# Eye movements and visual memory: Detecting changes to saccade targets in scenes

JOHN M. HENDERSON
*Michigan State University, East Lansing, Michigan*

and

ANDREW HOLLINGWORTH
*Yale University, New Haven, Connecticut*

Saccade-contingent change detection provides a powerful tool for investigating scene representation and scene memory. In the present study, critical objects presented within color images of naturalistic scenes were changed during a saccade toward or away from the target. During the saccade, the critical object was changed to another object type, to a visually different token of the same object type, or was deleted from the scene. There were three main results. First, the deletion of a saccade target was special: Detection performance for saccade target deletions was very good, and this level of performance did not decline with the amplitude of the saccade. In contrast, detection of type and token changes at the saccade target, and of all changes including deletions at a location that had just been fixated but was not the saccade target, decreased as the amplitude of the saccade increased. Second, detection performance for type and token changes, both when the changing object was the target of the saccade and when the object had just been fixated but was not the saccade target, was well above chance. Third, mean gaze durations were reliably elevated for those trials in which the change was not overtly detected. The results suggest that the presence of the saccade target plays a special role in transsaccadic integration, and together with other recent findings, suggest more generally that a relatively rich scene representation is retained across saccades and stored in visual memory.

In human vision, high-quality visual information is available only at the fovea, and saccadic eye movements are used to direct fixation toward important stimuli in the current scene (Buswell, 1935; Yarbus, 1967). Because vision is effectively suppressed during saccades, the visual system is confronted with a series of temporally discrete and spatially displaced glimpses of the world, each lasting about 300 msec on average (Buswell, 1935; for review, see Henderson & Hollingworth, 1998, 1999a). The generation of an overall scene representation would therefore seem to require that information acquired during one fixation be combined with information acquired from prior and subsequent fixations. An important question in visual per-

ception and cognition, then, is the nature of the information that is retained and combined across successive fixations. The traditional view in vision science has been that a complete sensory image is retained across each saccade and fused with the sensory image from the following fixation (Breitmeyer, Kropfl, & Julesz, 1982; Duhamel, Colby, & Goldbert, 1992). However, the evidence is quite strong that a visually veridical sensory image of a scene is not retained and fused across saccades (Irwin, 1991, 1992; Pollatsek & Rayner, 1992). In addition, evidence from a variety of change detection paradigms has often demonstrated seemingly remarkable insensitivity to visual changes across saccades (Grimes, 1996; Henderson, 1997) and other visual disruptions (O'Regan, Rensink, & Clark, 1999; Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1998), a phenomenon known as "change blindness" (Simons & Levin, 1997). Recognition of this phenomenon has led to the suggestion that our conscious experience of a complete visual scene is an illusion (Dennett, 1991; O'Regan, 1992), and that contrary to experience, nothing is retained in memory beyond the general gist of the scene, the identities of specific objects, a coarse representation of spatial layout, and the visual content of the currently attended object or scene region (O'Regan, 1992; Rensink, 2000a, 2000b; Rensink et al., 1997; Simons & Levin, 1997).

In contrast to this latter proposal, more recent evidence from change detection experiments suggests that the nature of the scene representation constructed dynamically

across fixations and stored in memory is more complete than change blindness originally implied (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001). Fixation position within a scene, in particular, appears to play an important role in determining whether visual information from the scene is encoded during one fixation, stored, and retrieved during a subsequent fixation, and hence whether changes will or will not be detected. For example, viewers are able to detect relatively subtle changes in scenes such as object rotations and token replacements (e.g., changing one telephone to another telephone) that take place during a saccade, as long as the changed object is fixated before and after the change (Hollingworth & Henderson, 2002). Cued recognition of object detail following scene viewing is also quite good if the critical object was fixated during initial scene viewing (Hollingworth & Henderson, 2002; see also Friedman, 1979; Nelson & Loftus, 1980). Detection of changes in complex real-world scenes during the flicker paradigm is similarly closely related to fixation position (Hollingworth, Schrock, & Henderson, 2001).

The location of the change with respect to the direction of a saccadic eye movement also affects the degree to which information in an image will be encoded, with better encoding of information from the scene region toward which the eyes are about to saccade (the saccade target) than the region from which the saccade was launched (the saccade source), or another region of the scene that was neither the saccade source nor the target. This effect may be the result of the allocation of attention toward the saccade target prior to a saccade (Henderson, 1996; Henderson, Pollatsek, & Rayner, 1989; Hoffman & Subramaniam, 1995; Irwin & Andrews, 1996; Rayner, McConkie, & Ehrlich, 1978; Shepherd, Findlay, & Hockey, 1986). However, there appear to be saccade target effects in addition to those that can be attributed to the allocation of attention: In Henderson and Hollingworth (1999b), participants were able to detect with very high accuracy the deletion of the saccade target when the deletion took place during that saccade. Deletion of the saccade source was not as detectable as deletion of the saccade target. Importantly, detection of saccade target deletions did not decrease with the amplitude of the saccade to that object, whereas detection of the deletion of an object that was the saccade source did fall off with saccade amplitude. In comparison, detection of another object change, the in-depth orientation of an object in the scene, was not as accurate (though it was still significantly above chance), and did drop off with saccade amplitude. Together, these results suggest that there may be something special about the coding of the presence of the saccade target in integrating information across eye movements.

The saccade target deletion effects observed in the change detection paradigm (high change detection rate and relative imperviousness of this high rate to the distance of the saccade source to the changing target) provide a potentially useful tool for investigating transsaccadic memory and integration during scene perception.

As a first pass, we can generate two hypotheses about what properties make the deletion of a saccade target so much more noticeable than changes to other visual characteristics of an object either at the target location or elsewhere in the scene. First, it could be that saccade target *presence* is specially coded, making the absence of the target following the saccade regardless of saccade amplitude particularly salient. For example, it could be that a content-free representation of the position of the saccade target (perhaps in a configuration with other such location markers) is used either solely or primarily to map the presaccade scene representation onto the postsaccade scene view during transsaccadic integration. The proposal that object configuration plays an important role in transsaccadic memory is consistent with recent work showing that visual short-term memory is preferentially accessed by information about spatial configuration (Jiang, Olson, & Chun, 2000), and is also consistent with evidence that spatially configured object tokens ("object files") are functional in transsaccadic integration during object identification (Henderson, 1994; Henderson & Anes, 1994; Henderson & Siefert, 2001). If object presence per se is specially coded for the target of a saccade, then only saccade target deletions should be well detected and show insensitivity to saccade distance.

Second, it could be that the identity or semantic category of the saccade target is specially coded, and that it is the change in identity or category at fixation following the saccade that is especially noticeable. For example, the transsaccadic system might code something like "apple" at the saccade target area prior to a saccade, and then check to determine that an object with that concept or identity is present after the saccade. If this hypothesis is correct, then changing the saccade target to an object with a different basic-level concept and identity should be detected as well as deletions, and should also show insensitivity to saccade amplitude.

To distinguish between these two hypotheses, the present study employed the *object boundary paradigm* introduced by Henderson and Hollingworth (1999b). In this paradigm, changes to prespecified critical objects are triggered when the eyes cross the boundary of a software-defined critical region surrounding the critical object. Figure 1 presents an example scene. In this example, the critical object is the phone on the desk. A change to the object within the critical region was made either during the first saccade entering this region (toward condition) or the first saccade exiting this region once the target was initially fixated for at least 90 msec (away condition). The toward condition is of primary interest because this is the condition that is diagnostic of the nature of the representation generated and retained at the saccade target location; the away condition serves as a control condition against which to compare saccade target changes.

