

# Eye movements during scene viewing: Evidence for mixed control of fixation durations

JOHN M. HENDERSON

*University of Edinburgh, Edinburgh, Scotland*

AND

GRAHAM L. PIERCE

*Michigan State University, East Lansing, Michigan*

Recent behavioral and computational research on eye movement control during scene viewing has focused on where the eyes move. However, fixations also differ in their durations, and when the eyes move may be another important indicator of perceptual and cognitive activity. Here we used a scene onset delay paradigm to investigate the degree to which individual fixation durations are under direct moment-to-moment control of the viewer's current visual scene. During saccades just prior to critical fixations, the scene was removed from view so that when the eyes landed, no scene was present. Following a manipulated delay period, the scene was restored to view. We found that one population of fixations was under the direct control of the current scene, increasing in duration as delay increased. A second population of fixations was relatively constant across delay. The pattern of data did not change whether delay duration was random or blocked, suggesting that the effects were not under the strategic control of the viewer. The results support a mixed control model in which the durations of some fixations proceed regardless of scene presence, whereas others are under the direct moment-to-moment control of ongoing scene analysis.

When viewing pictures of real-world scenes, we move our eyes about three times per second via saccadic eye movements (Henderson, 2003; Rayner, 1998). Most recent research on eye movement control in scenes has focused on the placement of fixations (e.g., Baddeley & Tatler, 2006; Itti & Koch, 2000; Parkhurst & Niebur, 2003; Tatler, Baddeley, & Vincent, 2006; Torralba, Oliva, Castelhano, & Henderson, 2006; Underwood & Foulsham, 2006). Far less research has focused on the control of fixation duration (Henderson, 2003). Fixation durations during scene viewing last about 300 msec on average, but vary substantially around this mean (for reviews, see Henderson & Hollingworth, 1998, 1999; Rayner, 1998). The degree to which variability in fixation duration during scene viewing is due to moment-to-moment changes in perceptual processing is currently unknown. In this study, we used a scene onset delay paradigm based on a similar technique developed in reading (Rayner & Pollatsek, 1981; see also Morrison, 1984; Shiori, 1993; Vaughan, 1982; Vaughan & Graefe, 1977) to investigate the degree to which the duration of a fixation is directly and immediately controlled by the visual stimulus present during that fixation.

As a first pass, we should distinguish two general classes of eye movement control theories (Rayner, 1998). *Direct control* theories posit that decisions about fixation

termination (saccade initiation) are made during a viewer's current fixation, based on properties of the stimulus that are available in that fixation. In contrast, *indirect control* theories posit that factors other than those in the current fixation determine fixation duration. Direct control requires that an evaluation of the success of ongoing perceptual and/or cognitive processing can feed back in real time to the saccade initiation system. Indirect control would not require real-time feedback. Variations of indirect control include *delayed control*, *general parameter control*, and *timing control*. Delayed control assumes that stimulus processing during a fixation is too slow to produce an immediate effect of the current stimulus on the current fixation. General parameter control suggests that fixation durations are based on global properties of the stimulus or viewing task; for example, the duration parameter might be set to a long value at the beginning of a difficult visual search (Hooge, Vlaskamp, & Over, 2007). Timing control assumes that individual fixation durations are determined by an internal timing mechanism that is designed to keep the eyes moving at a constant rate.

The strongest evidence for direct control of fixation durations is provided by the stimulus onset delay paradigm (Morrison, 1984; Rayner & Pollatsek, 1981; Shiori, 1993; Vaughan, 1982). In this paradigm, the stimulus is

---

J. M. Henderson, john.m.henderson@ed.ac.uk

---

removed from the display during a saccade, delaying its appearance at the beginning of the next fixation for some specified amount of time. After the delay, the stimulus is restored. Direct control predicts that fixation durations should increase with the duration of the delay. In contrast, indirect control predicts that the duration of the onset delay will not influence the duration of the fixation. Depending on the exact nature of indirect control, the onset delay might influence the duration of a fixation further downstream, or might not have any influence at all (Morrison, 1984; Rayner & Pollatsek, 1981; see also Vaughan, 1982).

### The Present Study

In the present study, we applied the stimulus onset delay paradigm to scenes. In our *scene onset delay* paradigm, participants viewed photographs of real-world scenes while preparing for a later memory test. On a critical 10% of the fixations in each trial, the scene was removed during the saccade prior to the beginning of the critical fixation, and was then presented again following a predetermined delay. If fixation durations during scene viewing are under the direct and immediate control of the current scene input, then fixation durations should increase monotonically with scene onset delay; in the extreme, the slope should approach 1.0 if fixation durations perfectly reflect the availability of useful visual information. In contrast, if fixation durations reflect delayed effects of properties of the visual input determined primarily by a general timing mechanism, or are largely preprogrammed, then we should observe no specific relationship between scene onset delay and fixation duration.

