Eye movements and picture processing during recognition

JOHN M. HENDERSON, CARRICK C. WILLIAMS,
MONICA S. CASTELHANO, and RICHARD J. FALK
Michigan State University, East Lansing, Michigan

Eye movements were monitored during a recognition memory test of previously studied pictures of full-color scenes. The test scenes were identical to the originals, had an object deleted from them, or included a new object substituted for an original object. In contrast to a prior report (Parker, 1978), we found no evidence that object deletions or substitutions could be recognized on the basis of information acquired from the visual periphery. Deletions were difficult to recognize peripherally, and the eyes were not attracted to them. Overall, the amplitude of the average saccade to the critical object during the memory test was less than 4.5° of visual angle in all conditions and averaged 4.1° across conditions. We conclude that information about object presence and identity in a scene is limited to a relatively small region around the current fixation point.

Visual acuity and color sensitivity are highest at the fixation point and fall off rapidly and continuously with visual eccentricity. Cognitive analysis and memory encoding are also most complete for the region of the scene at fixation and drop off monotonically from that point (Henderson & Hollingworth, 1999b; Nelson & Loftus, 1980). Because the human visual system is structured with a high-resolution central fovea and a lower resolution visual surround, human scene perception is active and dynamic: Three to four times each second, the viewer selects, via a saccadic eye movement, a specific region of the scene for priority in perceptual and cognitive processing (Buswell, 1935; Yarbus, 1967; for reviews, see Henderson & Hollingworth, 1998, and Rayner, 1998). An important question in active scene perception is the degree to which peripheral visual information is acquired and used to control the placement of the next fixation.

Parker (1978) reported a study of eye movements during scene viewing that involved analyzing the fixation patterns of viewers who were trying to find differences between previously learned and currently viewed scene images. The rationale for this method was that the distance at which various kinds of differences can be recognized is an indication of how far from fixation the information needed to note such differences is acquired and processed.

In Parker's (1978) study, participants were shown scenes that contained six objects. Each scene was presented in one form during an initial viewer-timed learning phase and then was shown 12 more times in a test phase (six *same* and six *different* trials). This sequence was repeated until the viewer had taken part in a session of 65 presentations of a given scene (5 learning and 60 test presentations). Viewers participated in seven such sessions, one for each scene. In the test presentations, the scenes were changed through manipulations to the objects, which were (1) deleted, (2) increased or decreased in size by 67%, (3) replaced with an object from a different conceptual category (type substitutions), or (4) replaced with an object from the same conceptual category (token substitutions).

Parker (1978) reported several important results. First, viewers typically fixated each of the six objects per scene, and the order in which they fixated the objects was quite consistent. Importantly, though, viewers were able to detect object deletions accurately and quickly, and on 85% of those trials, the deletions were noticed without the viewer's fixating the region of the scene that had contained the deleted object. Given that the objects were about 10° from each other in the displays, deletions were apparently detected in the periphery without fixation. Viewers also tended to "jump ahead" to a changed object in a scene; changed objects were fixated, on average, about 6.5% sooner than unchanged objects (1 fixation sooner out of an average of about 15.5 fixations). Because no background elements upon which viewers might fixate were present, these results suggested that viewers could detect changes and move to them sooner (or in the case of deletions, respond directly to them), given information acquired from

10° away. Parker concluded that "the results indicated that during recognition, information is encoded from a wide area and utilized both to direct additional eye fixations and to reach response decisions" (p. 284).

Parker (1978) adapted an interactive perceptual cycle model (Neisser, 1976) as a theoretical context for this work. In expanding this model to difference detection during scene recognition, Parker proposed that active picture recognition comprises four processes. First, information is acquired from a wide area of the test scene; second, the acquired information is compared with expectations generated from information stored in memory; third, a decision process assesses whether the evidence for a mismatch (or match) between acquired information and expectations is sufficient for a response; fourth, if the mismatch is not definitive enough to generate a response, a guidance process sends the eyes to a potential mismatch site.

The results from Parker (1978) are important because they provide strong evidence that information about the presence and categories of the specific objects in a scene is acquired and processed from quite a large region of the scene (with a radius of 10° centered at fixation). However, these results seem to be at odds with other evidence suggesting that the properties of objects are typically processed from a relatively small area of the scene around the current fixation point. For example, Nelson and Loftus (1980) recorded viewers' eye movements while they examined a set of scenes in preparation for a difficult memory test. Performance in the memory test was found to be a function of the nearest fixation to the tested object during learning, with good performance if the tested object had been fixated during learning and a rapid drop off to near-chance levels if the nearest fixation was beyond about 2° from the object subsequently tested.

Similar findings have been obtained in the *flicker paradigm*, in which participants are asked to detect and report scene changes across brief blank periods that alternate with two versions of the scene (Simons & Levin, 1997). The fact that viewers typically take a considerable amount of time to find changes in this paradigm suggests that there is severe restriction in the amount of information that can be acquired and processed during each fixation. Hollingworth, Schrock, and Henderson (2001) monitored eye position in the flicker paradigm and found that participants typically did not detect a scene change until after the changing object had appeared in foveal or near-foveal vision. Change detection in studies in which the change is effected on line in response to a saccadic eye movement have similarly shown that object changes are far more likely to be noticed if the changing object is fixated before and after the change (e.g., Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002; see Henderson & Hollingworth, 2003b, for a review).