To test the hypotheses concerning the nature of the representation retained from fixation to fixation at the saccade target location, the critical object was deleted, its semantic type was changed, or its instantiation as a visual token of a particular semantic type was changed while the

**Figure 1. Example scene used in Experiments 1A and 1B. The target object was the telephone on the desk. The phone changed to another phone, to a notebook, or was deleted.**

semantic type was maintained. The deletion condition provides evidence about the degree to which target presence is represented across saccades, the type change condition provides evidence about the degree to which identity and basic-level concept is represented, and the token change condition provides evidence about the degree to which specific visual information about the object is represented across a saccade. Evidence from the transsaccadic object identification literature suggests that visually specific representations are retained across saccades: Transsaccadic object identification is affected by changes to the specific details present in an object image before and after a saccade even when the identity and basic-level concept of the object remains unchanged. For example, transsaccadic preview benefits are reduced by token substitution (e.g., changing one dog to another dog, Henderson & Siefert, 2001) and mirror reflections (Henderson & Siefert, 1999, 2001). Evidence from change detection has similarly shown that the information necessary to discriminate one token from another can be retained across multiple views in a complex scene (Hollingworth & Henderson, 2002), though the degree to which these representations are specifically maintained for the saccade target has not yet been investigated.

The present study also allowed us to examine two additional questions related to scene representation and memory. First, as noted, the degree to which scene detail is preserved in memory has recently become controversial. According to one view, memory representations for viewed scenes include, at best, information about semantic gist, object identities, coarse spatial layout, and the visual content of the currently attended object (Rensink, 2000a, 2000b; Simons & Levin, 1997; Wolfe, 1999). The change blindness phenomenon has been taken to provide evidence for this view. In contrast, other sources of evidence, including recent findings using the transsaccadic change detection methodology, suggest that memory representations for scenes are more complete and detailed than has been inferred from the change blindness phenomenon (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001; see Henderson & Hollingworth, in press, for review). The present study provided further opportunity to investigate the degree to which visually specific representations are preserved during scene perception. On the basis of our previous studies, we expected that object changes, including visual changes (indexed by the token change condition), would be detected at rates significantly above chance. Particularly diagnostic of the preservation of visual representations in short-term memory would be the ability to detect token changes in the away condition, because in the away condition attention has been withdrawn from the critical object prior to the change (Hollingworth & Henderson, 2002). Thus, an ability to detect visual changes in this condition requires that a visual representation be preserved in memory from one fixation to the next.

Second, an overt change detection response is one indication that a viewer has retained scene information in

memory. However, a number of investigators have shown that overt response measures do not provide complete evidence about whether the information needed to detect a change is available (Fernandez-Duque & Thornton, 2000; Hayhoe, Bensinger, & Ballard, 1998; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001; Williams & Simons, 2000). The present study provided us with an opportunity to examine the extent to which overt measures of change detection underestimate the completeness of the underlying scene representation for the saccade target.

## EXPERIMENT

Participants were instructed to study complex real-world scenes to prepare for a memory test in which they would have to distinguish the previously viewed scenes from versions of the scenes in which only a small detail had been changed. Participants were also told to monitor for object changes and to press a button whenever such a change was detected (Grimes, 1996; Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002). To investigate the representation and retention of information about a saccade target across eye movements within a scene, changes were made to prespecified objects as a function of (1) the location of the change with respect to the direction of the eye movement at the time of the change and (2) the nature of the object change itself. For the manipulation of the location of the change with respect to eye movement direction, changes took place either during the first saccade that brought the eyes to the critical object (toward condition) or during the saccade that took the eyes away from the critical object immediately after it had been fixated the first time (away condition). To investigate the nature of the information that is encoded and retained across a saccade, three types of object changes were compared. In the type change condition, the critical object was replaced by another object that differed in identity and basic-level semantic category. These changes also involved changes to visual characteristics of the objects, though object size was maintained across the change. In the token change condition, an object that was a member of the same basic-level category but that differed in visual detail replaced the critical object. Finally, in the deletion condition, the critical object was removed from the scene; any background that was occluded by the critical object was revealed at the time of the deletion. No-change catch trials were included to provide an assessment of the false alarm rate in the experiment.

To ensure sufficient statistical power, the study was divided into two subexperiments, each examining a subset of the change conditions in a different set of participants. In Experiment 1A, the location of the change with respect to eye movement direction was crossed with the type change and deletion conditions. In Experiment 1B, the location of the change with respect to eye movement direction was crossed with the type and token change conditions. In this way, the type change condition provided a common condition across subexperiments, and all three conditions could be examined with the limited number of scenes available.

## Method

**Participants**. Thirty Michigan State University undergraduate students participated in the experiment for course credit, 15 each in Experiments 1A and 1B. All participants had normal vision and were naive with respect to the hypotheses under investigation.

**Stimuli**. Thirty-five scene images were computer rendered from 3-D wire-frame models using a commercial rendering program. Wire-frame models were acquired commercially, donated by 3-D graphic artists, or developed in house. Each model depicted a typical human-scaled environment. Base scenes were rendered from these models. To create the type change, token change, and deletion conditions, the critical objects were replaced or removed in the models, and the scenes were rerendered. All scene images subtended 15.8° × 11.9° visual angle at a viewing distance of 1.13 m. Critical objects subtended 2.43° on average along the longest axis. The objects used for the type and token changes were chosen to be similar in size to the initial critical object in each scene. Figure 1 shows a sample stimulus scene. As can be seen in Figure 1, the critical objects were placed in an uncluttered region of the scene offset from the center so that saccades to them (and fixations on them) could be easily identified. Nothing about the critical objects themselves identified them as different from the other objects in the scenes.

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

Stimuli were displayed at a resolution of 800 × 600 pixels × 256 colors on an NEC Multisync P750 monitor driven by a Hercules Dynamite 128/Video graphics card. The screen refresh rate was 143 Hz. The room was dimly illuminated by an indirect, low-intensity light source. With this display equipment and these viewing conditions, the scene changes used here cannot be detected from phosphor persistence, as shown via an electronic shutter test (Henderson & Hollingworth, 1999b).

Buttonpresses were collected with a button panel connected to a dedicated input/output (I/O) card. The eyetracker, display monitor, and I/O card were interfaced with a 90-MHz Pentium-based computer. The computer controlled the experiment and maintained a complete record of eye position and time values, as well as buttonpress events and times, over the course of each trial.

**Procedure**. Upon arriving for the experimental session, participants were given a written description of the experiment along with a set of instructions. The description informed participants that their eye movements would be monitored while they viewed images of naturalistic scenes on a computer monitor. Participants were instructed to view each scene in preparation for a memory test that would be given after all scenes had been shown. They were told that "on the test, you will have to distinguish the scenes presented in the experiment from new versions of the scenes that may differ in only a small detail of a single object." They were further told that while they were viewing each scene, a change might occur to a single object. For each subexperiment, the two possible types of changes were described using a sample scene. Participants were instructed to press a response button as soon as such a change was detected and that if a change were to occur in a scene, it would occur only once. Finally, participants were told that on some trials, no change would occur.

Following review of the instructions, the experimenter calibrated the eyetracker. Calibration was considered accurate if the computer's estimate of the current fixation position was within ±5' arc of each

marker. The participant then completed the experimental session. Calibration was checked every three to four trials, and the eyetracker was recalibrated between trials when necessary.

A trial consisted of the following events. First, a fixation screen was shown. When the participant fixated a central box in this screen (as indicated by a computer-generated display of its estimated fixation position), the experimenter started the trial. The initial version of the scene, containing the prechange version of the critical object, was displayed until the participant's gaze crossed the boundary into the critical object region (toward condition) or crossed the boundary exiting the critical region after that region had been fixated for a minimum of 90 msec (away condition). The boundary region was 0.36° larger on each side than the smallest rectangle enclosing the critical object. When the eyes crossed the boundary, the scene image changed so that it contained the postchange version of the critical object, or no critical object in the case of the deletion condition. In the vast majority of trials this change was completed during a saccade. Viewing continued for 20 sec, or until the participant pressed the response button, indicating that a change had been detected.