We used the scene onset delay paradigm to investigate the control of fixation durations during scene viewing. In Experiment 1, delays were 0 (control), 40, 75, 100, 250, and 450 msec. In Experiment 2, they were 150, 200, 300, 400, 600, and 750 msec. In Experiment 3, they were 0 (control), 100, 200, 600, and 1,200 msec. The longer delays allowed us to investigate the nature of anticipatory saccades, as well as the degree to which fixation durations can be “stretched.” To reduce the chances of strategy effects, we delayed scene onset during only 10% of all fixations in each of the experiments. In Experiments 1 and 2, delay was chosen from fixation to fixation in a pseudorandom manner. In Experiment 3, we used a random versus blocked delay manipulation to investigate directly whether any observed delay effects were open to viewing strategies (Morrison, 1984).

In summary, we conducted three experiments over a total of 36 independent participants. To facilitate comparison, we combine our report of the data from the three experiments within the Method, Results, and Discussion sections.

## METHOD

### Participants

Thirty-six undergraduate students at Michigan State University—12 per experiment, all naive to the purposes of the study—participated as volunteers or in exchange for monetary payment.

### Apparatus

Eye movements were monitored by an SR Research EyeLink II eyetracker sampling at 500 Hz (Experiments 1 and 2) or by an SR Research EyeLink 1000 eyetracker sampling at 1000 Hz (Experiment 3). Viewing was binocular, but only the right eye was tracked. The images were presented on a 21-in. Viewsonic CRT monitor with a refresh rate of 160 Hz (Experiments 1 and 2) or 150 Hz (Experiment 3). Scenes subtended  $26.79^\circ \times 20.58^\circ$  at a viewing distance of 84 cm. The experiment was controlled with SR Research Experiment Builder software.

### Stimuli

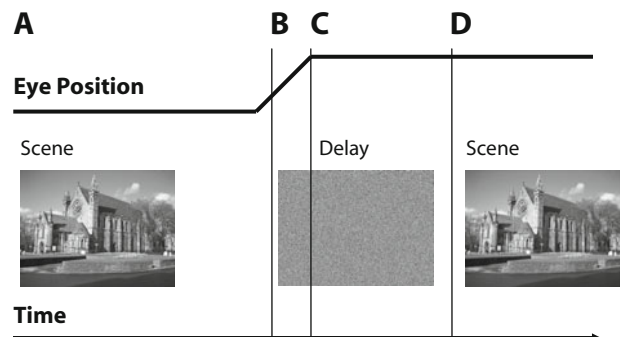
Participants were presented 40 unique, full-color  $800 \times 600$  pixel photographs of real-world scenes (20 outdoor, 20 indoor) from a variety of scene categories.

### Procedure

Each participant viewed all 40 scenes. Each scene was presented for 101 saccades (40–100 sec). Participants viewed the scenes in preparation for a memory test to be administered later and were told to ignore any occasional flicker they might notice.

Scene onset delay was implemented by a saccade-contingent display change: The scene was removed from the CRT during the saccade just prior to the critical fixation (see Figure 1). With this method, the scene was not present when the critical fixation began. Scene onset delays took place every 10th fixation.<sup>1</sup> To avoid phosphor persistence, a color noise mask was presented during the delay.<sup>2</sup> Following the delay, the scene reappeared. In Experiments 1 and 2, saccades were identified as eye movements whose velocity exceeded  $30^\circ/\text{sec}$  and acceleration exceeded  $8,000^\circ/\text{sec}^2$ . In Experiment 3, saccades were identified as eye movements whose velocity exceeded  $50^\circ/\text{sec}$ . Fixations that began prior to the display change completion and those on which the participant blinked were removed from analysis; this resulted in the removal of 19% of the data from Experiments 1 and 2 and 11% from Experiment 3.

Participants in Experiment 1 were presented delays of 0 (control), 40, 75, 100, 250, and 450 msec. Participants in Experiment 2 were given delays of 150, 200, 300, 400, 600, and 750 msec. Delay values in Experiments 1 and 2 were chosen pseudorandomly for each critical fixation within each scene. Participants in Experiment 3 were given delays of 0 (control), 100, 200, 600, and 1,200 msec; the delays were varied within each scene pseudorandomly for one block of trials (random block) and remained constant within each scene for a



**Figure 1.** Illustration of the scene onset delay paradigm. At the first time-point (A), the eyes are in the ninth fixation following the last delay. When the next saccade is detected (B), the display is changed, so that when the eyes begin the following critical fixation (C), the scene has been removed from view. Following the specified delay (D), the scene reappears. The duration of the remaining saccade following the display change (C – B) is subtracted from the specified delay (D – B) to generate the actual delay that is plotted in subsequent figures. Scenes and mask were presented in color.