The results of two additional studies suggest that semantic information about the objects appearing in a scene is acquired from a relatively small region around the current fixation point (De Graef, Christiaens, & d'Ydewalle,

1990; Henderson, Weeks, & Hollingworth, 1999). In both of these studies, participants examined scenes in which either semantically consistent or semantically anomalous objects had been embedded. If semantic information can be acquired and processed from the visual periphery, the eyes should be attracted to anomalous objects from relatively distant fixation sites (Loftus & Mackworth, 1978). Contrary to this prediction, saccades were not attracted to an anomaly in the periphery, nor were saccades to either type of object especially long. For example, Henderson et al. found that the average saccade length to an object was about 3.5° regardless of its semantic status.

In summary, although Parker's (1978) results are provocative, they appear anomalous in the context of other scene perception studies. Several design features of Parker's experiment are likely to have led to an overestimate of the degree to which object properties are peripherally processed in natural scene perception. First, the scenes used by Parker were schematic in format, consisting of an array of six discrete line-drawing objects against a blank background. The objects were large (judging from other measurements reported in the paper, they subtended about 7° on average) and were relatively distant from each other (10° center to center). Prior studies have shown that much smaller line drawings of objects presented in an uncluttered field can be identified 10° from fixation (Pollatsek, Rayner, & Collins, 1984). The sparseness of the scenes and the large distances between objects may also have led to saccade lengths that were inflated, as compared with those typically observed in more natural scene images (Henderson & Hollingworth, 1998, 1999a). In more complex images, pattern perception is limited by lateral masking as well as by visual eccentricity (Bouma, 1978). Second, Parker's experiment included only seven base scenes, and variants of each of these scenes were presented 65 times to each viewer. Given the simplicity of the scenes and the large number of repetitions, viewers may have been able to learn specific visual cues that would be sufficient to detect object differences in the periphery. Finally, the results were based on data from only 3 participants.

EXPERIMENT

In the present study, we followed Parker (1978) and used the recognition of changes to objects in memorized pictures as a diagnostic tool to investigate the role of eye movements in object processing during scene viewing. In contrast to Parker, the pictures were full-color images of complex scenes, and each image was tested only once. In an initial study phase, the participants were shown 33 full-color, meaningful, naturalistic scenes for a total viewing time of 20 sec each. We have previously demonstrated that very good memory representations of objects in scenes can be generated that are based on fewer than 20 sec of scene-viewing time (Castelhano & Henderson, 2001; Hollingworth & Henderson, 2002). Furthermore, we have shown that fewer than 20 sec of viewing during

learning leads to good memory performance for the specific critical objects and scenes used in this experiment, particularly when the critical object is fixated during learning (Hollingworth & Henderson, 2002). In the present study, to increase the memory representation of the scenes further, we distributed 20 sec of study time over two consecutive blocks of learning trials, taking advantage of spaced practice to enhance the initial memory trace of the scene.

In the subsequent test phase, each scene was shown a third time, and the participants were asked to determine whether the current scene was different from the studied version. The scenes in the test phase were identical to the learned scenes, had one object removed from them via a deletion, or had an object from a different basic-level conceptual category substituted for one of the objects in the learned scene. In the substitution condition, the original object was not semantically central in defining the meaning of its scene, and both the original and the replacement objects fit in their scene equally well.[1] The deletion and substitution conditions replicated the equivalent conditions used in Parker (1978), but with full-color complex scenes. The participants were asked to press a *yes* button as quickly as possible if and when they determined that the test scene was different from the original studied version and to press a *no* button if they believed that the test scene was the same as the studied version. Our primary dependent measures were the accuracy of recognizing that the test scene was different from the studied version, the time to make this determination, the amplitude of the saccade to the manipulated object when the manipulation was recognized following fixation, and the tendency to "jump ahead" to the manipulated object when the manipulation was noticed. If the model and data presented by Parker generalize to complex scenes, deletions should be noticed accurately and quickly in the visual periphery without fixation. Furthermore, substitutions (and to the extent they are fixated, deleted objects) should be noticed relatively distantly in the periphery, and the eyes should tend to jump ahead to substitutions with relatively long initial saccades. On the other hand, data from other scene perception studies lead to the expectation that object processing will be more restricted for peripheral objects in complex scenes (e.g., Henderson et al., 1999; Nelson & Loftus, 1980; see Henderson & Hollingworth, 1998, 1999a, for reviews).

Method

Participants. Twenty-four participants were recruited from undergraduate psychology classes or through an advertisement in the university newspaper at Michigan State University. Five participants were removed from the study for high false alarm rates in the memory test, and 1 was removed for mixing up the response buttons in the memory test. The data reported below were, therefore, based on the remaining 18 participants. The participants received either course credit or $7 remuneration for participating in the experiment.

Materials and Design. The materials for this experiment consisted of 33 realistic three-dimensional rendered color scenes (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002;

Hollingworth et al., 2001). An example scene is shown in Figure 1. Scenes subtended 15.69° horizontally $\times$ 11.76° vertically at a viewing distance of 1.13 m. A critical object was chosen in each scene, and manipulations to this object created the three experimental conditions. During the test phase of the experiment, the critical object was removed from the scene (*deletion* condition), replaced by another object of a different basic-level conceptual type (*substitution* condition), or remained unchanged (*identical* condition).

Apparatus. The stimuli were displayed at a resolution of 800×600 pixels $\times$ 32,768 colors on an NEC Multisync XE 17-in. monitor driven by a Hercules Dynamite Pro super video graphics adapter card operating at a refresh rate of 143.6 Hz.