In both Experiments 1A and 1B, each participant viewed 35 scenes for 20 sec each. Twenty-eight of the scenes changed as a function of the $2 \times 2$ factorial combination of eye movement condition (toward vs. away) $\times$ change condition (for Experiment 1A, type change vs. deletion; for Experiment 1B, type change vs. token change). An additional seven scenes did not change in each subexperiment; these trials provided an opportunity to determine the false alarm rate. Within each subexperiment, scenes were assigned to eye movement and change conditions via a Latin square design so that each scene appeared in each condition an equal number of times across participants. The order of scene presentation (and hence the order of condition presentation) was determined randomly for each participant within each subexperiment. Participants were assigned to subexperiment using a pseudorandom procedure; each participant took part in only one experiment. Each subexperiment lasted approximately 35 min.

## Results and Discussion

Eye movement data files consisted of time and position values for each eyetracker sample. Saccades were defined as changes in eye position greater than 8 pixels (about 8.8′ of arc) in 15 msec or less. Samples that did not fall within a saccade were considered part of a fixation. During a fixation, eye position does not remain perfectly still; the position of each fixation was calculated as the mean of the position samples (weighted by the duration of time at each position) that fell between consecutive saccades (Henderson, McClure, Pierce, & Schrock, 1997). Fixation duration was calculated as the elapsed time between consecutive saccades. Fixations less than 90 msec and greater than 2,500 msec were eliminated as outliers. Trials were eliminated if the eyetracker lost track of eye position prior to the change or if the change was not completed before the beginning of the next fixation on the scene. Eliminated trials accounted for 15.4% of the data in Experiment 1A and 14.7% of the data in Experiment 1B.

Figure 2 shows detection performance in all conditions across the two subexperiments. The solid bars in the figure show target detections that took place within 1,500 msec of a change, and the hatched extensions show late detections, defined as those that did not occur within the first 1,500 msec after the change.[1] Change detection analyses were conducted over all detections. Overall false alarm rates were under 8% in both experiments (7.6% in Exper-

iment 1A and 1.9% in Experiment 1B). Change detection in all change conditions was reliably better than their false alarm rates (all $ps < .05$).

As can be seen in Figure 2, change detection was generally poorer when the eyes had just fixated but were moving away from the critical object at the time of the change than when they were moving toward the critical object. This result replicates the earlier finding that motivated the present study (Henderson & Hollingworth, 1999b). In both Experiments 1A and 1B, there was a main effect of the location of the critical object change with respect to eye movement direction [$F(1,14) = 8.83, p < .01$, and $F(1,14) = 11.89, p < .005$, respectively]. Eye movement condition and change condition did not produce a reliable interaction in either subexperiment ($Fs < 1$).

**The saccade target deletion effect**. In terms of the primary issue of change detection for the saccade target, participants were quite sensitive to object deletions, particularly when the deletions took place during the saccade toward the deleted object, with an overall toward deletion detection rate of 91.5%. The high level of performance in the toward deletion condition replicates the results of our earlier study (Henderson & Hollingworth, 1999b). In contrast, type changes in the toward condition were detected reliably less often than deletions, at a rate between 50% and 60% in both Experiment 1A (53.5%) and Experiment 1B (58.8%). The toward deletion and toward type conditions (Experiment 1A) reliably differed [$F(1,14) = 18.29, p < .005$]. This difference suggests that the detection of deletions in the toward condition was not based only on retention of information about the identity or basic-level category of the saccade target. If it had been, then both deletions and type changes should have been detected at the same rate. Instead, the advantage of the deletion condition over the type change condition suggests that the detection of saccade target deletions is based on the retention of additional information beyond identity or basic-level category.

A signature of the special nature of the saccade target in the transsaccadic change detection paradigm is the lack of a reduction in deletion detection as a function of saccade amplitude (Henderson & Hollingworth, 1999b). In the present study, we again observed this effect. Figure 3 depicts detection performance in all conditions as a function of the spatial extent of the saccade that triggered the change across the two subexperiments. To make the regression on saccade length meaningful, we included only data from trials in which the change was detected immediately following (within 1,500 msec of) the image change. As can be seen in the top line of Figure 3, we once again observed no drop-off in detection rate as a function of saccade amplitude in the toward deletion condition. In fact, there was some tendency for performance to increase with saccade amplitude in this condition [$R_{pb} = .20, t(66) = 1.68, p = .097$].[2] The failure to observe a drop-off in the toward deletion condition is perhaps most striking when compared with the away deletion condition, where there was a clear decrease in detection performance with in-
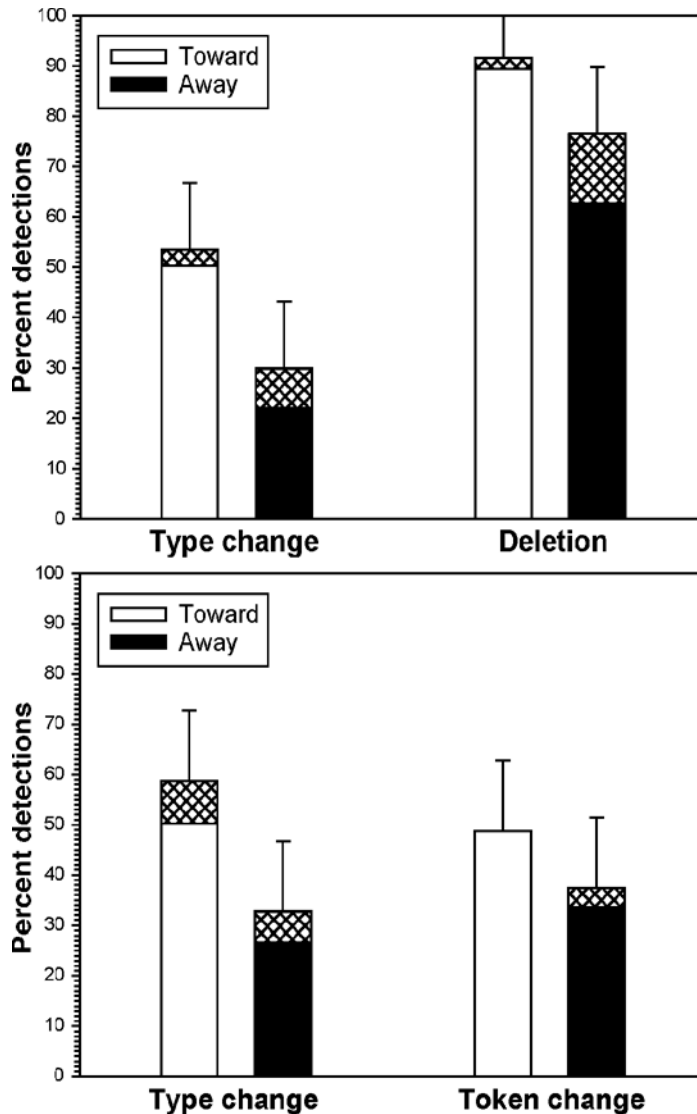
**Figure 2. Change detection performance (percent detections) in Experiment 1A (top panel) and Experiment 1B (bottom panel). The full bars represent immediate target detections (within 1,500 msec of a change), and the hatched extensions represent late detections (later than 1,500 msec following a change). Error bars are 95% confidence intervals based on the error term of the interaction between saccade direction and change.**