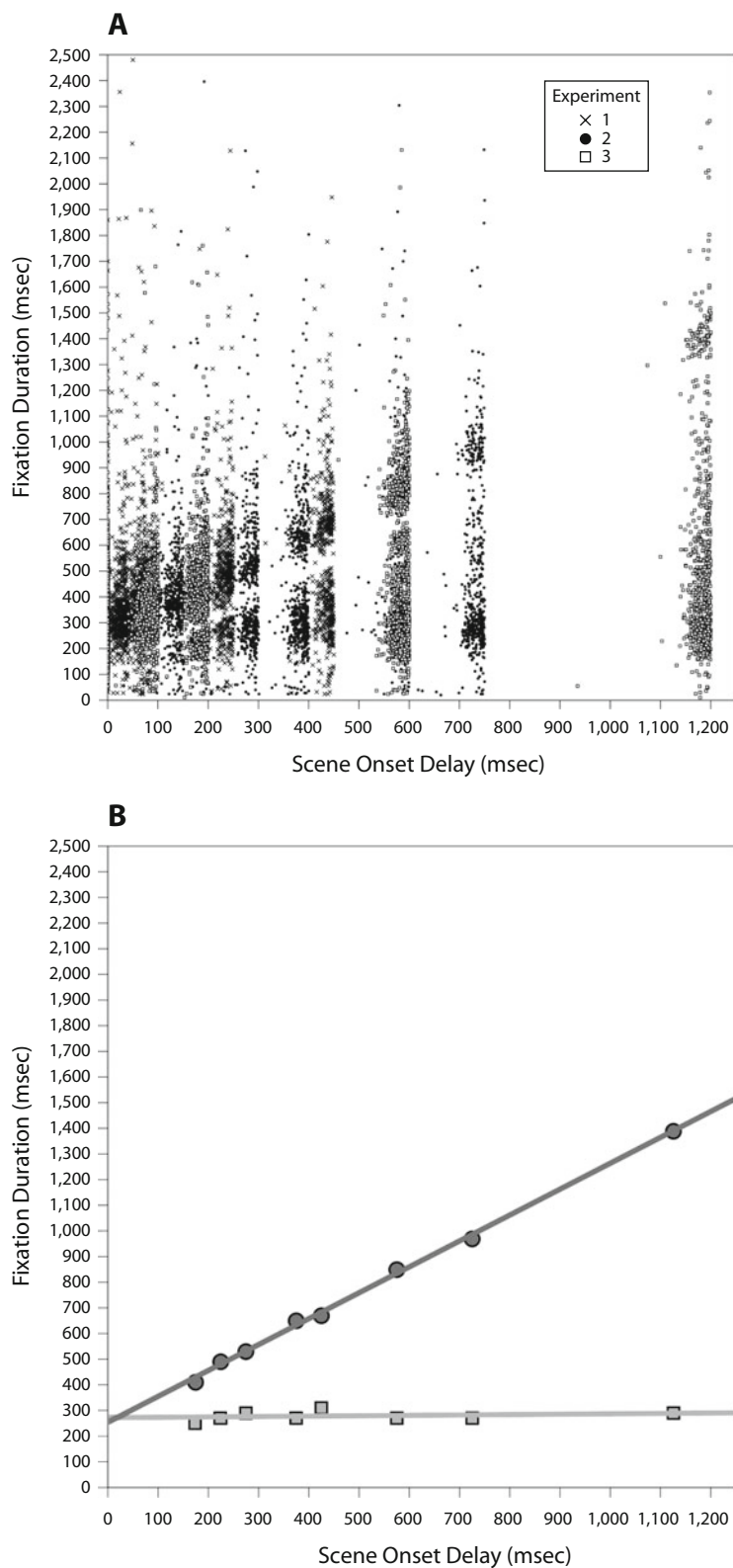


Figure 2. (A) Fixation duration as a function of scene onset delay in Experiments 1, 2, and 3. Each point represents one critical fixation, defined as the first fixation following the saccade in which the scene was removed from the display (point C in Figure 1). (B) Regression lines fit to the two modes at each scene onset delay across the three experiments.

second block of trials (constant block). Block order was counterbalanced across participants.

## RESULTS

### Onset Delay Results

Figure 2 shows the duration of all fixations across the three experiments as a function of scene onset delay. Each prescribed onset delay generated a small range of actual onset delays because the delay timer began during the saccade prior to the critical fixation (see Figure 1). The time of the remaining saccade prior to the critical fixation was therefore subtracted from the nominal delay in order to generate the actual delay. Two fixation duration populations are apparent in Figure 2 at each delay: One population increased monotonically with delay (the upper cluster of fixation durations apparent at each delay), whereas the second population remained constant across delay (the lower cluster of fixation durations). Also apparent in Figure 2 is a gap in the fixation duration distribution at each delay. The timing of this gap is consistent with the saccadic inhibition effect previously reported for display changes and is likely due to the motion transient created by the reappearance of the scene after the delay (Reingold & Stampe, 2002).

