

Research Article

GLOBAL TRANSSACCADIC CHANGE BLINDNESS DURING SCENE PERCEPTION

John M. Henderson¹ and Andrew Hollingworth²

¹*Department of Psychology and Cognitive Science Program, Michigan State University, and*

²*Department of Psychology, Yale University*

Abstract—*Each time the eyes are spatially reoriented via a saccadic eye movement, the image falling on the retina changes. How visually specific are the representations that are functional across saccades during active scene perception? This question was investigated with a saccade-contingent display-change paradigm in which pictures of complex real-world scenes were globally changed in real time during eye movements. The global changes were effected by presenting each scene as an alternating set of scene strips and occluding gray bars, and by reversing the strips and bars during specific saccades. The results from two experiments demonstrated a global transsaccadic change-blindness effect, suggesting that point-by-point visual representations are not functional across saccades during complex scene perception.*

Human vision is dynamic and active: On average, observers move their eyes three times each second in very fast saccadic eye movements. Between the saccades are fixations, each lasting about 300 ms on average, during which visual information is acquired from the environment (see Henderson & Hollingworth, 1999a, for review). Recent research has shown that a change to a specific region of a scene often goes unnoticed when that change takes place during a saccade (Grimes, 1996; Henderson & Hollingworth, 1999b; McConkie & Currie, 1996). Similar effects are observed if a scene element changes during a visually blank period of time that simulates a saccade (Rensink, O'Regan, & Clark, 1997) or is otherwise masked or hidden from view at the time of the change (e.g., Simons & Levin, 1998). This *change-blindness* effect is striking because it seemingly undermines a long-standing assumption in vision science that the visual system constructs a complete and integrated representation of the visual world across glimpses. Furthermore, the effect has been taken to call into question the intuition that perceptual experience directly reflects the nature of the underlying visual representation; instead, change blindness appears to indicate that observers' experience of a complete and detailed visual world is based on what is in fact a sparse and incomplete visual representation (Dennett, 1991).

Recent theoretical treatments of scene perception based on the change-blindness effect have converged on two assumptions concerning visual representation. First, all forms of visual representation of a scene element are assumed to be lost once attention is withdrawn from that element (Rensink, 2000a, 2000b; Wolfe, 1999). This assumption explains change blindness: Changes are missed because visual representations of scene elements are not retained and accumulated during viewing. Instead, a scene representation is taken to consist only of a transient visual representation of the currently attended element, the se-

mantic gist of the scene, gross spatial layout of scene elements, and an inventory of object concepts and identities (Irwin, 1991; Rensink, 2000a, 2000b; Wolfe, 1999). Second, it is assumed that a highly detailed, point-by-point representation is created for the currently attended scene element (e.g., object), and that this veridical point-by-point representation can survive a transient disruption (e.g., a saccade) as long as attention remains directed toward that element (Rensink, 2000a, 2000b). This assumption is taken to explain the fact that once a change has been detected, it is thereafter very apparent (Rensink, 2002; Rensink et al., 1997; Simons 2000).

Contrary to the first of these assumptions, recent evidence has demonstrated that relatively detailed visual representations of objects in scenes can in fact be retained after attention has been withdrawn. These representations are visual in the sense that they encode visual properties of objects, such as viewpoint-specific orientation and the visual detail necessary to identify specific members of a general object class (Henderson & Hollingworth, 1999b, 2003a; Hollingworth, in press; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001). The conclusion that these visual representations are retained after the withdrawal of attention is based on three main findings. First, change blindness for object orientation and token replacement is reduced or even eliminated when care is taken to ensure that a changing object is attended both before and after the change, even if attention has been withdrawn in the meantime (e.g., Henderson & Hollingworth, 1999a, 2003a; Hollingworth, in press; Hollingworth & Henderson, 2002). Second, in those cases in which change blindness persists when the changing object is attended both before and after the change, more subtle measures of change detection, such as fixation duration, reveal that the change has been noted by the visual system at some level (Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002; Hollingworth et al., 2001; see also Hayhoe, Bensinger, & Ballard, 1998). Third, when memory is directly probed following scene viewing, evidence for the retention of visual representations can be found over retention spans of several minutes (Hollingworth & Henderson, 2002).