creasing saccade distance [$R_{pb} = -.39, t(71) = -3.60, p < .001$]. The difference in performance between the toward deletion and away deletion conditions suggests that it is not simply that deletion is particularly noticeable at further eccentricities (see also Henderson & Hollingworth, 1999b). Similarly, in contrast to the results observed in the toward deletion condition, change detection fell as a function of saccade amplitude in the toward type condition [$R_{pb} = -.17, t(141) = -2.08, p < .05$, pooling observations from Experiments 1A and 1B], and there was a nonsignificant trend in the same direction in the toward token condition [$R_{pb} = -.13, t(75) = -1.16, p = .25$]. Fi-

nally, a drop-off in detection performance as a function of saccade amplitude was observed in the away type condition [$R_{pb} = -.27, t(132) = -3.26, p < .005$, pooling observations from Experiments 1A and 1B] and a marginally reliable drop-off was observed in the away token condition [$R_{pb} = -.23, t(71) = -1.97, p = .052$]. This overall pattern of data is strikingly similar to the pattern that we observed in our earlier study, where we found both away deletion detection and toward and away rotation detection were reduced as saccade amplitude increased (Henderson & Hollingworth, 1999b). Across both that earlier study and the present study, change detection was relatively im-
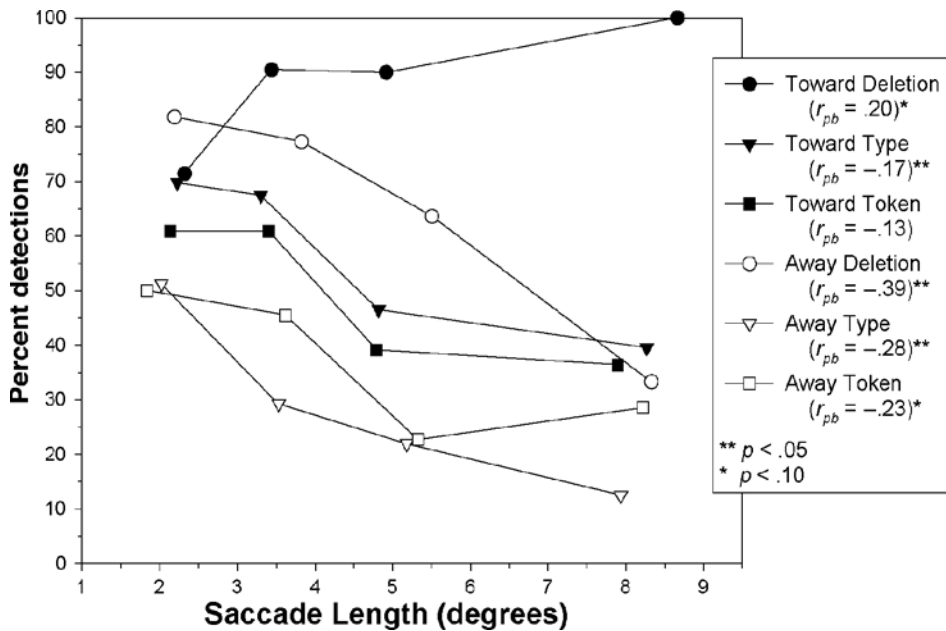
**Figure 3.** Change detection performance (percent detections) as a function of the length of the saccade triggering the change in Experiments 1A and 1B. For the type change conditions, data from Experiments 1A and 1B were combined. In each condition, the mean of each saccade length quartile is plotted against the mean percentage detections in that quartile.

pervious to saccade amplitude only in the toward deletion condition.

In addition to our focus on the nature of the saccade target deletion effect, the present study also gave us the opportunity to explore two additional questions concerning the nature of the scene representations that are generated and stored from fixation to fixation: How visually detailed are the object representations that are preserved in memory across saccades? To what degree does overt change detection completely reflect the underlying object representations preserved in visual memory?

**The visual specificity of object representations**. If participants can detect changes only on the basis of nonvisual information such as scene gist and object identities, then they should be able to detect type but not token changes. On the other hand, if visually specific representations can be preserved in scene memory, then detection of both type and token changes should be reliably above chance. In Experiment 1B, although performance in the token change condition was not perfect, it was well above the false alarm rate of 1.9%, both when the eyes were moving toward the critical object (48.8% correct) [$F(1,14) = 50.46, p < .001$] and away from the critical object (37.4% correct) [$F(1,14) = 25.37, p < .001$]. Because token changes did not alter the critical object's identity or basic-level semantic category, did not modify the overall gist of the scene, and did not change the spatial relations among the entities in the scene, but did change the visual details of the critical object, these data suggest that visually specific representations can be preserved across saccades. Importantly, token changes could be detected even when

the change took place during the saccade away from the changing object. In this away condition, attention would be directed away from the changing object (and toward the saccade target) prior to and following the change. Thus, these detections could not be based on continuously attending to the changing object during the change, a requirement for detection that has been proposed in the change blindness literature (Rensink, 2000a, 2000b; Rensink et al., 1997; see also Wolfe, 1999). The present results converge with other recent evidence suggesting that quite specific visual representations are retained across saccadic eye movements, as well as over longer periods of time during scene viewing, even in the absence of the continuous allocation of attention to the changing object (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001).[3]

**Covert change detection**. Recent evidence suggests that overt change detection does not adequately reflect the completeness of the underlying scene representation (Fernandez-Duque & Thornton, 2000; Williams & Simons, 2000).[4] For example, the time that the eyes remain fixated on a critical object is often increased by the presence of a change even when the change is not overtly reported (Hayhoe et al., 1998; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001). To investigate whether covert detection effects were present in the present study, and particularly whether such effects were present for the saccade target, we examined the degree to which changes would be registered in gaze durations given that a participant failed to press the change button. More specifically, gaze duration (the sum of all fixation durations

in an object region from entry to exit) was measured for the first entry of the eyes into the critical region following a change when that change was not explicitly reported. Miss trials were compared with the equivalent entry in the no-change condition. The results of this analysis are shown in Figure 4.