The upper tail of each fixation duration distribution that was observed for each delay became longer as the delay increased (Figure 2). These fixations showed direct control. Also, a substantial population of fixations remained relatively constant over delay. Since they had been launched before the scene reappeared, the latter fixations were consistent with the anticipatory saccades observed in the stimulus onset delay paradigm in reading (Morrison, 1984; Rayner & Pollatsek, 1981). The proportion of anticipatory saccades increased as delay increased, as shown in Table 1; this is also consistent with prior results from studies of reading, suggesting that the proportion of fixations held for the entire delay decreases as the delay increases (Morrison, 1984; Rayner & Pollatsek, 1981).

Following Rayner and Pollatsek (1981), regression lines were fit to the two populations for all delays over 150 msec (Figure 2). Since computing means for the two populations would have required assigning each fixation to one or the other population a priori—an assignment that would have required assumptions about the exact nature of

the boundaries of what were likely to be two overlapping populations—we chose instead to compute the regression equation over the modes of the distributions. The regression analysis for the upper population showed that these fixation durations increased linearly and in a one-to-one relationship with delay (slope = 1.01, intercept = 252.09,  $R^2 = 1.00$ ,  $p < .01$ ). These data provide strong evidence for the existence of a population of fixations whose durations are under the direct control of the current scene input. In contrast, the durations of the lower population of fixations remained constant across delay (slope = 0.02, intercept = 268.95,  $R^2 = .09$ , n.s.), suggesting that this second population of fixations with shorter durations was not influenced by delay.

Could it be that fixation durations were elevated in the delay conditions solely because of saccadic inhibition? Although the upper modes in each distribution might potentially have been due to a combination of fixations that had elevated durations and fixations that were displaced by saccadic inhibition, this cannot entirely account for the stimulus onset delay effect. First, the longer fixation durations in the longer delay conditions, which became apparent earlier than the saccadic inhibition “gap,” cannot be accounted for by saccadic inhibition. In fact, the gap itself in the longer delay conditions existed only because fixation durations had been lengthened to the point of the gap. If the second (elevated) population of fixations were simply a displacement of fixations from the tail of a constant underlying distribution due to saccadic inhibition, the underlying distribution would have to have had a tail where the gap appeared. The data from the zero- and short-delay conditions show that it did not. In summary, the gap in the longer delay conditions appeared only because fixations in those delay conditions had been lengthened in duration.