The present study was designed to test the second assumption: that point-by-point sensory representations of a scene element can be retained across a saccade as long as attention is continuously allocated to that scene element during the saccade (Rensink, 2000a, 2000b, 2002). Prior studies have demonstrated that simple visual patterns cannot be visually integrated across saccades (e.g., Irwin, 1991, 1992). Similarly, the contours of single objects cannot be integrated across saccades (Henderson, 1997). However, the degree to which point-by-point representations are functional across saccades during natural scene perception has not yet been investigated.

In the present study, we examined this issue using a saccade-contingent global display-change paradigm. Viewers studied pictures of complex real-world scenes while their eye movements were recorded. Each scene was presented as an alternating series of scene

Address correspondence to John M. Henderson, Psychology Research Building, Michigan State University, East Lansing, MI 48824-1117; e-mail: john@eyelab.msu.edu.

strips and occluding gray bars (Fig. 1). Eye movements were monitored with a fast and accurate eyetracker, and each time a viewer's direction of gaze crossed an invisible software-defined boundary, the display changed such that the occluding gray bars revealed the previously occluded scene image, and the scene image previously visible became occluded by gray bars. These display changes took place during the saccade, when vision is suppressed (Thiele, Henning, Kubischik, & Hoffman, 2002; Volkman, Schick, & Riggs, 1968), and were completed before the onset of the next fixation so that the changes could not be detected from visual transients. The occlusion manipulation ensured that every pixel and every visual feature in the image was replaced each time a change took place, but that higher-level visual properties of the scene, such as depicted surface color, viewpoint, direction of lighting and shadow, object shapes, and spatial relationships among scene elements, remained the same across the change. Viewers were instructed to memorize the scenes for a later memory test, and to indicate immediately via button press each time the image changed. The exact nature of the change that might occur was described. Subjects were told that on some trials more than one change would occur, and that they should respond as quickly as possible to each change while withholding response when no change occurred. In Experiment 1, the region boundaries divided each of the scenes into three equally sized regions. In Experiment 2, the boundaries were placed around three specific objects in each scene so that both overt (button press) and covert (fixation duration) measures of change detection could be examined.

During a fixation-saccade-fixation sequence generated as gaze shifts from one object to another, attention-saccade dynamics ensure that a single location is attended immediately prior to and immediately following the saccade. That is, prior to a saccade, attention mandatorily and exclusively precedes the eyes to the target of that impending saccade; the eyes then move to the attended location, and attention and fixation are recoupled (Deubel & Schneider, 1996; Henderson, Pollatsek, & Rayner, 1989; Hoffman & Subramanian, 1995; Irwin & Andrews, 1996; Kowler, Anderson, Doshier, & Blaser, 1995; Rayner, McConkie, & Ehrlich, 1978; Shepherd, Findlay, & Hockey, 1986). Therefore, attention is oriented to and maintained at a single location before and after each saccade (the saccade target prior to the saccade and that same location once it is foveally acquired following the saccade). If a point-by-point representation is functional across saccades for attended scene elements, then the scene changes used in this study should be highly salient and easily detected. If instead only a higher-level visual representation of the scene, abstracted away from the point-by-point image, is functional across saccades even for attended scene elements, then these scene changes should not be detectable.

GENERAL METHOD

Subjects

Eighteen Michigan State University undergraduate students, naive with respect to the hypotheses under investigation, participated for course credit (10 in Experiment 1, 8 in Experiment 2).

Apparatus

The right eye was tracked using a Generation 5.5 SRI Dual Purkinje Image Eyetracker (spatial resolution < 1 arcmin, Crane & Steele, 1985) interfaced with a 90-MHz computer (sampling rate >

1000 Hz) driving a Hercules Dynamite 128/Video graphics card and NEC Multisync P750 monitor (refresh rate = 143 Hz, saccade-contingent change completed in 7–14 ms). Viewing was binocular, and the head was restrained with a bite bar and forehead rest. A shutter test (Irwin, 1994) demonstrated that display changes could not be detected on the basis of phosphor persistence (Henderson & Hollingworth, 1999a).