First, we examined miss trials in the toward condition, calculating gaze duration for the first entry of the eyes into the critical region. There were not enough miss trials to assess the toward deletion condition. For the toward type condition, data from Experiments 1A and 1B were combined, treating experiment as a between-participants factor. Two participants were excluded from this analysis due to an empty cell for misses. Mean gaze duration was 655 msec for toward type misses and 429 msec for the no-change control, a reliable difference of 226 msec [$F(1,26) = 12.84, p < .005$]. For the toward token condition, taken from Experiment 1B, 2 participants were again excluded due to an empty cell for misses. Mean first entry gaze duration was 668 msec for toward token misses, 229 msec longer than the mean for the no-change control (439 msec), a difference that approached reliability [$F(1,12) = 3.59, p = .08$]. Second, we examined miss trials in the away condition, calculating gaze duration for the second entry of the eyes into the critical region (i.e., the first entry after the change). Again, there were too many empty cells to assess the away deletion condition. For the away type condition, data were pooled across experiments. In order to maintain equal $n$ across the two experiments, an empty cell for 1 participant in Experiment 1B
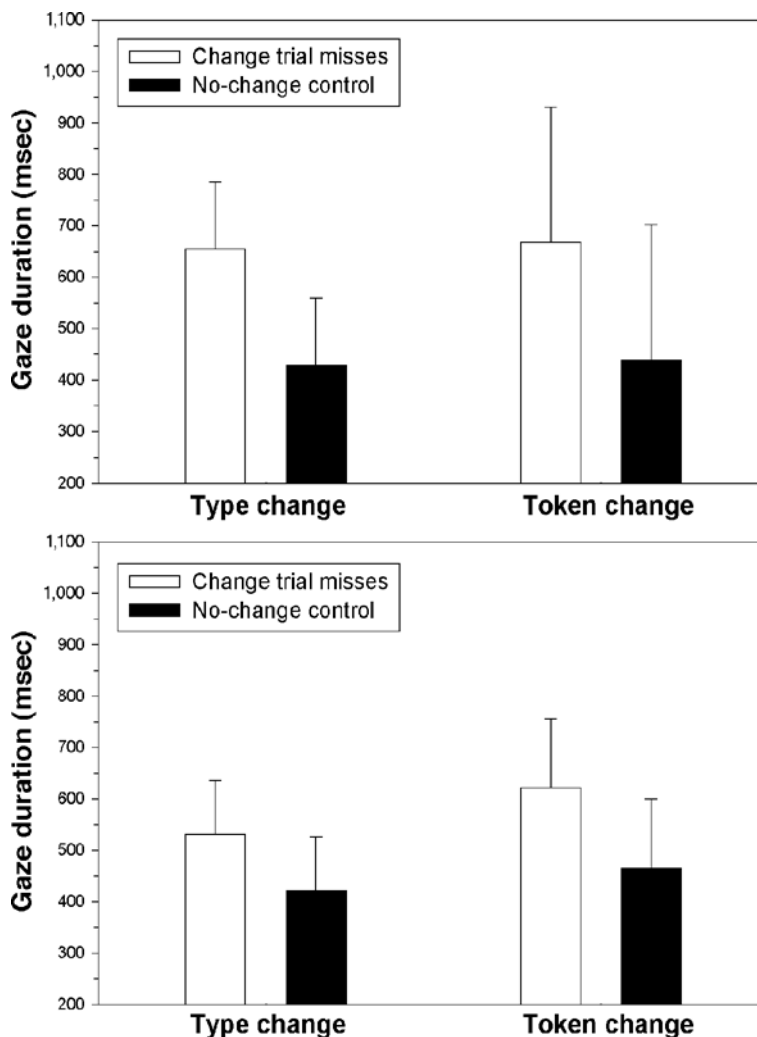


Figure 4. Gaze duration (in milliseconds) for the first entry of the eyes into the critical region after the change in the toward condition (top panel) and the away condition (bottom panel). Mean gaze duration for miss trials in each of the change conditions is contrasted to the corresponding mean gaze duration for the no-change control. Error bars are 95% confidence intervals based on the error term for each contrast. Type change data are collapsed across Experiments 1A and 1B.

was replaced by the mean of the other participants in that condition. Mean gaze duration was 532 msec for away type misses and 421 msec for the no-change control, a reliable difference of 111 msec [$F(1,28) = 4.25, p < .05$]. For the away token condition, 1 participant was excluded from the analysis due to an empty cell for misses. Mean gaze duration was 622 msec for away token misses and 465 msec for the no-change control, a reliable difference of 157 msec [$F(1,13) = 6.33, p < .05$]. Overall, these results provide a strong replication of similar effects of change on gaze duration in the absence of overt detection (Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001). The present data extend these earlier studies by demonstrating that effects of change on gaze duration in the absence of overt detection can be observed when the change occurs for the saccade target during the first saccade toward that target (toward condition). Overall, these results support the view that overt change detection is not fully representative of the completeness of the representations that are generated and retained across saccades (Henderson & Hollingworth, in press; Hollingworth & Henderson, 2002). An alternative explanation for the increase in gaze durations on changed objects has recently been suggested.[5] This explanation goes as follows: First, assume that during scene memorization, those objects that are fixated for less time the first time they are encountered (first pass) tend to be fixated for more time the second time they are encountered (second pass). This might be considered the "conservation of total encoding time" assumption. Second, assume that changes to objects that were fixated longer in the first pass are more likely to be overtly detected. The latter relationship was initially reported by Hollingworth and Henderson (2002) and was replicated here: In the present study, there was a reliable positive correlation between gaze duration and detection performance in the away type condition [$R_{pb} = .21$, $t(162) = 2.78, p < .01$] and the away token condition [$R_{pb} = .23, t(86) = 2.21, p < .05$]. In the away deletion condition, there was no relationship [$R_{pb} = -.17, t(86) = -1.62, p = .11$].

Putting these assumptions together, a selection artifact could be producing the gaze duration effect. In this explanation, when an object is fixated for a relatively short amount of time during first pass, a change to that object is less likely to be overtly noticed during second pass. At the same time, gaze duration on that object during second pass will be relatively long due to the conservation of processing time. When an object is fixated for a relatively long time during first pass, on the other hand, a change to that object is more likely to be noticed, eliminating what would otherwise be short gazes (due to conservation of processing time) from the data set. The result will be that second-pass gaze durations in the change condition will have fewer short fixations than would normally be part of a second-pass gaze duration distribution. In contrast, in the no-change condition, no second-pass fixations are eliminated, so the entire distribution of gaze durations expected on the basis of the conservation of fixation time assumption will be included. The consequence will be longer mean gaze durations in the change condition. If this line of reasoning is correct, then the increased gaze durations we have observed in miss trials do not provide evidence for the preservation of an underlying scene representation.

There are at least three sources of evidence that argue against this explanation for the observed covert change detection effects on gaze duration. First, the reliable covert effects in the toward type condition in the present study cannot be accounted for by this artifact explanation, because the covert effect is derived from *first-pass* gaze durations on these objects. Second, the conservation of total encoding time assumption predicts a negative correlation between first-pass and second-pass gaze durations on an object. In a direct test of this prediction, we have previously found that there is either no relationship, or a positive relationship, between first- and second-pass gaze duration on an object in a scene (Hollingworth & Henderson, 2002). To examine this relationship in the present study, we examined gaze durations in the control (no-change) conditions in Experiments 1A and 1B. In Experiment 1A, we observed a marginally reliable positive correlation between first-pass and second-pass gaze durations [$R_{pb} = .19, t(74) = 1.68, p = .097$]. In Experiment 1B, the correlation was negligible and not reliable [$R_{pb} = -.04$, $t(68) = -.37, p = .72$]. Thus, again, we find no support for the negative correlation that is assumed in the artifact explanation. Third, we have found that objects that produce longer first-pass gaze durations due to their semantic consistency within a scene also produce longer second-pass gaze durations (Henderson, Weeks, & Hollingworth, 1999). These results suggest that both first- and second-pass gaze durations are influenced similarly by cognitive processing difficulty, and that the gaze duration relationship across encounters is positive, not negative. Note that this positive relationship would tend to mask a covert gaze duration effect for the same reason that the conservation of processing time assumption might produce an artifactual one. Thus, our confidence is increased that the elevated gaze durations observed in the miss trials are a robust result of covert detection of the change.

## DISCUSSION

A central question in visual cognition is the nature and completeness of the scene representation that is generated over time and across successive glimpses (Henderson & Hollingworth, 1999b). One important technique for investigating this question is to examine a viewer's ability to detect a scene change that takes place during a saccadic eye movement (e.g., Currie, McConkie, Carlson-Radvansky, & Irwin, 2000; Grimes, 1996; Hayhoe et al., 1998; Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001; McConkie & Currie, 1996). In an earlier study we reported that viewers are more sensitive to the deletion of a saccade target than to the rotation of a saccade target or to the deletion of an object that was not the saccade target

(Henderson & Hollingworth, 1999b). Furthermore, the ability of a viewer to detect the deletion of a saccade target was relatively unaffected by the amplitude of the saccade to that target. This insensitivity to saccade amplitude for target deletions contrasted with the effect of amplitude for other types of target changes, which showed a decrease in detection performance as saccadic amplitude increased (Henderson & Hollingworth, 1999b), suggesting that there is something special about the coding of saccade target presence.