Eye movements were monitored using a Generation 5.5 Fourward Technologies Dual Purkinje Image Eye Tracker (Crane, 1994; Crane & Steele, 1985), which has a resolution of 1 min of arc and a linear output over the visual display range used here. A bite bar and forehead rest maintained the participant's viewing position and distance. Viewing was binocular, although only the position of the right eye was tracked. Signals were sampled from the eye tracker, using the polling mode of the Date Translations DT2803 analog-to-digital converter, producing a sampling rate slightly greater than 1000 Hz.

Procedure. The participants began the experiment by reading the instructions for the study. A bite bar composed of dental compound was then fashioned to minimize head movements during the experiment. The experimenter orally described the eye-tracking equipment and restated the instructions. After any remaining questions were answered, the eye tracker was calibrated. Calibration consisted of having the participant fixate four markers at the top, bottom, left, and right sides of the display area. Calibration was checked by displaying six test positions and a fixation marker that indicated the computer's estimate of the current fixation position; calibration was considered successful when the fixation marker was within 0.10° of the test positions as they were fixated. Once calibration was completed, the experimental trials began.

The experiment consisted of two blocks of study trials (study phase) followed by a block of memory test trials (test phase). Prior to the study phase, the participants were instructed that they would see a sequence of scenes that they were to memorize for a later difficult memory test. Each of the 33 scenes was then presented for 10 sec in each of two blocks, for a total study time of 20 sec. The calibration screen was presented between each scene, and calibration was checked throughout each block. A short break was given between the two study phase blocks and between the study and the test phases. During the test phase, the participants were reminded that they would be shown each scene again and that they were to decide whether the current scene was identical to the original version or, instead, had been altered. The participants were informed that in the case of altered scenes, an object that had been present during the study session would be either deleted or replaced by a different kind of object. The participants indicated whether the scene was the same or different by pressing one of two response buttons on a button box. The scene was terminated when a button was pressed. The participants' eye movements were tracked throughout each trial. Following completion of the test phase, the participants were debriefed about the purpose of the experiment.

A given participant saw all 33 scenes, 11 in each condition (identical, deletion, and substitution). Assignment of scene to condition was counterbalanced so that each scene appeared equally often in each condition across participants. The order of scene presentation in each of the blocks of the study phase and in the test phase was determined randomly for each participant. The experiment lasted approximately 1 h.

Results

Study phase. Memory for the details of objects in scenes is highly affected by whether or not an object is
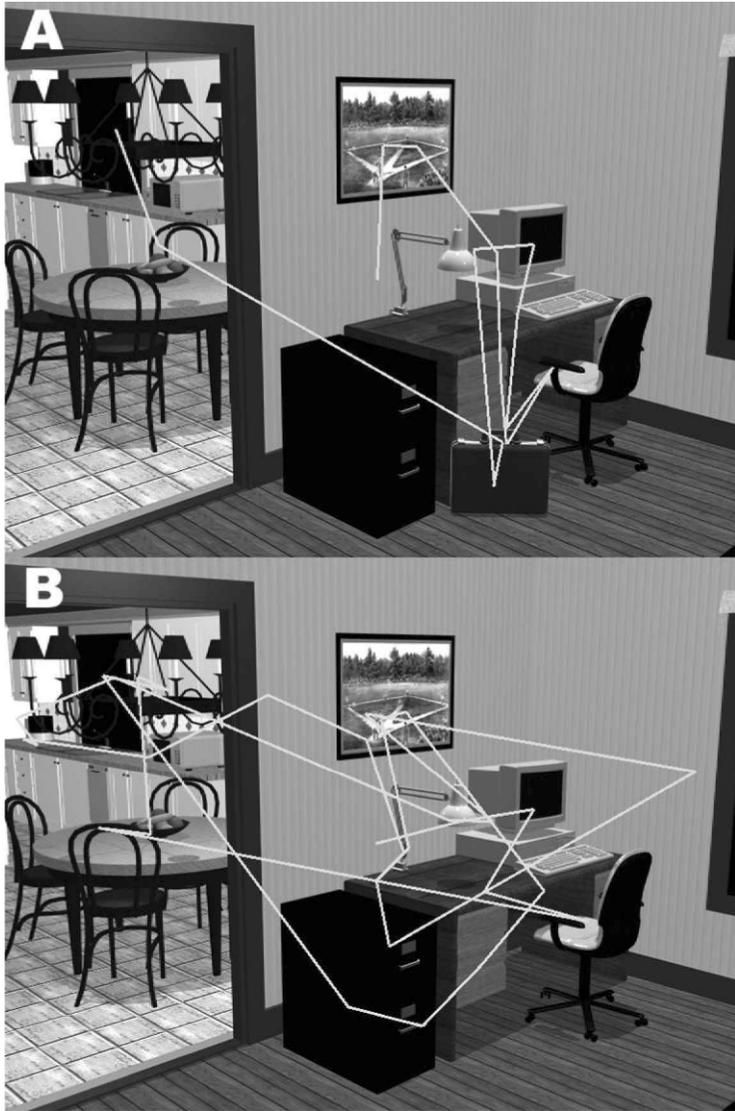
Figure 1. Example scene and scan paths in the test phase. Panel A shows the scan path of a participant who detected a substitution (briefcase), and panel B shows the scan path of a participant who failed to notice a deletion.

fixated during initial viewing (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002; Nelson & Loftus, 1980). Therefore, to ensure that a good memory representation had been formed for the critical objects, we included, in the test phase, analyses only of trials in which the critical object was fixated at least once during the study phase. For all measures that involved fixating the critical object (both in the study and the test phases), a critical region for each scene was defined as the smallest rectangular region that could be constructed around the superposition of the critical object and its substitution object. Any fixation falling within that region was considered to be a fixation on the critical object.