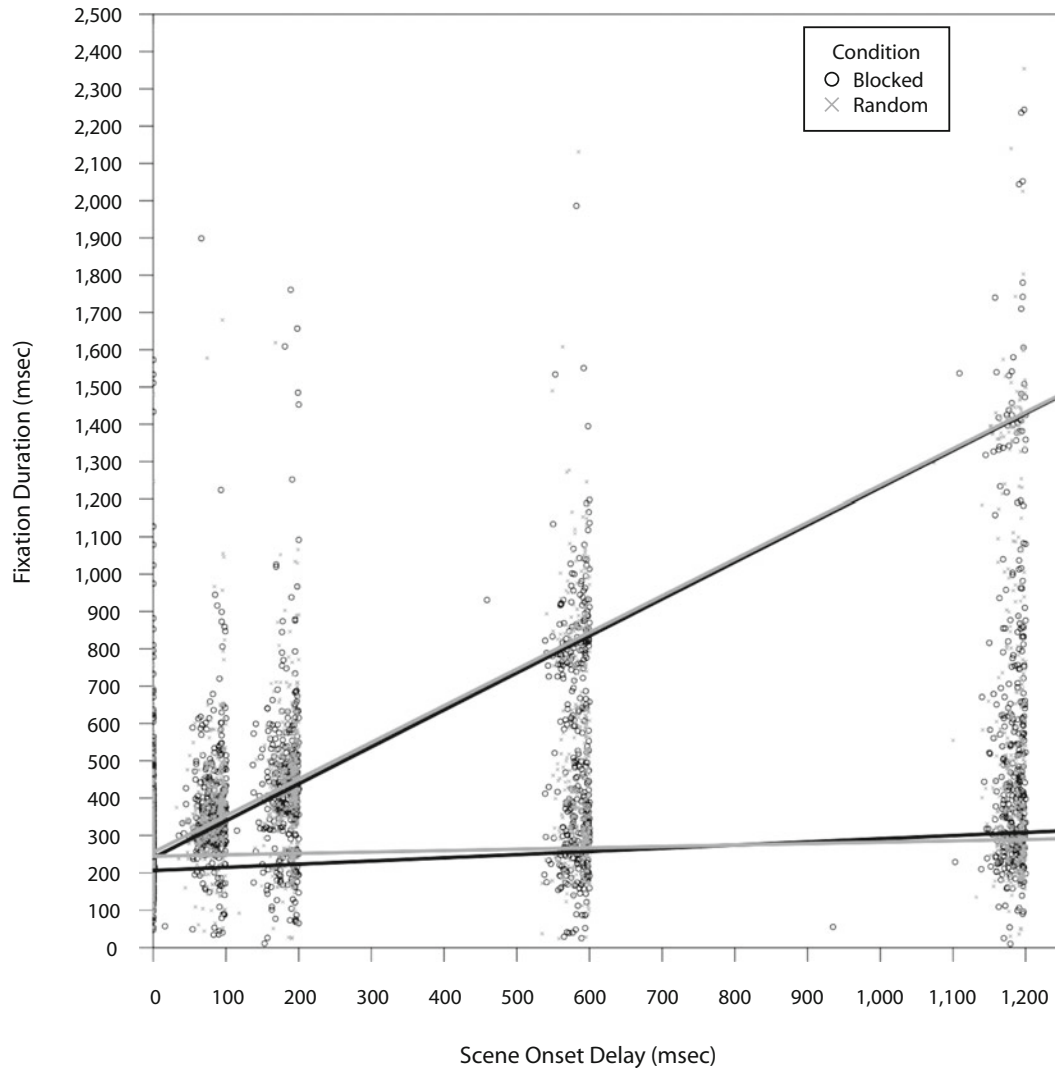
### Delay Spillover

We also analyzed the fixation  $n+1$  durations—those immediately following the critical fixations ( $n$ )—as a function of the scene onset delay during the critical fixation. Unlike with the critical fixations, the resulting distributions of durations for  $n+1$  fixations were monomodal. Fixation  $n+1$  was longer in duration (292 vs. 247 msec) when the critical fixation was directly controlled (i.e., outlasted the delay) than when it was anticipatory, but these durations did not differ as a function of the scene onset delay in the critical fixation (slopes of  $-.01$ , n.s., for regression of fixation  $n+1$  duration on fixation  $n$  scene onset delay). By two fixations downstream, no effects of delay were apparent. These data provide no evidence of a “spillover” effect on fixation  $n+1$  of the scene onset delay in the critical fixation. Taken together, the data show that when scene onset delay influenced fixation duration, the influence was restricted to the fixation containing the delay. This finding suggests that the extra processing time needed during a fixation is completely accounted for by the duration of that fixation and does not support a version of delayed control in which the influence of processing in a fixation is reflected in the subsequent fixation.

**Table 1**  
Percentage of Fixations in the Lower and Upper Populations  
at Each Scene Onset Delay (in Milliseconds)

Scene Onset Delay	Lower Population Percentage	Upper Population Percentage
175	20.89	79.11
225	24.21	75.79
275	47.48	52.52
375	58.78	41.22
425	48.82	51.18
575	69.98	30.02
725	75.96	24.04
1,175	87.24	12.76

Note—Data were separated by the saccadic inhibition gap in the distributions at each onset delay.



**Figure 3.** Fixation duration as a function of scene onset delay for blocked (Experiment 2) and random (Experiment 3) delays. The two distributions are indistinguishable, indicating that the effect of onset delay on fixation duration was not a result of participant strategy.