In the present study, we used a viewer's sensitivity to the deletion of a saccade target as a tool to investigate the nature of the information that is retained and combined across saccadic eye movements during complex scene perception. We manipulated two factors, the location of the changing object with respect to the direction of the saccade at the time of the change (toward or away from the target) and the nature of the change (type change, token change, or deletion). The primary question was whether saccade target presence per se is special, as the saccade target deletion effects suggest, or whether instead other sorts of changes to a saccade target are also more easily detected and show relative insensitivity to saccade distance, as should be the case if properties of the saccade target other than presence are the basis for the deletion effect.

The results of the present study were quite clear. First, deletions were better detected when the deleted object was the target of a saccade than when the deleted object had just been fixated but was not the target of a saccade at the time of the change. Second, saccade target deletions were better detected than either saccade target type or token changes. Third, the amplitude of the saccade taking the eyes to the target did not adversely affect the detection of saccade target deletions. The latter finding replicates our previous study (Henderson & Hollingworth, 1999b) and contrasts with the clear effect observed here of saccade amplitude on the detection of type and token changes for the saccade target. The invariance of deletion detection for the saccade target with amplitude also contrasts with the clear effect of amplitude on deletion, as well as type and token changes, for saccades away from the changing object. The finding that type changes were considerably more difficult to detect than deletions, and that they did not show the same invariance with saccade amplitude exhibited by deletions in the toward condition, strongly suggests that it is not the retention of the saccade target's identity or basic-level semantic category that underlies the special status of the saccade target. Instead, these results suggest that there is something special about the coding and retention across a saccade of saccade target presence itself. One theoretical account of transsaccadic visual perception that places special emphasis on the target of a saccade is the *saccade target theory of visual stability*. We next consider our results in the context of that theory.

## Saccade Target Theory of Visual Stability

According to the saccade target theory of visual stability (Currie et al., 2000; Irwin, McConkie, Carlson-Radvansky, & Currie, 1994; McConkie & Currie, 1996),

an object in the visible scene is selected as the saccade target prior to each saccade. The location of that target is coded within an internal representation of the scene and retained in visual short-term memory, along with features of the object that allow it to be found after the saccade (*locating information*). After the saccade, the visual system engages in a search for the locating information. This search is constrained to a limited *initial search region* around the landing position following the saccade (McConkie & Currie, 1996). Once the saccade target is located in the current image, it provides the basis for remapping the stored scene representation to the present retinal input, which leads to the sense of visual stability experienced across saccades.

Although saccade target theory was proposed specifically to account for the experience of visual stability, it might also provide an explanation for the special nature of saccade target deletions in change detection paradigms. Specifically, the toward deletion effect could be accommodated by the theory with the additional assumption that the locating information generated from the presaccade image includes a content-free representation of target presence; for example, the saccade target might be coded simply as a blob, a point representing the target's center of gravity, or a FINST-like positional index (Pylyshyn, 1989). If a presence marker such as this were created for the saccade target prior to a saccade, and were then checked against a similar content-free representation of presence at fixation following the saccade, then the deletion of the saccade target would be relatively salient because it would lead to an inability to find the locating information following the saccade. In contrast, in both the type and token changes, the locating information (the content-free presence marker) would still be available after the saccade and stability processes could proceed.

This version of saccade target theory suggests that it is the content-free coding of the saccade target alone that is important in mapping the pre- and postsaccade images. A problem with this view is that in the case of deletions, there are typically other objects nearby that might (to the system seeking the locating information) be indistinguishable from the saccade target. In other words, given the type of scheme just outlined, the visual system would be faced with a correspondence problem across the saccade in trying to determine which of several content-free position markers identified following a saccade should be tied to the position marker coded prior to the saccade. One way around this problem, and a possibility that is more in keeping with current work on the representation of spatial information across views, is that the presence of the saccade target is coded with respect to its place in a spatial configuration that includes the positions of other nearby objects. As before, to accommodate the special status of presence per se, the hypothesis would be that the markers coding the positions of the saccade target and other nearby salient objects are abstracted away from other visual or semantic information. Unlike the single-marker hypothesis, though, these markers together would be combined into a content-free representation of the configuration of the tar-

get and its nearby neighbors. If a representation such as this were created for the saccade target and its neighbors prior to a saccade, and were then checked against a similar representation following the saccade, then the deletion of the saccade target would be salient because it would change the overall configuration of the set of markers tied to the saccade target. Changes to the visual features or identity of the saccade target (or the neighboring objects that are coded in the configuration), in contrast, would not change the configuration. Furthermore, in this view, the detection of changes to a configuration of positions should be less affected by saccade amplitude than should the detection of changes to visual form or meaning, because the encoding of the presence of the target and its neighbors does not require that the same level of visual detail be resolvable.

The hypothesis we are proposing, then, is that an allocentric representation of the target's position with respect to other nearby objects is preferentially used as the locating information to map the new retinal information acquired after a saccade to the internal scene representation that was generated prior to that saccade. This emphasis on allocentric spatial representation across saccades is consistent with other work showing that during transsaccadic object identification, visual properties are bound to position defined by an allocentric reference frame (Henderson, 1994; Henderson & Anes, 1994; Henderson & Siefert, 2001). It is also consistent with recent evidence that the information stored in visual short-term memory from one view is preferentially accessed via information about spatial configuration rather than information about other visual characteristics such as color or shape (Jiang et al., 2000). Importantly, the latter finding has been shown to generalize to transsaccadic memory (Carlson, Covell, & Warapius, 2001), providing additional evidence that the process that compares pre- and postsaccade representations preferentially relies on spatial configuration rather than on visual or semantic content for initial access. It is important to note, however, that spatial configuration is not the only type of information maintained and integrated across saccades (see also the next section, below), only that a representation of configuration plays a significant primary role, perhaps providing the retrieval key used to access other visual and semantic information (Henderson, 1994; Henderson & Anes, 1994; Henderson & Siefert, 2001; see also Carlson et al., 2001; Jiang et al., 2000; Pylyshyn, 2000).

We note that one result in the literature appears to call into question saccade target theory's primary assumption that the saccade target is preferentially used to map pre- and postsaccade images. This result was found in a transsaccadic change detection paradigm in which the viewer was asked to execute a saccade from one point-light walker to another, both of which were displayed on a computer monitor (Verfaillie & De Graef, 2000). During the saccade, the position or the in-depth orientation of either the saccade source (the walker that is fixated prior to the saccade) or the saccade target (the walker that is fixated following the saccade) was changed. Detection of a change in position and orientation was found to be equivalent for the two walkers. This finding is at odds with the clear advantage for the saccade target observed in the present study as well as in Henderson and Hollingworth (1999b). One possible explanation for the difference in results is based on the nature of the attentional and saccade dynamics in ongoing scene perception versus in a single-saccade paradigm. In the present study, the viewer was engaged in an ongoing series of fixations and saccades during temporally extended scene viewing. In contrast, in the study reported by Verfaillie and De Graef, viewers were executing a single saccade between two objects. In the former case, the viewer may be basing responses on representations that are naturally generated by the perceptual system as it ties together successive views. In the latter case, participants may be able to overcome the natural attention-saccade dynamics and strategically allocate attention to (and encode and compare) the saccade source and target objects equivalently on each trial. This hypothesis makes two clear predictions. First, a saccade target advantage should be observed for point-light walkers (or any other objects) if those walkers were to be viewed during an ongoing series of fixations and saccades executed as part of extended scene perception. Second, equivalent source and target performance should be observed for objects in natural scenes of the sort used in the present study if participants were to make a single saccade between the source and target and then indicate if either had changed. Both of these predictions await empirical test.