During the study phase, the fixation criterion was met for 92.4% of the critical objects. Conversely, there were

45 cases (7.6% of all the trials) in which a participant did not fixate the critical object during either block of the study phase. These scenes were eliminated from the test phase analyses for those participants. The number of trials lost in the test phase as a result of this criterion by condition is given in Table 1.

For the remaining study phase trials, the participants moved their eyes to the critical object from other areas of the scene (i.e., entered the critical region) about 1.68 times on average. Table 2 shows the average number of critical region entries as a function of condition. These values were marginally different from each other [$F(2,34) = 2.787$, $MS_e = 0.06693$, $p = .08$]. The average total fixation time (sum of the durations of all fixations) on the critical objects in the study phase, conditionalized on the ob-

Table 1
Summary Data for the Test Phase

	Measure			
Condition	No. of Items Excluded Because Critical Region Never Fixated	Percent Correct	A'	Response Time for Correct Trials (sec)
Identical	20	82	–	9.0
Substitution	9	66	.818	5.1
Deletion	16	50	.742	6.4

ject's receiving a fixation, was 744 msec. Table 2 shows the average total fixation time values by condition. These values did not reliably differ ($F < 1$).

The marginally reliable difference in critical region entries across conditions could not have been due to any experiment-specific variables; as far as the participants were concerned, all the study phase trials were the same. Nevertheless, the fact that the objects that would ultimately be deleted in the test phase were fixated numerically more often and for numerically more time in the study phase could influence the degree to which deletions would be noticed in the periphery, as compared with substitution or identical objects, and so could affect our estimate of peripheral processing of object presence. This possibility turned out not to raise problems, because it biases the results toward better memory performance for the deletions, an effect that (as will be discussed below) was not observed in the test phase data.

Test phase. Important evidence concerning peripheral object processing in Parker's (1978) study derived from the condition in which an object was deleted from the scene in the test phase. Specifically, the participants quickly and accurately noticed object deletions in the visual periphery without moving their eyes to the empty scene region. In the present study, we sought to replicate and extend this result in complex scenes. We examined the degree to which deletions were noticed, the time to correctly notice a deletion, and in cases in which the deletion was noticed without fixating the critical region, the likelihood of fixating close to the region.

Table 1 shows the accuracy of recognizing the object manipulations as a function of condition. First, in contrast to Parker (1978), we observed that deletions were more difficult to notice than substitutions. As can be seen in Table 1, the participants recognized the deletions on only about 50% of the trials, whereas the recognition rate in the substitution condition was 66%. The false alarm rate in the identical condition was 18%. A one-way analysis of variance showed that these three rates differed reliably [$F(2,34) = 51.51, MS_e = 0.0205, p < .001$]. The recognition rate for deletions was reliably higher than the false alarm rate [$F(1,17) = 46.93, MS_e = 0.04, p < .001$], and the recognition rate for the substitutions was reliably higher than that for deletions [$F(1,17) = 9.249, MS_e = 0.0446, p < .01$]. A signal detection analysis using nonparametric A', shown in Table 1, indicated reliable sensitivity to substitutions [$A' = .818; t(17) =$

15.417, $p < .001$], as well as to deletions [$A' = .742$; $t(17) = 8.038, p < .001$]. The signal detection analysis also showed that sensitivity for substitutions was reliably higher than that for deletions [$F(1,17) = 6.102, MS_e = 0.008547, p < .05$]. Overall, then, contrary to the results reported by Parker, deletions were more difficult to notice than were substitutions.

In addition to reporting that object deletions were recognized more accurately than other object manipulations, Parker (1978) also reported that viewers were able to respond to deletions relatively quickly, suggesting that deletions were easily noticed (i.e., tended to "pop out") in the visual periphery. To investigate this issue in the present study, we analyzed the response times to correct trials as a function of condition. As is shown in Table 1, mean response time for correct trials differed by condition [$F(2,34) = 12.848, MS_e = 5,665,261, p < .001$]. Contrary to Parker, however, we again found no evidence that deletions were recognized more readily than substitutions. Mean response time for correct recognition of the deletions (6.4 sec) was numerically slower than that for substitutions (5.1 sec) by 1.3 sec, although the difference was not reliable [$F(1,17) = 2.876, MS_e = 10,779,870, p > .10$]. The longer response times in the identical control condition (9.0 sec) indicates that the participants engaged in a relatively exhaustive search of the scenes for differences.

Finally, we examined eye movements for those trials in which the deletion was noticed without fixation. We found that on 93% of such trials, the eyes fixated within 3° of the critical region, and on 97% of such trials, the eyes fixated within 4° of the critical region. Thus, we have no evidence that deletions were detected in the visual periphery.