### Blocked Versus Random Delays

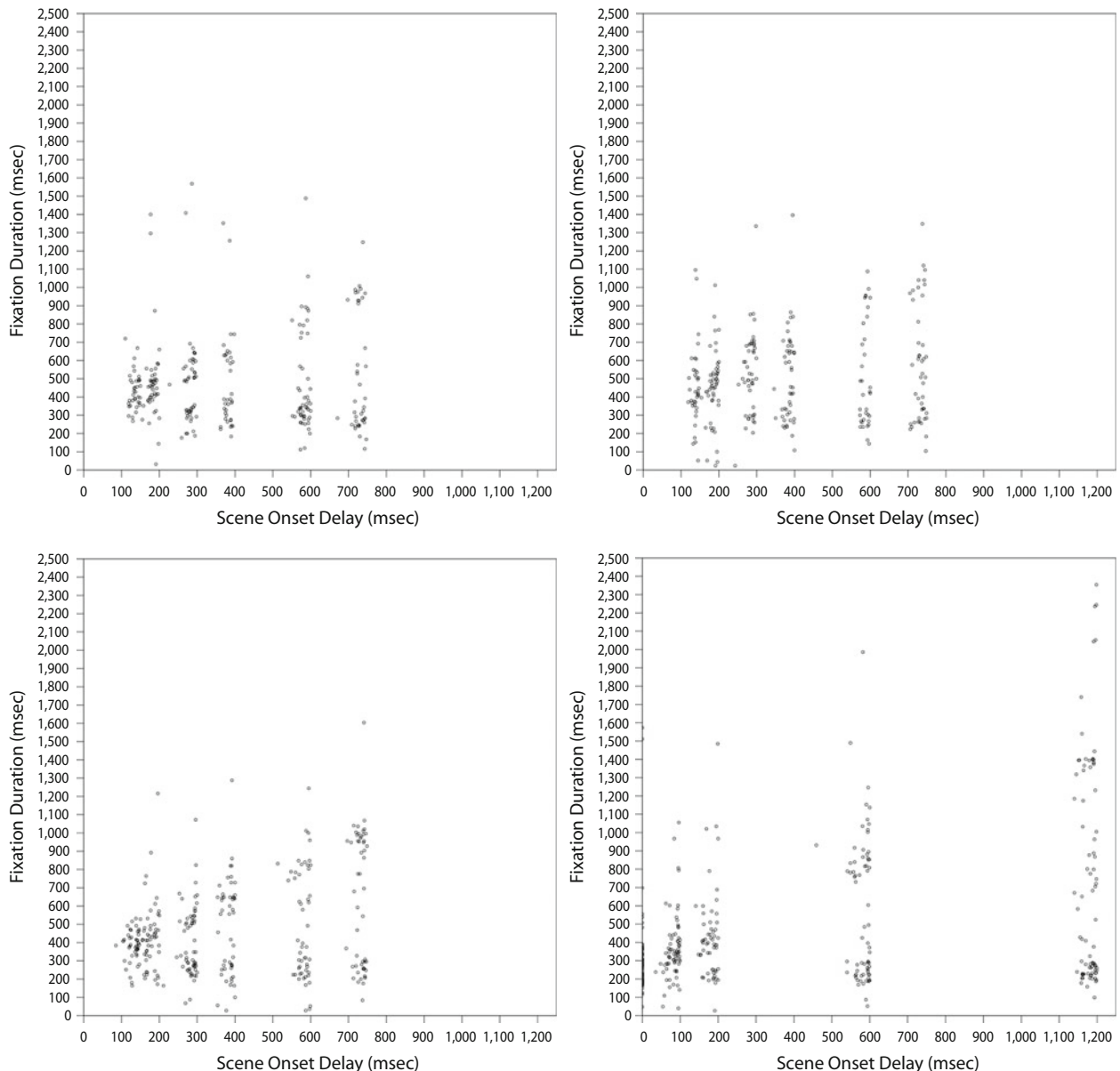
The distributions of fixation durations in the random and constant-delay conditions from Experiment 3, along with the best-fitting regression equations, are shown in Figure 3. To determine whether scene onset delay affected fixation duration differently when delay was blocked versus random, we conducted two analyses. First, a 5 (stimulus onset delay)  $\times$  2 (blocked vs. random)  $\times$  2 (blocked vs. random first) ANOVA showed that, whereas the effect of scene onset delay was significant [ $F(4,100) = 727.28, p < .01$ ], neither the blocking condition nor condition order produced significant effects; nor were any interactions involving these factors significant ( $F_s < 1$ ). Second, we fit separate regression equations to the data from the blocked and random conditions, also shown in Figure 3. For the random data, the regression analysis for the upper population showed that these fixation durations increased linearly and in a one-to-one relationship with

delay (slope = 0.982, intercept = 253.49,  $R^2 = .995, p < .05$ ), whereas for the lower population, no relationship was observed (slope = 0.037, intercept = 243.03,  $R^2 = .645, p > .40$ ). The regressions for the blocked condition were very similar (slope = 0.989, intercept = 238.42,  $R^2 = .993, p = .05$ , and slope = 0.084, intercept = 202.63,  $R^2 = .738, p > .30$ ). These results indicate that the pattern of data was not the result of participant strategies.

### Participant Consistency

The two populations of fixation durations at each onset delay could reflect similar participant-by-participant distributions, or could be due to a mixture of data produced by two subsets of participants, with each participant generating only one of the two populations. Participant-by-participant plots of fixation duration as a function of scene onset delay were consistent with the former interpretation (see Figure 4). Each participant showed the same division





**Figure 4.** Fixation duration as a function of scene onset delay for 8 representative participants (Experiments 1 and 2). The two populations of fixations at each onset delay were present in each participant's data set, demonstrating that these fixation populations do not reflect two populations of participants. (Figure 4 continues on p. 572.)

of fixations into two populations as a function of scene onset delay, with one population of fixations unaffected by scene onset delay and a second population increasing as delay increased. These data are similar to those reported by Morrison (1984) in a study of reading and suggest that the direct control of fixation duration was not limited to a subset of the participants.

## DISCUSSION

Fixations differ in their durations during real-world scene viewing (Henderson, 2003; Henderson & Hollingworth, 1998, 1999; Rayner, 1998). The present study provides the first unambiguous evidence that a substantial proportion

of this variability is due to the immediate control of fixation durations by currently available scene information. For these directly controlled fixations, an increase in the time needed for scene analysis within the fixation was entirely accounted for by an increase in the duration of that fixation. Furthermore, the delay of scene analysis in one fixation did not spill over into the next fixation. These findings, together, are consistent with a direct control hypothesis, according to which scene analysis taking place within a fixation is completed prior to the termination of that fixation.