## Visual Memory Theory and Change Detection

In addition to the special coding of the saccade target's presence, we also have good evidence that other properties of objects in a scene, including properties that allow a viewer to distinguish one visual token from another, can be retained and compared across eye movements and over longer periods of time. To account for this fact, we have recently proposed a *visual memory theory* of dynamic scene representation (Hollingworth & Henderson, 2002; see also Henderson & Hollingworth, in press). According to visual memory theory, a relatively detailed scene representation is built up in memory across eye fixations. The scene representation is retained both over the shorter term in short-term memory (e.g., Irwin, 1992; Irwin & Andrews, 1996) and over the longer term in long-term scene memory (e.g., Shepard, 1967; Standing, Conezio, & Haber, 1970; see also Friedman, 1979). Importantly, these scene representations are not to be construed as sensory in nature. Instead, we draw a distinction between sensory representations and abstract visual representations. In our view, the representations retained and integrated across saccades can be visually specific, but abstract (see Henderson & Hollingworth, in press; Hollingworth & Henderson, 2002). We take an *abstract visual representation* to be a nonmaskable and non-iconic visual description encoded in the vocabulary of visual computation. Abstract visual representations are visual in the sense that they represent visual properties such as object shape. For example, structural descriptions (e.g., Biederman, 1987; Marr, 1982; Palmer, 1977) and hierarchical feature representations

(e.g., Riesenhuber & Poggio, 1999) are examples of abstract visual representational systems that have been proposed to underlie object recognition. Recent evidence suggests that structural descriptions may form at least part of the representation of object shape that is retained across saccades (Carlson-Radvansky, 1999; Carlson-Radvansky & Irwin, 1995). Such representations are not equivalent to conceptual representations, which code semantic properties of the viewed scene, nor are they linguistic descriptions of scene properties.

Succinctly, in visual memory theory, the detection of a change to an object in a scene is a function of initial attention to and encoding of a representation of the prechanged object, retention of that representation either in an active state in short-term memory and/or in an inactive state in long-term memory, generation of a new representation following the change to compare with the stored representation, and retrieval of the stored representation from long-term memory if it is not currently active in short-term memory. In visual memory theory, the initial allocation of attention to an object gates sensory processing of that object and leads to the generation of (1) an abstract representation of the object's visual properties; (2) a representation of the object's position, including its allocentric position with respect to other nearby objects; and (3) semantic representations and an identity code. A limited number of these representations can be actively maintained in a limited-capacity short-term memory store (Irwin & Andrews, 1996), perhaps as integrated object representations (Carlson et al., 2001; Luck & Vogel, 1997) or "object tokens" (Henderson, 1994; Henderson & Anes, 1994; Henderson & Siefert, 2001). Furthermore, processing in short-term memory leads to consolidation of both visual and semantic representations as part of the overall scene representation, and to storage of this information into a more stable long-term memory representation of the scene (Hollingworth & Henderson, 2002).

In this view, transsaccadic processing proceeds as follows: Prior to a saccade, attention is obligatorily allocated to the saccade target (e.g., Deubel & Schneider, 1996; Henderson, 1996; Henderson et al., 1989; Hoffman & Subramaniam, 1995; Irwin & Andrews, 1996; Kowler, Anderson, Dosher, & Blaser, 1995; Rayner et al., 1978; Shepherd et al., 1986; see also Sheinberg & Logothetis, 2001, for recent single-unit work in macaque demonstrating preferential encoding of the saccade target in natural scenes). Sensory processing of the saccade target is gated, the representations just discussed are generated and stored in short-term memory for that object, and processes of consolidation and transfer to long-term memory are initiated. Once the eyes begin to move, the sensory representations quickly decay (Sperling, 1960), leaving only the abstracted visual and semantic representations in short-term and long-term memory. When the eyes land, the stored representations are compared with information encoded in the current fixation. If the information is different, an error signal is generated and the change is noted (either overtly or covertly). Otherwise, the fixated information is integrated into the current scene representation. Because these stored representations include both abstract visual and semantic information, changes to form, meaning, and identity can all be detected. The fact that saccade target deletions are particularly salient can be accommodated by positing that the locating information preferentially used to map the new retinal input to the stored representation is the saccade target's position with respect to other local objects, consistent with saccade target theory as just described as well as with the type token theory of transsaccadic object integration (Henderson, 1994; Henderson & Anes, 1994; Henderson & Siefert, 2001).

Because the information needed to detect a change can be retrieved from long-term memory, changes to objects that are not the target of a saccade can also be detected as long as the information relevant to the change was successfully consolidated and stored during a previous fixation, and as long as that information is retrieved from memory and compared with the present stimulus (e.g., Hollingworth & Henderson, 2002). Redirecting attention back to the changed scene region following the change greatly increases the probability that the change will be detected because (1) it ensures that the postchange object is encoded, and (2) it increases the probability that information about the original version of that object will be retrieved from long-term memory. This occurs because local information in the scene provides a strong retrieval cue. Even if the object has been deleted from the region, the spatial position of fixation and the coding of nearby scene information can provide a retrieval cue for the missing object (Henderson & Hollingworth, 1999b). This basic assumption accounts for the strong tendency for late detections in the change detection paradigm to take place only after the changed object has been refixated (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001; see Henderson & Hollingworth, in press, for review).

Visual memory theory contrasts with what we have called *localist-minimalist* theories of scene representation, in which scene representations consist of transient visual representations of attended objects and nonvisual representations of scene gist, spatial layout, and conceptual information outside of the focus of attention (Rensink, 2000a, 2000b; Wolfe, 1999). On the one hand, localist-minimalist theories might not have particular difficulty accounting for the good change detection performance at the saccade target location, given the strong evidence that attention precedes a saccade to the saccade target location. That is, good saccade target change detection could be explained by assuming that attention is allocated to the saccade target both before and after the saccade, and that this continuous allocation of attention helps maintain a visual representation of the target. However, even in this case, such representations would have to be abstract, given the preponderance of evidence that precise sensory images cannot be retained and integrated across saccades (e.g., Irwin, 1992; Pollatsek & Rayner, 1992). Localist-minimalist theories have a more difficult time accounting for performance when the change takes place during a saccade away from the target, because in that case attention is allocated to a

nonchanging object (the target of the saccade) immediately before and after the saccade. The fact that viewers can detect token changes and rotations in the away condition, even when those detections come several seconds and many fixations after the change takes place (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002), cannot easily be accommodated by localist-minimalist theories, but can be naturally explained by visual memory theory.

Finally, in the present study we also found that overt change detection performance underestimated the degree to which object representations were retained in memory across fixations. We operationally defined *covert detection* as an increase in gaze duration on a critical object when it had changed relative to a control condition in which it had not changed, for those trials in which the viewer did not overtly respond to the change with a buttonpress. We found that gaze durations increased when a change had taken place but was not overtly reported, replicating our prior results (Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001; see also Hayhoe et al., 1998) and extending them to situations in which the changed object was the saccade target. Furthermore, these effects were not due to a selection artifact based on the initial fixation time spent on the critical objects. Finally, evidence for covert change detection was observed both for type changes and for token changes. The latter results provide additional evidence that relatively detailed visual representations of the objects in a scene are retained over time and across multiple eye fixations.

## REFERENCES

BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.

BREITMEYER, B. G., KROPFL, W., & JULESZ, B. (1982). The existence and role of retinotopic and spatiotopic forms of visual persistence. *Acta Psychologica*, **52**, 175-196.

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

CARLSON, L. A., COVELL, E. R., & WARAPIUS, T. (2001). Transsaccadic coding of multiple objects and features. *Psychologica Belgica*, **41**, 9-27.