In summary, there was no evidence in either response accuracy or response latency to indicate that object deletions were easily recognized in the visual periphery. In-

Table 2
Summary Data for the Study Phase

	Measure	
Condition	No. of Critical Region Entries	Total Critical Region Fixation Time (msec)
Identical	1.68	745
Substitution	1.57	735
Deletion	1.78	753

stead, object deletions were noticed less accurately than object substitutions, and to the extent that they were noticed at all, they required more time to find than did substitutions. Taken as a whole, these data provide no support for the hypothesis that the absence of an expected object in a scene is easily recognized in the visual periphery.

A second finding taken by Parker (1978) to suggest that object processing during picture recognition extends into the visual periphery was that viewers tended to move their eyes about one fixation sooner to a changed object than to an unchanged object. Although saccade distances were not reported in Parker, given that the scenes consisted of objects that were about 10° apart and given that there was nothing in the scenes to look at other than the six objects, the data suggest that the information needed to detect object differences was acquired from a scene area with a diameter of up to 20° of visual angle.

In the present study, we sought to replicate and extend this finding to complex color scenes. We computed three measures of *peripheral saccadic attraction* in the test phase: the mean length of the saccade that brought the eyes to the critical region, the mean elapsed time for the eyes to reach the critical region the first time, and the mean number of fixations to reach the critical region the first time. The length of the saccade to the changed region provides an estimate of the scene region in which object processing is taking place. In addition, the tendency to "jump ahead" to a change in the periphery, as indexed by an earlier, longer saccade to the changed region, would suggest that a relatively large peripheral area of the scene had been processed in the prior fixation (Parker, 1978). Because the use of recognition as a tool for investigating eye movements in object and scene processing requires accurate memory for the original scene, each of these measures was contingent on a correct response in the test phase.

The peripheral saccadic attraction measures can be computed only given that the critical region was fixated at test. We therefore first determined the proportion of trials on which the eyes entered the critical region during the test phase as a function of condition. These data are shown in Table 3. There was a large and reliable difference in these proportions across the three conditions [$F(2,34) = 74.54$, $MS_e = 0.02077$, $p < .001$]. Contrasts against the identical baseline revealed that there was no difference in the proportion of critical regions entered (critical objects fixated) in the substitution condition

(.87) versus the identical control condition (.86; $F < 1$). However, the critical region in the deletion condition was fixated far less frequently (.35) than in the identical control condition [$F(1,17) = 78.15$, $MS_e = 0.05838$, $p < .001$]. In fact, in the test phase, 1 participant never fixated the critical region that had contained the deleted object but fixated the critical region on 73% of the substitution trials and on 91% of the identical condition trials. A comparison of the cumulative number of participants who fixated the critical region at a given percentage-of-trials criterion also revealed the low fixation rate in the deletion condition. These data, conditionalized on fixating the region in the study phase and correct response in the test phase, are shown by condition in Figure 2. In the identical control condition, 10 out of 18 participants fixated the critical region on 100% of the possible trials; in the substitution condition, 9 out of 18 participants did. In contrast, in the deletion condition 1 out of 18 participants fixated the critical region on 100% of the possible trials.

At first glance, the low fixation rate in the deletion condition appears to be consistent with the results of Parker (1978), in which it was found that participants were able to recognize deletions without fixating the deleted region. However, recall that unlike Parker, we observed that deletions were in fact noticed less often than object substitutions. In fact, the fixation rate for the deletions (35%) plus the guessing rate (18% false alarms) accounts well for the 50% deletion detection rate. Thus, the lower rate of fixation of the deleted object region was not the result of greater peripheral detection without fixation. Instead, these data suggest that regions from which objects had been deleted neither attracted fixations nor were readily recognized as having changed.

To investigate peripheral saccadic attraction, we first computed the mean amplitude of the initial saccade into the critical object region as a function of condition. This analysis was contingent on a correct response and on the critical object's receiving a fixation.[2] The mean amplitude of the first saccade to the critical object across conditions was about 4.2°. As can be seen in Table 3, there was some tendency for saccades to the critical region in the substitution condition to be greater in amplitude than those in the identical and deletion conditions, but this difference did not approach statistical reliability ($F < 1$). Furthermore, even in the substitution condition, the average amplitude was only 4.4°, a value that is less than

Table 3
Eye Movement Data for the Test Phase

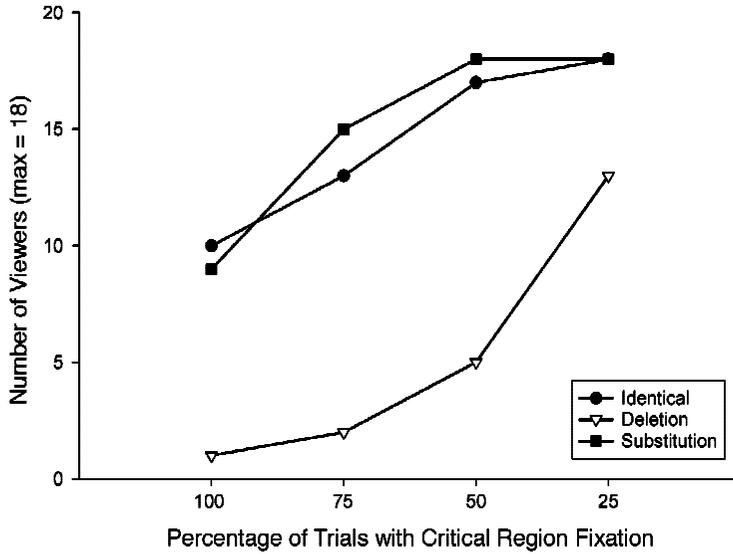| Condition	Measure			
	Proportion of Critical Objects Fixated	Amplitude of First Saccade to Critical Object Region (deg)	No. of Fixations to Initial Critical Object Region Fixation	Elapsed Time to Initial Critical Object Region Fixation (sec)
Identical	.86	3.96	7.5	2.35
Substitution	.87	4.40	5.2	1.62
Deletion	.35	3.97	11.8	3.87

Figure 2. Cumulative number of viewers reaching a criterion level of critical region fixations, by condition.

half of the object-to-object distances in Parker's (1978) study. The present saccade amplitude value is consistent with other reports in the literature (see Henderson & Hollingworth, 1998, 1999a). There was no tendency for the amplitudes of saccades to be greater to a scene region from which an object had been deleted than to a scene region containing a replacement object.