In addition to the directly controlled fixations, the durations of a second population of fixations were independent of current visual input. These fixations might have been indirectly controlled by a general fixation duration param-

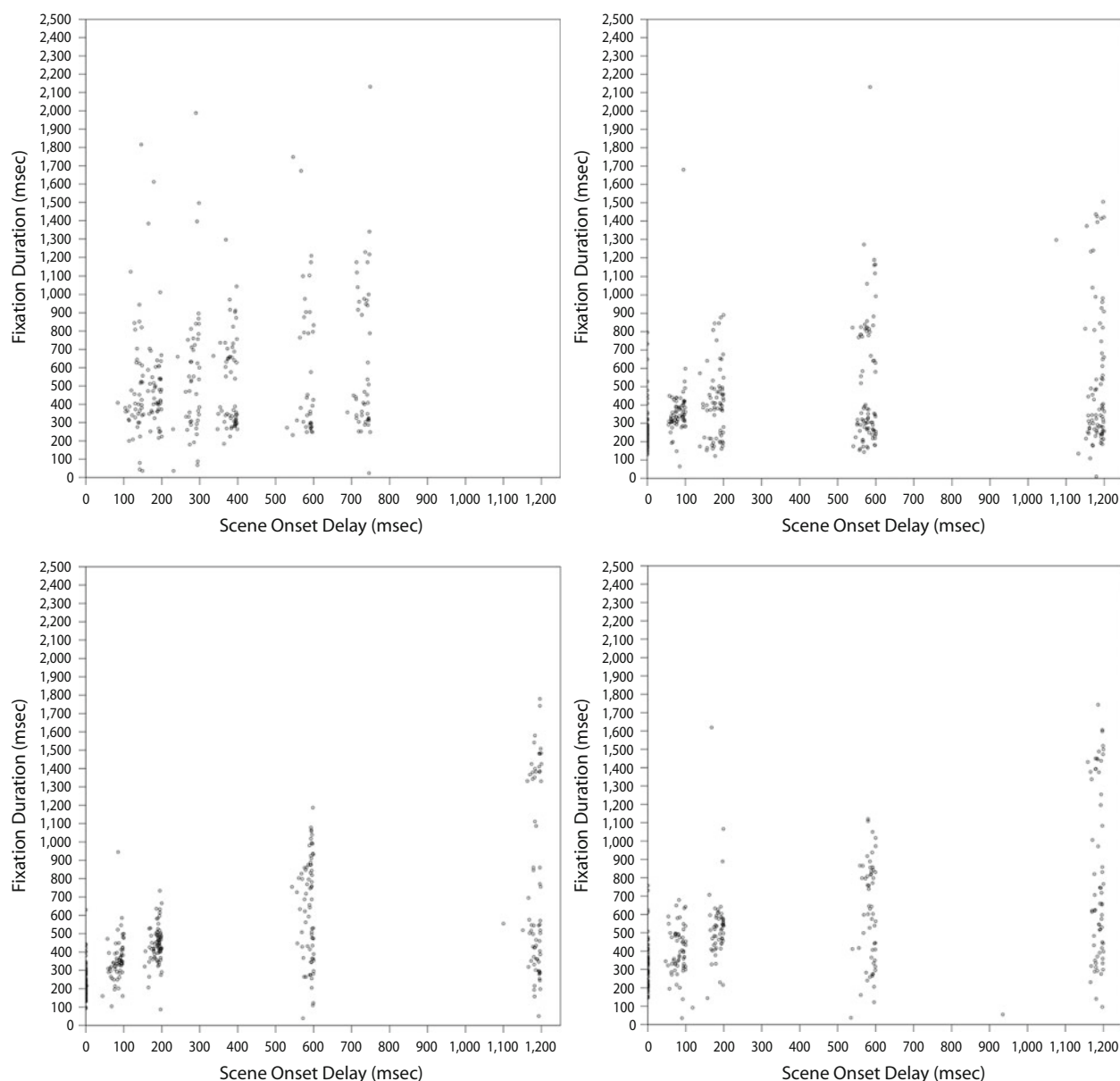


Figure 4 (continued).

eter that is set according to the nature of the stimulus and task (Hooge & Erkelens, 1998). Alternatively, they might have been determined by a timing mechanism that is uninfluenced by the current stimulus and task, and that sometimes wins out over the direct control system (Engbert, Nuthmann, Richter, & Kliegl, 2005; Findlay & Walker, 1999; Yang & McConkie, 2001). Additional research will be needed to distinguish between these two hypotheses.