Stimuli

Pictures of real-world scenes (35 in Experiment 1, 36 in Experiment 2) subtending $15.8^\circ \times 11.9^\circ$ at a viewing distance of 1.13 m were rendered from three-dimensional wire-frame models. Rendered images were divided into 40 vertical strips, each subtending 0.395° . In one version of each scene, alternating bars were replaced with uniform gray in a square-wave pattern; the other version of each scene reversed the positions of the gray bars and the displayed scene.

Procedure and Design

Once the subject fixated a central box on a fixation screen, the experimenter started the trial. The initial version of a scene (A) was displayed until the subject's direction of gaze crossed one of two invisible boundaries (Experiment 1) or into one of three invisible object regions (Experiment 2). The scene then changed to the alternate version (A'). Viewing continued for a total of 10 s per scene, with images alternating between A and A' each time direction of gaze crossed a region boundary (Experiment 1) or the first time direction of gaze entered each object region (Experiment 2). No-change control trials were included (7 in Experiment 1, 18 in Experiment 2), with scenes assigned to change and no-change conditions via Latin square across subjects. Order of scenes and conditions was determined randomly for each subject. Subjects were instructed to view each scene in preparation for a memory test, and to press a response button immediately upon detecting a change; the nature of the change that might occur was described in detail. Subjects were told that more than one change would occur on some trials, and that no change would occur on others. They were to respond as quickly as possible for each change and to withhold response when no change occurred.

EXPERIMENT 1

In Experiment 1, the display changed each time the subject's direction of gaze crossed either of two invisible change-generating boundaries that divided each scene into three equally sized regions. The subject was instructed to press the response button immediately whenever a change was detected.

Each subject executed multiple saccades across the boundaries in each trial (Fig. 1a). Trials were eliminated when the change could not be completed before the eyes came to rest in fixation following the saccade that initiated the display change. Summing over the 10 subjects in the experiment, 1,691 useable scene changes took place (range = 106–234 across viewers), for an average of 6.48 changes per trial, or an average of 1 change every 1.54 s of viewing. Subjects detected only 45 out of these 1,691 changes, for an overall detection rate of 2.66%, with a range of 0% to 5.26% across subjects (0.0%, 0.74%, 1.0%, 1.89%, 2.56%, 3.14%, 3.47%, 4.0%, 4.48%, 5.26%). Thus, the results demonstrate nearly complete global change blindness.

EXPERIMENT 2

Although overt responses to image changes have sometimes been taken to reflect completely the degree to which viewers notice changes, more recent evidence has demonstrated that overt responses underestimate change sensitivity (Hayhoe et al., 1998; Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002; Hollingworth et al., 2001). Therefore, it is possible that although the point-by-point changes were not overtly reported in the first experiment, they may still have been detected. Fixation duration has been shown to be a sensitive measure of covert (i.e., unreported) change detection (Hayhoe et al., 1998; Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002; Hollingworth et al., 2001), but this measure could not be examined in Experiment 1 because there were very few no-change baseline trials, and because the use of large boundary regions did not allow for control of the position of fixations following boundary crossings, an important factor in determining fixation durations (Henderson & Hollingworth, 1999b).

Method

In Experiment 2, a modified version of the boundary paradigm was employed. In the change condition, the display change used in Experiment 1 took place whenever the viewer's direction of gaze entered any of three predefined object regions in each scene (Henderson & Hollingworth, 1999b; Hollingworth & Henderson, 2002; Hollingworth et al., 2001), as shown in Figure 1b. In addition, we equated the number of change and no-change trials to increase the power of the fixation-duration comparison. Only data from the first entry into each region were analyzed, as an additional control over fixation durations.