Figure 3 shows the frequency histogram of saccade amplitudes for all first saccades to the critical region in the test phase. This plot includes only correct trials and collapses the data across the three test conditions. It is clear that the modal saccade length to the critical region was about 4° and that there were very few saccades equal to or greater than 8°.[3]

As a second measure of peripheral saccadic attraction, we computed the mean number of fixations between the onset of the scene and the first fixation within the critical region for those trials in which the changes were recognized. Recall that Parker (1978) reported that participants fixated the changed objects about 1 fixation sooner than the unchanged objects out of an average of 15.5 fixations. In the present study, if changed regions in the visual periphery were able to attract saccades, the mean number of fixations should be less for the object manipulation conditions than for the identical control condition. Moreover, we would expect that if deletions are particularly able to attract saccades, the mean number of fixations to the critical region from which an ob-
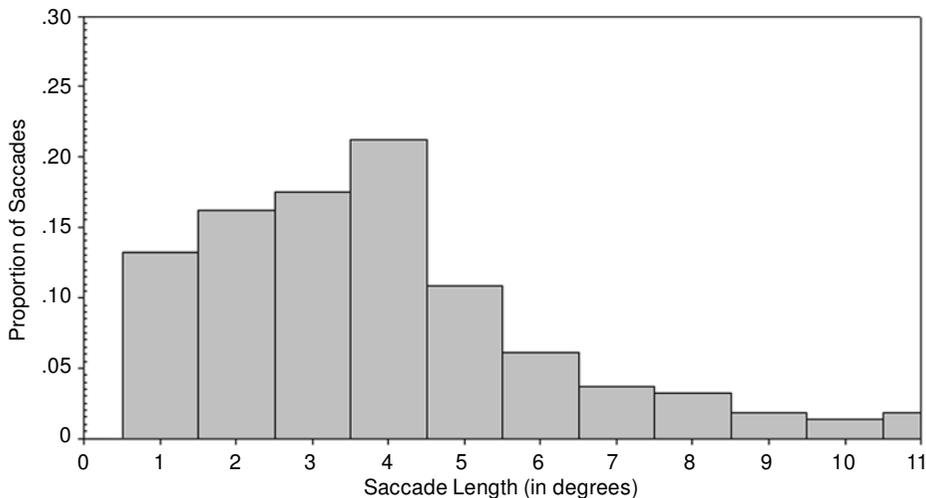

Figure 3. Histogram of saccade amplitudes in the test phase.

ject had been deleted should be less than that to a region containing an object substitution. Table 3 shows the mean number of fixations to the critical region for correct trials in the three conditions. Overall, there was a reliable effect of condition on number of fixations [$F(2,34) = 6.744$, $MS_e = 30.296$, $p < .005$]. Consistent with the hypothesis that a change attracts the eyes, the eyes first reached a substituted object about 1.5 fixations sooner than they first reached the critical object in the identical control condition [$F(1,17) = 12.75$, $MS_e = 7.657$, $p < .01$]. However, as was reported above and is shown in Table 3, the amplitude of the saccade taking the eyes to the critical object was equivalent in the substitution and identical conditions, suggesting that the lower fixation number in the former case was based on object processing for a relatively restricted spatial region around the current fixation position. It appears that once the eyes landed within about 4° of an object, there was increased likelihood that the object would be fixated if it was a substitution. In contrast to the substitution results, the eyes took slightly over 4 fixations *longer* to reach the deleted region than to reach the critical object in the control condition [$F(1,17) = 3.718$, $MS_e = 89.812$, $p = .07$]. Therefore, we again found no evidence that the distance at which the deletion of an expected object is recognized is greater than the distance at which other types of object manipulations are recognized.

A third measure of peripheral saccadic attraction, related to number of fixations, is the elapsed time between the onset of the scene and the first fixation within the critical region. These data are presented in Table 3. Overall, there was a reliable effect of condition on elapsed time [$F(2,34) = 5.674$, $MS_e = 4,172,159$, $p < .01$]. Consistent with the hypothesis that a manipulation attracts the eyes, a replaced object was fixated 0.731 sec sooner than the object in the identical control condition [$F(1,17) = 11.209$, $MS_e = 858,039$, $p < .01$]. Again, however, it is important to keep in mind that the saccades leading to fixation on the critical object were no longer in the substitution condition than in the identical control condition. Furthermore, the eyes were 1.5 sec *slower* to reach the deleted region than they were to reach the critical object in the control condition, although this difference was only marginally reliable [$F(1,17) = 3.436$, $MS_e = 12,059,452$, $p = .08$]. Together with the analysis of the number of fixations to the critical region, the elapsed time data provide no support for the hypothesis that the absence of an expected object is detected at a greater distance from fixation than are changes to other expected object properties.

