statistically equivalent across the learning conditions, and although performance was numerically better in the intentional condition, the difference was not great, varying from a high of 9% in Experiment 1 to 4% in Experiment 2 and 1% in Experiment 3. The lack of large memory differences across conditions is striking given that participants were motivated to try to encode the critical objects into memory in the memorization condition, but simply had to reject the critical objects as targets in the search task. Furthermore, as noted above, fixation time on the critical objects was greater on average for the memorization than the search conditions in all three experiments, providing more time for memory encoding in the former case. Despite these advantages, performance for the critical objects was statistically equivalent following the memorization and search tasks. This result must be treated with some caution because the experiments were not designed to investigate this difference directly, and the result itself is based on a null effect. Nevertheless, the result is intriguing, and at the least suggests that memory performance in the implicit learning condition was approaching the ceiling level of performance set by the factors that limited performance in the intentional memorization condition.

Visual search does have memory

The results of the present study are also relevant to the issue of memory for distractor items in visual search. Horowitz and Wolfe (1998) presented a series of experiments apparently demonstrating that participants did not retain in memory information about previously viewed distractor items in a visual search task. In these studies, a target letter that had been a distractor in previous search trials was not found more rapidly than a target letter that was new to the search array. The lack of improvement in search efficiency suggested to Wolfe (1999) that attending to a distractor item on prior search trials did not establish a memory representation that could later be used to facilitate search for that item. In contrast to that conclusion, the data from the present study demonstrate that participants do have visual memory for previously viewed distractors in a visual

search task. These results are consistent with more recent results from the array search literature, which have shown in various ways that visual search does produce memory (Gibson, Li, Skow, Brown, & Cooke, 2000; Kristjánson, 2000; Peterson, Kramer, Wang, Irwin, & McCarley, 2001; Wolfe, Kemplen, & Dahlen, 2000). The present study contributes to this body of work by demonstrating that attention to a distractor object in search through a natural scene similarly leaves a lingering memory trace of the visual details of that object.

Level of analysis: Objects and scene patches

We have emphasized memory for objects because objects are the targets of fixations (Buswell, 1935; Yarbus, 1967), and in the memorization and visual search tasks used in this study the object was the task-relevant level of analysis. However, it could be argued that our tests of object memory were contaminated by the inclusion of a small amount of the surrounding scene in the memory test. We would argue that the main conclusions hold whether the level of analysis is objects or scene patches containing objects. In either case, visual details must be encoded to discriminate the target patches from the foils; in the case of orientation discrimination, the visual information in the target and foil patches was identical except for orientation, so discrimination required memory for the specific orientation of that patch. At both of these levels of analyses, the conclusion is that relatively detailed visual representations have been formed and retained in both intentional and incidental learning viewing tasks.

CONCLUSION

The present study was designed to test competing predictions from localist-minimalist and visual memory theories of scene representation. The differences between these views reflect a fundamental contrast in perspectives on the functional operation of the visual system. In the former view, the concept of visual representation is unnecessary at best and misguided at worst. In the latter view, visual representation is a necessary and natural consequence of vision. The results of the present study strongly suggest that visual information acquired from attended objects during scene viewing is stored in memory whether or not the viewer intends to remember that information, consistent with the latter view.

REFERENCES

- Bahrick, H. P., & Boucher, B. (1968). Retention of visual and verbal codes of the same stimuli. *Journal of Experimental Psychology*, *78*, 417–422.
- Ballard, D. (1996). On the function of visual representation. In K. Akins (Ed.), *Vancouver studies in cognitive science: Vol. 5. Perception* (pp. 111–131). New York: Oxford University Press.
- Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, *13*, 207–230.

- Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*, 719–722.
- Brooks, R. A. (1987). Intelligence without representation. *Artificial Intelligence*, *47*, 139–159.
- Buswell, G. T. (1935). *How people look at picture: A study of the psychology of perception in art*. Chicago: University of Chicago Press.
- Chun, M. M., & Jiang, Y. H. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.
- Chun, M. M., & Jiang, Y. H. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 224–234.
- Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch & J. L. Davis (Eds.), *Large-scale neuronal theories of the brain*. Cambridge, MA: MIT Press.
- Crane, H. D. (1994). The Purkinje image eyetracker, image stabilization, and related forms of stimulus manipulation. In D. H. Kelley (Ed.), *Visual science and engineering: Models and applications* (pp. 15–89). New York: Macel Dekker.
- Crane, H. D., & Steele, C. M. (1985). Generation-V dual-Purkinje-image eyetracker. *Applied Optics*, *24*, 527–537.
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355.
- Gibson, B. S., Li, L., Skow, E., Brown, K., & Cooke, L. (2000). Searching for one versus two identical targets: When visual search has memory. *Psychological Science*, *11*, 324–327.
- Goodman, G. (1980). Picture memory: How the action schema affects retention. *Cognitive Psychology*, *12*, 473–495.
- Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Vancouver studies in cognitive science: Vol. 5. Perception* (pp. 89–110). New York: Oxford University Press.
- Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object identification. *Psychological Science*, *8*, 51–55.
- Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, *10*, 438–443.
- Henderson, J. M., & Hollingworth, A. (2003a). Global transsaccadic change blindness during scene perception. *Psychological Science*, *14*, 493–497.
- Henderson, J. M., & Hollingworth, A. (2003b). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception and Psychophysics*, *65*, 58–71.
- Henderson, J. M., & Hollingworth, A. (2003c). Eye movements, visual memory and scene representation. In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects and scenes: Analytic and holistic processes*. New York: Oxford University Press.
- Hock, H. S., Romanski, L., Galie, A., & Williams, C. S. (1978). Real-world schemata and scene recognition in adults and children. *Memory and Cognition*, *6*, 423–431.
- Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes perception. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 388–403.
- 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., & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin and Review*, *8*, 761–768.
- Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *394*, 575–577.
- Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, *23*, 420–453.

- Irwin, D. E. (1992a). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 307–317.
- Irwin, D. E. (1992b). Visual memory within and across fixations. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 146–165). New York: Springer-Verlag.
- Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication*. Cambridge, MA: MIT Press.
- Irwin, D. E., Yantis, S., & Jonides, J. (1983). Evidence against visual integration of information across saccadic eye movements. *Perception and Psychophysics*, *34*, 35–46.
- Kristjánson, A. (2000). In search of remembrance: Evidence for memory in visual search. *Psychological Science*, *11*, 328–332.
- Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended object in motion pictures. *Psychonomic Bulletin and Review*, *4*, 501–506.
- Mandler, J. M., & Parker, R. E. (1976). Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 38–48.
- Mandler, J. M., & Ritchey, G. H. (1977). Long-term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 386–396.
- McConkie, G. W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes in reading* (pp. 137–162). Newark, DE: International Reading.
- McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception and Psychophysics*, *25*, 221–224.
- Neisser, U. (1967). *Cognitive psychology*. Appleton: Century-Crofts.
- Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 391–399.
- Nickerson, R. S. (1965). Short-term conceptual memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology*, *19*, 155–160.
- 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., & Levy-Shoen, A. (1983). Integrating visual information from successive fixations: Does trans-saccadic fusion exist? *Vision Research*, *23*, 765–768.
- O'Regan, J. K., & Noë, A. (2001). A sensorimotor theory of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*, 939–1031.
- O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change blindness as a result of “mud-splashes”. *Nature*, *398*, 34.
- 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.
- Pezdek, K., Maki, R., Valencia-Laver, D., Whetstone, T., Stoeckert, J., & Dougherty, T. (1988). Picture memory: Recognizing added and deleted details. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 468–476.
- Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 587–595.
- Pollatsek, A., & Rayner, K. (1992). What is integrated across fixations? In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 166–191). New York: Springer-Verlag.
- Rayner, K., & Pollatsek, A. (1983). Is visual information integrated across saccades? *Perception and Psychophysics*, *34*, 39–48.
- Rensink, R. A. (2000a). Seeing, sensing, and scrutinizing. *Vision Research*, *40*, 1469–1487.
- Rensink, R. A. (2000b). The dynamic representation of scenes. *Visual Cognition*, *7*(1–3), 17–42.
- Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, *53*, 245–277.

- 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.
- Shepard, R. N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Behavior*, *6*, 156–163.
- Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people in real-world interaction. *Psychonomic Bulletin and Review*, *5*, 644–649.
- Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, *25*, 207–222.
- Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, *19*(2), 73–74.
- Wolfe, J. M. (1999). Inattentional amnesia. In V. Coltheart (Ed.), *Fleeting memories* (pp. 71–94). Cambridge, MA: MIT Press.
- Wolfe, J. M., Klempe, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 693–716.
- Yarbus, A. I. (1967). *Eye movements and vision*. New York: Plenum Press.