Current theories of gaze control in scene viewing (and the computational models that implement them) focus solely on accounting for *where* the eyes move, rather than *when* they move (e.g., Itti & Koch, 2000; Torralba et al., 2006). If individual fixation duration can be treated as a constant, there is no need to extend these models in order to explicitly account for duration. Furthermore, effects

of visual or cognitive processes on more global measures of fixation time, such as gaze duration (the sum of all fixations in a region from entry to exit of that region), could be accounted for within these current models, because if individual fixation durations are a constant, then increased gaze durations must entirely be the result of an increased probability of refixation. This data pattern is, in fact, often observed in scene viewing (e.g., Henderson & Hollingworth, 2003; Henderson, Weeks, & Hollingworth, 1999; Hollingworth, Williams, & Henderson, 2001). The same logic holds true for total fixation duration (the sum of all fixations in a region, including region reentries). Since current gaze control models already account for the locations of fixations, they can also, in principle, account for the locations of *refixations*, giving them an ability to

explain differences in gaze durations or total fixation time by differences in refixations within a given scene region. Given the present finding—that individual fixation durations are also, to some extent, under direct control in scene viewing—current computational models of scene viewing need to be extended in order to explicitly account for when the eyes move (Henderson, 2003).

In summary, the present study is the first to produce unambiguous evidence that individual fixations can be directly and immediately controlled by the current visual stimulus during scene viewing. These results provide an existence proof that fixation durations can be directly controlled, and so open the possibility that other sorts of processes (e.g., feature analysis, object recognition, spatial analysis, memory encoding) could, in principle, exert a similar influence. Also, it is clear that some fixation durations are not under direct control, and an open question is whether these fixations reflect delayed control, timing control, preprogramming, or some other mechanism. These results have important implications for areas of research that use eye movements to study how underlying processes unfold in real time, as well as for understanding how attention and fixation are directed through a complex visual scene.

#### AUTHOR NOTE

We thank Fernanda Ferreira, Robin Hill, George Malcolm, Antje Nuthmann, Tim Smith, and the Edinburgh Eye Movement Users Group for their feedback on this research. We also thank Geoffrey Underwood and an anonymous reviewer for their thoughtful comments on the manuscript. Correspondence concerning this article should be addressed to J. M. Henderson, Psychology Department, 7 George Square, University of Edinburgh, Edinburgh EH8 9JZ, Scotland (e-mail: john.m.henderson@ed.ac.uk).

#### REFERENCES

- BADDELEY, R. J., & TATLER, B. W. (2006). High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis. *Vision Research*, **46**, 2824-2833.
- ENGBERT, R., NUTHMANN, A., RICHTER, E. M., & KLIEGL, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, **112**, 777-813.
- FINDLAY, J. M., & WALKER, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral & Brain Sciences*, **22**, 661-721.
- HENDERSON, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, **7**, 498-504.
- 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. 269-293). Oxford: Elsevier.
- HENDERSON, J. M., & HOLLINGWORTH, A. (1999). High-level scene perception. *Annual Review of Psychology*, **50**, 243-271.
- HENDERSON, J. M., & HOLLINGWORTH, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics*, **65**, 58-71.
- HENDERSON, J. M., WEEKS, P. A. JR., & HOLLINGWORTH, A. (1999). Effects of semantic consistency on eye movements during scene viewing. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 210-228.
- 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.
- HOOGE, I. T. C., & ERKELENS, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, **38**, 1295-1302.
- HOOGE, I. T. C., VLASKAMP, B. N. S., & OVER, E. A. B. (2007). Saccadic search: On the duration of a fixation. In R. G. P. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movement research: Insights into mind and brain* (pp. 581-595). Amsterdam: Elsevier.
- ITTI, L., & KOCH, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, **40**, 1489-1506.
- MORRISON, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 667-682.
- PARKHURST, D. J., & NIEBUR, E. (2003). Scene content selected by active vision. *Spatial Vision*, **16**, 125-154.
- RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.
- RAYNER, K., & POLLATSEK, A. (1981). Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology*, **33A**, 351-373.
- REINGOLD, E. M., & STAMPE, D. M. (2002). Saccadic inhibition in voluntary and reflexive saccades. *Journal of Cognitive Neuroscience*, **14**, 371-388.
- SHIORI, S. (1993). Postsaccadic processing of the retinal image during picture scanning. *Perception & Psychophysics*, **53**, 305-314.
- TATLER, B. W., BADDELEY, R. J., & VINCENT, B. T. (2006). The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task. *Vision Research*, **46**, 1857-1862.
- TORRALBA, A., OLIVA, A., CASTELHANO, M. S., & HENDERSON, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, **113**, 766-786.
- UNDERWOOD, G., & FOULSHAM, T. (2006). Visual saliency and semantic incongruity influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, **59**, 1931-1949.
- VAUGHAN, J. (1982). Control of fixation duration in visual search and memory search: Another look. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 709-723.
- VAUGHAN, J., & GRAEFE, T. (1977). Delay of stimulus presentation after the saccade in visual search. *Perception & Psychophysics*, **22**, 201-205.
- YANG, S.-N., & MCCONKIE, G. W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research*, **41**, 3567-3585.

#### NOTES

1. Participants were unaware that there was a relationship between their eye movements and scene blanking, and so were unaware that blanking took place every 10th fixation.
2. We have replicated the same data pattern with a uniform gray filler field.

(Manuscript received October 29, 2007;  
revision accepted for publication January 2, 2008.)