GENERAL DISCUSSION

In the present study, the participants were shown complex color images depicting real-world scenes. Each scene was shown twice during a learning phase and then, again, a third time in a test phase, either in the original version or with a manipulation to one object. In the latter case, an object from the original scene was deleted or replaced with an object from a different conceptual cat-

egory. The experiment was an attempt to replicate and extend the finding, originally reported by Parker (1978), that scene processing during memory recognition for such properties as object presence and conceptual type extends into the visual periphery.

Three primary results in Parker's (1978) study initially supported the conclusion that object properties in scene recognition are acquired and processed from a relatively large region of a scene. First, recognition of object deletions was very rapid once a scene with a deleted object appeared. Second, these rapid detections occurred without the need to fixate the region from which the object had been removed. Third, in the course of scene viewing during the recognition test, the participant's gaze tended to "jump ahead" to the region of the scene containing a changed object. These results led Parker to the conclusion that the processing of object properties extends over a wide area of a scene.

We found several results that contrast with those reported by Parker (1978). First, object deletions were relatively difficult to notice, and furthermore, recognition time for correctly noticing deletions was numerically longer than it was for noticing that a different object had been substituted for the original object. The conclusion seems to be that deleted objects are easily missed, rather than easily noticed. Second, although there was some tendency for changed regions to be fixated sooner than unchanged regions, the average distance of the saccade to the changed region was about 4°, considerably smaller than the average distances between the objects in Parker. These results suggest that in complex pictures of natural scenes, object properties are acquired from a region with a radius of about 4° from fixation, on average. This conclusion is also supported by analyses of the saccade distributions in the present study, which showed that the modal interobject saccadic amplitude across conditions was 4°, with saccades of 10° being extremely infrequent (see Figure 3).

How do we account for the difference between the present results and those reported by Parker (1978)? First, with respect to deletions, it seems likely that the use of simple arrays of objects as stimuli in Parker made it relatively easy for participants to notice the removal of an object from the image. Because each scene comprised an array of six objects shown against a blank field, the display could easily be encoded as a spatial configuration of occupied positions. Simons (1996) has demonstrated that changes to the spatial configuration of a small set of objects are much more detectable than changes to object identities, a finding that mirrors Parker's results. In contrast, the deletion of an object from the kind of complex natural scene used in the present study creates much less change to the overall configuration of the image. This occurs, in part, because the removal of one object changes only a small fraction of the overall configuration and, in part, because portions of other objects that were previously occluded in the scene are revealed by the deletion, reducing the overall impact of the con-

figural change. In fact, a common finding in the memory literature is that the deletion of visual information from a complex scene is much more difficult to notice than is the addition of visual information (Pezdek et al., 1988; see also Hearst, 1991). It seems, then, that the participants in the deletion condition of Parker were detecting changes to the spatial configuration of items and that detection of such changes can be accomplished over a relatively wide area of an image. In contrast, detecting the absence of an object from a complex scene is much more spatially restricted, with an average radius closer to 4° of visual angle from fixation, at least for the type of complex scene images used here.

The results of the present study are in line with other reports of object processing in the scene perception literature (e.g., De Graef et al., 1990; Henderson et al., 1999; Nelson & Loftus, 1980; see Henderson & Hollingworth, 1998, 1999a, for reviews). For example, as was described in the introduction, Henderson et al. observed that the amplitudes of saccades to objects that were semantically anomalous within their scenes were no greater than those to semantically consistent objects, with an average amplitude of 3.5° in both conditions. These results suggested that the processing of object meaning is restricted to the foveal and parafoveal regions of the scene. In the present experiment, our estimate for noticing type substitutions was a bit larger, with average saccadic amplitudes to the critical region in the substitution condition of about 4.4°. However, we cannot be certain that the type substitutions in the present study were detected entirely on the basis of conceptual information; it could be that visual differences between the original and the substitution objects were noticed and drove eye movements, a possibility that was not open in our earlier study. Alternatively, it may be that processing of object semantics in line drawings takes place over a slightly smaller scene region than it does for the same information in full-color images. In either case, the present data clearly indicate that the participants were not acquiring and comparing object information from a perceptual span of 10° from fixation, as Parker's (1978) results suggested, but rather were acquiring object information from less than half that distance.

It is important to note that we are not suggesting that the area in which scene information is processed is equally limited for all types of information. For example, it is likely that gist (i.e., the overall meaning of the scene) and global layout are processed over a significantly larger scene area than are the identities of specific objects (Oliva & Schyns, 1997; Sanocki & Epstein, 1997). An analogous situation exists in reading, where the total perceptual span region (including word length cued by the spaces between words) is considerably larger (about 3–4 characters to the left of fixation and 15 to the right) than it is for word identities (the *word span*), which is, at best, two words and probably closer to one (Rayner, 1998). Similarly, in scene perception, it is likely that different types of information can be acquired at different distances from fixation. For example, recognition of the deletion of an object might be based on visual information (e.g., visual features present in that scene region are no longer present or have been replaced by other features that were previously occluded), semantic and conceptual information derived from the visual information (e.g., "the stapler is missing"), or information about spatial layout (e.g., a region of the scene that was occupied is now empty, or a configuration of items changed its shape). Type substitutions, in comparison, cannot be recognized on the basis of changes to scene layout but could reflect recognition of changes to the relatively gross visual shape information needed to distinguish these objects or could reflect recognition of changes to associated semantic and conceptual information.