CARLSON-RADVANSKY, L. A. (1999). Memory for relational information across eye movements. *Perception & Psychophysics*, **61**, 919-934.

CARLSON-RADVANSKY, L. A., & IRWIN, D. E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1441-1458.

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: Marcel Dekker.

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

CURRIE, C. B., McCONKIE, G. W., CARLSON-RADVANSKY, L. A., & IRWIN, D. E. (2000). The role of the saccade target object in the perception of a visually stable world. *Perception & Psychophysics*, **62**, 673-683.

DENNETT, D. C. (1991). *Consciousness explained*. Boston: Little, Brown.

DEUBEL, H., & SCHNEIDER, W. X. (1996). On the nature of the span of apprehension. *Psychological Research*, **55**, 29-39.

DUHAMEL, J. R., COLBY, C. L., & GOLDBERT, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, **255**, 90-92.

FERNANDEZ-DUQUE, D., & THORNTON, I. M. (2000). Change detection without awareness: Do explicit reports underestimate the representation of change in the visual system? *Visual Cognition*, **7**, 324-344.

FRIEDMAN, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, **108**, 316-355.

GRIMES, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception* (Vancouver Studies in Cognitive Science, Vol. 5, pp. 89-110). Oxford: Oxford University Press.

HAYHOE, M. M., BENSINGER, D. G., & BALLARD, D. H. (1998). Task constraints in visual working memory. *Vision Research*, **38**, 125-137.

HENDERSON, J. M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General*, **123**, 410-426.

HENDERSON, J. M. (1996). Visual attention and the attention-action interface. In K. Aikens (Ed.), *Perception* (Vancouver Studies in Cognitive Science, Vol. 5, pp. 290-316). Oxford: Oxford University Press.

HENDERSON, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, **8**, 51-55.

HENDERSON, J. M., & ANES, M. D. (1994). Effects of object-file review and type priming on visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 826-839.

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). New York: 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. (in press). Perception of faces, objects, and scenes: Analytic and holistic processes. In M. A. Peterson & G. Rhodes (Eds.), *Eye movements, visual memory, and scene representation*. New York: Oxford University Press.

HENDERSON, J. M., McCLURE, K., PIERCE, S., & SCHROCK, G. (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception & Psychophysics*, **59**, 323-346.

HENDERSON, J. M., POLLATSEK, A., & RAYNER, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics*, **45**, 196-208.

HENDERSON, J. M., & SIEFERT, A. B. C. (1999). The influence of enantiomorphic transformation on transsaccadic object integration. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 243-255.

HENDERSON, J. M., & SIEFERT, A. B. C. (2001). Types and tokens in transsaccadic object integration. *Psychonomic Bulletin & Review*, **8**, 761-768.

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.

HOFFMAN, J. E., & SUBRAMANIAM, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, **57**, 787-795.

HOLLINGWORTH, A., & HENDERSON, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 113-136.

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

HOLLINGWORTH, A., WILLIAMS, C. C., & HENDERSON, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, **8**, 761-768.

IRWIN, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, **23**, 420-456.

IRWIN, D. E. (1992). 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. (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* (pp. 125-155). Cambridge, MA: MIT Press.

IRWIN, D. E., McCONKIE, G. W., CARLSON-RADVANSKY, L. A., & CURRIE, C. (1994). A localist evaluation solution for visual stability across saccades. *Behavioral & Brain Sciences*, **17**, 265-266.

JIANG, Y., OLSON, I. R., & CHUN, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 683-702.

KOWLER, E., ANDERSON, E., DOSHER, B., & BLASER, E. (1995). The role of attention in the programming of saccades. *Vision Research*, **35**, 1897-1916.

LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.

MARR, D. (1982). *Vision.* San Francisco: W. H. Freeman.

McCONKIE, G. W., & CURRIE, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 563-581.

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

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., RENSINK, R. A., & CLARK, J. J. (1999). Change blindness as a result of "mudsplashes." *Nature*, **398**, 34.

PALMER, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, **9**, 441-474.

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.

PYLYSHYN, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, **32**, 65-97.

PYLYSHYN, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, **4**, 197-207.

RAYNER, K., McCONKIE, G. W., & EHRLICH, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 529-544.

RENSINK, R. A. (2000a). The dynamic representation of scenes. *Visual Cognition*, **7**, 17-42.

RENSINK, R. A. (2000b). Seeing, sensing, and scrutinizing. *Vision Research*, **40**, 1469-1487.

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.

RIESENHUBER, M., & POGGIO, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, **2**, 1019-1025.

SHEINBERG, D. L., & LOGOTHETIS, N. K. (2001). Noticing familiar objects in real world scenes: The role of temporal cortical neurons in natural vision. *Journal of Neuroscience*, **21**, 1340-1350.

SHEPARD, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning & Verbal Behavior*, **6**, 156-163.

SHEPHERD, M., FINDLAY, J. M., & HOCKEY, R. J. (1986). The relationship between eye movements and spatial attention. *Quarterly Journal of Experimental Psychology*, **38A**, 475-491.

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

SIMONS, D. J., & LEVIN, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, **5**, 644-649.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74** (11, Whole No. 498).

STANDING, L., CONEZIO, J., & HABER, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2,500 visual stimuli. *Psychonomic Science*, **19**, 73-74.

VERFAILLIE, K., & DE GRAEF, P. (2000). Transsaccadic memory for position and orientation of saccade source and target. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1243-1259.

WILLIAMS, P., & SIMONS, D. J. (2000). Detecting changes in novel 3D objects: Effects of change magnitude, spatiotemporal continuity, and stimulus familiarity. *Visual Cognition*, **7**, 297-322.

WOLFE, J. M. (1999). Inattentional amnesia. In V. Coltheart (Ed.), *Fleeting memories* (pp. 71-94). Cambridge, MA: MIT Press.

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

## NOTES

1. We have previously found that when detection does not take place within 1,500 msec of the change, detection either does not occur at all or occurs only once the changed region is refixated (Henderson & Hollingworth, 1999b). Therefore, it appears that 1,500 msec provides a reasonably conservative cutoff value for "immediate" detections.

2. In this and subsequent regression analyses, we regressed saccade amplitude against the dichotomous detection variable (yielding a point-biserial coefficient). Each trial was treated as an observation. Since each participant contributed more than one sample to the analysis, variation due to differences in participant means was removed by including participant as a categorical factor (implemented as a dummy variable) in the model.

3. If meaning in addition to visual information is retained in transsaccadic memory, then one might expect that type changes (which change both meaning and visual information) would be detected better than token changes (which maintain meaning). In the present study, the detection rate for type changes was no greater than for token changes ($F < 1$ collapsed over the toward and away conditions), nor did the two factors (type/token $\times$ toward/away) interact [$F(1,14) = 1.261$, $MS_e = 0.0640$, $p = .28$]. This result could be taken to suggest that *only* visual information is preserved. However, this conclusion is unwarranted because the degree of visual change in the type and token change conditions was not controlled.

4. We use here *overt* versus *covert* detection, rather than *explicit* versus *implicit* detection, because the latter terms are evocative of the distinction between explicit and implicit memory and so may be taken to suggest a proposal of separate underlying functional and neural systems. We do not want to imply a commitment to the theoretical stance that overt and covert change responses reflect separate change detection systems. Gaze duration effects may reflect trials on which participants detect the change but are not confident enough to respond positively. Alternatively, these effects might be based on representations that are available to motor systems (in this case the oculomotor system) but not to perceptual or other decision processes. This is an issue that awaits further empirical investigation. For now, we use here what we hope are more theoretically neutral terms.

5. We thank Dan Simons for bringing this possibility to our attention.