Results

Changes were eliminated from analysis when either (a) the change could not be completed before the eyes came to rest in fixation following the saccade that initiated the display change or (b) the saccade crossed the region boundary but the following fixation was not within the object region (i.e., the saccade passed through the region). The first criterion was more difficult to meet in this experiment than in Experiment 1 because of the restriction on the following landing position. Summing across the 8 viewers and 36 scenes, there were 318 useable change-inducing first region entries in the change trials (2.2 changes/trial on average), and 337 equivalent region entries in the no-change control trials (2.3 entries/trial on average). Five of the 8 viewers had detection rates that were similar to those observed in Experiment 1 (0%, 4.17%, 4.47%, 6.67%, 7.15%), 1 viewer had an elevated detection rate (13.65%), and 2 viewers detected even more of the changes (22.8% and 29.6%). Prior research using the occlusion manipulation with single objects had shown that viewers occasionally adopt a detection strategy based on abstract verbal coding (e.g., "the briefcase handle was present, and now it is not") that allows them to perform at higher-than-usual levels (Henderson, 1997). Debriefing suggested that the latter 3 viewers performed better than the others because they used such a strategy. It is clear that the remaining 5 viewers detected very few, if any, of the changes.

Covert detection was assessed by comparing fixation duration on the object within the change-inducing region in the no-change control trials and in the change trials in which the change was not overtly reported via button press. Only those trials in which the eyes came to rest inside the region were scored. The average duration of the first fixation in each of

the critical regions was no different in the (undetected) change and no-change conditions (319 ms vs. 318 ms, respectively, $F < 1$). Similarly, gaze durations (the sum of the durations of all fixations the first time each of the critical regions was entered) did not differ across these conditions (750 ms vs. 730 ms, respectively, $F < 1$); the 20-ms difference was an order of magnitude smaller than the gaze-duration effects that have been observed in experiments demonstrating covert change detection with these same scenes (Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002).

GENERAL DISCUSSION

A prevalent theoretical assumption in the current change-blindness literature is that attended scene elements can be retained across saccades in a veridical point-by-point format. Contrary to this view, the present study demonstrated that even when every pixel and every point-by-point visual feature in a scene changed, these changes were extremely difficult to detect across saccades. In two experiments, change detection rates, as indexed by overt detection responses, were very low, approaching zero in the first experiment for all viewers, and approaching zero in the second experiment for the majority of viewers. Furthermore, in Experiment 2, which equated the change and no-change conditions both in number of trials and in fixation positions following each change, there was no evidence from fixation durations indicating that viewers covertly detected the changes. Because in this paradigm a change took place across the entire scene image, changes necessarily occurred wherever attention (and the saccade) happened to be directed, so the failure to detect change could not have been due to a failure to attend to the changing region. Together, these results constitute the first demonstration of a global transsaccadic change-blindness effect, and indicate that point-by-point representations are not functional across saccades during complex real-world scene perception.

The global change-blindness effect reported here is particularly striking because viewers were informed about the nature of the changes beforehand and were presented with multiple changes on each trial. This global change-blindness effect is also observed for highly knowledgeable viewers. For example, despite our own familiarity with the images and complete understanding of when the changes take place, we regularly experience the phenomenon, and can successfully detect the global change only by adopting a verbal coding strategy explicitly describing some feature of the scene.

The results of the present study contrast with recent evidence demonstrating that relatively detailed representations of the visual properties of objects in scenes can be retained across saccades and stored in memory. For example, we (Hollingworth & Henderson, 2002) reported a study in which a specific token of an object class was replaced with another token from that same class during a saccade (see also Henderson & Hollingworth, 2003a; Hollingworth et al., 2001). Three results from this previous study strongly suggested that relatively detailed visual object representations were retained in memory. First, the overt detection rate for these token substitutions was 28%. Second, when the token substitutions were not overtly detected, there was evidence for covert detection: Gaze durations were elevated by an average of 145 ms compared with trials on which the same objects were unchanged. Third, a forced-choice memory test showed that participants could discriminate the object token that had appeared in a scene from another token of the same type with an average accuracy rate of 87%. These findings are even more striking

when one considers that the objects used in the token-substitution experiments occupied only about 3% of the area of each scene on average (e.g., Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002). In contrast, the global image changes used in the present study took place across the entire scene image, yet were almost impossible to detect.