The results of the present study seem at odds with one recent report of eye movements in changed and unchanged scenes (Ryan, Althoff, Whitlow, & Cohen, 2000). In that study, participants learned complex scenes and then viewed the same scenes again either in their original version or in a manipulated version. The manipulations included object addition, object deletion, and a left–right object shift. The main result of direct relevance to the present study was a change in viewing patterns for the manipulated regions, which the authors called a *relational manipulation effect*. For example, the participants had a higher proportion of fixations in the manipulated region than in the equivalent control region, as well as a higher number of transitions into and out of the manipulated region. In addition, the relational manipulation effect appeared to be restricted to those cases in which a change was not explicitly noticed. In the present experiment, we followed Parker (1978) in emphasizing eye movement patterns associated with explicit recognition of the object manipulation. Although Ryan et al. found no relational manipulation effects when the viewers explicitly noticed the manipulations, we found clear eye movement effects despite explicit recognition. However, in the present study, our emphasis was on the *initial* look to the manipulated region, whereas the relational manipulation effect emphasizes the *overall* time spent attending the manipulated region. That is, we were interested in whether the eyes would be initially drawn to a changed region and, if so, when and from what distance that would occur. Ryan et al., in contrast, reported such measures as the total numbers of fixations in the changed region and the proportions of all of the transitions into and out of the region. The latter measures collapse over multiple entries into the critical region, making it impossible to disentangle initial entry from overall entries. It could be that a region that is less likely to be entered (e.g., a region containing a deletion) still receives a higher total number of fixations than does a region that is more likely to be entered, because for the latter case, when such a region is entered, it tends to be reentered multiple times (e.g., Henderson et al., 1999). Unfortunately, a direct comparison of the overlapping condition (deletion) in the present study with the equivalent con-

dition in Ryan et al. is not possible, because Ryan et al. reported their data only collapsed across their three manipulation conditions.

In summary, the present study demonstrates that the information needed to recognize object deletions and type substitutions is acquired from a relatively restricted area of a picture of a complex, full-color, real-world scene. These results suggest that the processing of object information in a scene is spatially relatively limited, consistent with other recent evidence. The results do not, however, undermine the possibility that other sorts of scene information, such as gist, can be acquired from a more expansive area of a viewed scene.

REFERENCES

BOUMA, H. (1978). Visual search and reading: Eye movements and functional visual field. A tutorial review. In J. Requin (Ed.), *Attention and performance VII* (pp. 115-147). Hillsdale, NJ: Erlbaum.

BUSWELL, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

CASTELHANO, M. S., & HENDERSON, J. M. (2001, November). *Eye movements, viewing task, and scene memory.* Paper presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.

CRANE, H. D. (1994). The Purkinje image eyetracker, image stabilization, and related forms of stimulus manipulation. In D. H. Kelly (Ed.), *Visual science and engineering: Models and applications* (pp. 15-89). New York: Marcel Dekker.

CRANE, H. D., & STEELE, C. M. (1985). Generation-V dual-Purkinje-image eyetracker. *Applied Optics*, **24**, 527-537.

DE GRAEF, P., CHRISTIAENS, D., & d'YDEWALLE, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, **52**, 317-329.

HEARST, E. (1991). Psychology and nothing. *American Scientist*, **79**, 432-443.

HENDERSON, J. M., & HOLLINGWORTH, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269-283). Oxford: Elsevier.

HENDERSON, J. M., & HOLLINGWORTH, A. (1999a). High-level scene perception. *Annual Review of Psychology*, **50**, 243-271.

HENDERSON, J. M., & HOLLINGWORTH, A. (1999b). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, **10**, 438-443.

HENDERSON, J. M., & HOLLINGWORTH, A. (2003a). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics*, **65**, 58-71.

HENDERSON, J. M., & HOLLINGWORTH, A. (2003b). Eye movements, visual memory, and scene representation. In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (pp. 356-383). New York: Oxford University Press.

HENDERSON, J. M., WEEKS, P. A., JR., & HOLLINGWORTH, A. (1999). The effects of semantic consistency on eye movements during scene viewing. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 210-228.

HOLLINGWORTH, A., & HENDERSON, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of*

Experimental Psychology: Human Perception & Performance, **28**, 113-136.

HOLLINGWORTH, A., SCHROCK, G., & HENDERSON, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*, **29**, 296-304.

LOFTUS, G. R., & MACKWORTH, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 565-572.

NEISSER, U. (1976). *Cognition and reality*. San Francisco: Freeman.

NELSON, W. W., & LOFTUS, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 391-399.

OLIVA, A., & SCHYNS, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, **34**, 72-107.

PARKER, R. E. (1978). Picture processing during recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 284-293.

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, & Cognition*, **14**, 468-476.

POLLATSEK, A., RAYNER, K., & COLLINS, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, **113**, 426-442.

RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.

RYAN, J. D., ALTHOFF, R. R., WHITLOW, S., & COHEN, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, **11**, 454-461.

SANOCKI, T., & EPSTEIN, W. (1997). Priming spatial layout of scenes. *Psychological Science*, **8**, 374-378.

SIMONS, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, **7**, 301-305.

SIMONS, D. J., & LEVIN, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, **1**, 261-267.

YARBUS, A. L. (1967). *Eye movements and vision*. New York: Plenum.