The global change-blindness phenomenon reported here, together with recent demonstrations of good detection across saccades of token substitutions and other visual changes to attended objects, can be accommodated by a theory of dynamic scene perception in which information is integrated across saccades by visual representations that are abstracted away from the point-by-point image (Henderson & Hollingworth, 2003b; Hollingworth & Henderson, 2002). These representations are visual in the sense that they are orientation-specific (Henderson & Hollingworth, 1999a; Hollingworth & Henderson, 2002; see also Henderson & Siefert, 2001) and capable of coding the information needed to discriminate between different tokens of the same object class (Henderson & Hollingworth, 2003a; Hollingworth & Henderson, 2002). However, as demonstrated in the present study, these representations are not point-by-point sensory images of the kind that might be formed in primary visual cortex (see also Henderson, 1997; Irwin, 1991, 1992). Instead, the abstract visual representations that operate across saccades are consistent with functional accounts of more anterior brain areas, such as medial and inferior temporal cortex, where abstract visual categories seem to be represented and stored (e.g., Kreiman, Koch, & Fried, 2000; Logothetis & Sheinberg, 1996; Tanaka, 1996; Vuilleumier, Henson, Driver, & Dolan, 2002).

In summary, point-by-point visual representations are not functional across saccades, even for attended scene regions. More abstract visual representations, however, are functional across saccades as well as over longer periods of time, and contribute to an integrated scene representation that can support visual memory and dynamic visually guided activity.

Acknowledgments—This work was supported by grants from the National Science Foundation (BCS-0094433 and ECS-9873531) and the Army Research Office (DAAD19-00-1-0519; the opinions expressed in this article are those of the authors and do not necessarily represent the views of the Department of the Army or any other governmental organization; reference to or citations of trade or corporate names does not constitute explicit or implied endorsement of those entities or their products by the authors or the Department of the Army). We thank Dan Simons and two anonymous reviewers for their comments on the manuscript.

REFERENCES

- Crane, H.D., & Steele, C.S. (1985). Accurate three-dimensional eyetracker. *Applied Optics*, *17*, 691–705.
- Dennett, D.C. (1991). *Consciousness explained*. Boston: Little, Brown.
- Deubel, H., & Schneider, W.X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827–1837.
- Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Atkins (Ed.), *Perception* (Vancouver Studies in Cognitive Science Vol. 5, pp. 89–110). Oxford, England: 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. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, *8*, 51–55.
- 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.), *The perception of faces, objects, and scenes: Analytic and holistic processes* (pp. 356–383). New York: Oxford University Press.
- 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. (2001). Types and tokens in transsaccadic object integration. *Psychonomic Bulletin & Review*, *8*, 753–760.
- Hoffman, J.R., & Subramanian, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*, 787–795.
- Hollingworth, A. (in press). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*.
- Hollingworth, A., & Henderson, J.M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 113–136.
- Hollingworth, A., Williams, C.C., & Henderson, J.M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, *8*, 761–768.
- 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. (1994). On the measurement of phosphor persistence in oscilloscopic displays. *Vision Research*, *34*, 1623.
- 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.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*, 1897–1916.
- Kreiman, G., Koch, C., & Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature Neuroscience*, *3*, 946–953.
- Logothetis, N., & Sheinberg, D. (1996). Visual object recognition. *Annual Review of Neuroscience*, *19*, 577–621.
- McConkie, G.W., & Currie, C.B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 563–581.
- Rayner, K., McConkie, G.W., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and 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. (2002). Change detection. *Annual Review of Psychology*, *53*, 245–277.
- Rensink, R.A., O'Regan, J.K., & Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368–373.
- 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. (2000). Current approaches to change blindness. *Visual Cognition*, *7*, 1–15.
- Simons, D.J., & Levin, D.T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, *5*, 644–649.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139.
- Thiele, A., Henning, M., Kubischik, M., & Hoffman, P. (2002). Neural mechanisms of saccadic suppression. *Science*, *295*, 2460–2462.
- Volkman, F., Schick, A., & Riggs, L. (1968). Time course of visual inhibition during voluntary saccades. *Journal of the Optical Society of America*, *58*, 1310–1414.
- Vuilleumier, P., Henson, R.N., Driver, J., & Dolan, R.J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, *5*, 491–499.
- Wolfe, J.M. (1999). Inattention blindness. In V. Coltheart (Ed.), *Fleeting memories* (pp. 71–94). Cambridge, MA: MIT Press.

(RECEIVED 7/28/02; REVISION ACCEPTED 11/27/02)