

# Neural Correlates of Fixation Duration during Real-world Scene Viewing: Evidence from Fixation-related (FIRE) fMRI

John M. Henderson and Wonil Choi

## Abstract

■ During active scene perception, our eyes move from one location to another via saccadic eye movements, with the eyes fixating objects and scene elements for varying amounts of time. Much of the variability in fixation duration is accounted for by attentional, perceptual, and cognitive processes associated with scene analysis and comprehension. For this reason, current theories of active scene viewing attempt to account for the influence of attention and cognition on fixation duration. Yet almost nothing is known about the neurocognitive systems associated with variation in fixation duration during scene viewing. We addressed this topic using fixation-related fMRI, which involves coregistering high-resolution eye tracking and

magnetic resonance scanning to conduct event-related fMRI analysis based on characteristics of eye movements. We observed that activation in visual and prefrontal executive control areas was positively correlated with fixation duration, whereas activation in ventral areas associated with scene encoding and medial superior frontal and paracentral regions associated with changing action plans was negatively correlated with fixation duration. The results suggest that fixation duration in scene viewing is controlled by cognitive processes associated with real-time scene analysis interacting with motor planning, consistent with current computational models of active vision for scene perception. ■

## INTRODUCTION

When we actively view a scene, our eyes move from location to location, fixating objects and other scene elements for varying amounts of time (Henderson, 2011, 2013; Henderson & Hollingworth, 1999). Fixation duration averages 300 msec during active scene viewing, with considerable variability around this mean (Henderson, 2011; Rayner, 2009; Land & Hayhoe, 2001; Buswell, 1935). An important research question in scene perception is therefore the nature of the processes that determine fixation duration (Nuthmann, Smith, Engbert, & Henderson, 2010). Fixation duration is affected by visual quality manipulations, with longer durations for lower-quality images (Loftus, 1985). Similarly, viewing task influences fixation duration. For example, longer average durations are observed during scene memorization than search (Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011; Nuthmann et al., 2010; Castelano, Mack, & Henderson, 2009; Henderson, Weeks, & Hollingworth, 1999). These average effects could be due to differences in global fixation duration parameters set for a specific level of scene clarity or task or to moment-to-moment oculomotor decisions concerning current visual and cognitive processing. The hypothesis that fixation duration reacts to (and reflects) transitory aspects of scene

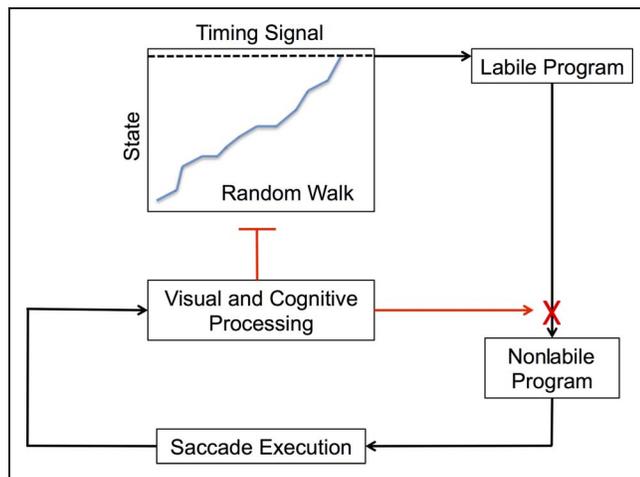
analysis in real time has been referred to as direct control (Henderson & Pierce, 2008).

Consistent with direct control, a series of studies has demonstrated that fixation duration is sensitive to the availability of useful scene information in that fixation. These studies initially used a scene onset delay paradigm (Morrison, 1984), in which a saccade-contingent display change technique allows the viewed scene to be occasionally removed from the display while the participant's eyes are in saccadic movement from one location to another (Luke, Nuthmann, & Henderson, 2013; Henderson & Smith, 2009; Henderson & Pierce, 2008). In this way, the scene is not visible at the beginning of the next critical fixation. After a predetermined delay, the scene returns to the display. The duration of the delay is varied, and the influence of the delay on the duration of the critical fixation is measured. In studies using this paradigm, a large population of fixations shows a strong positive linear relationship between delay duration and fixation duration, suggesting that these fixations are controlled directly and in real time by the current scene image.

Results from several other paradigms provide converging evidence that fixation duration is under direct control during scene viewing. For example, in the scene degradation paradigm, the image quality of a scene is reduced during a saccade just before a critical fixation so that quality is lower when the eyes land (Henderson, Nuthmann, & Luke, 2013). Then, during the saccade that terminates the critical fixation, the scene returns to its base quality.

When quality is lowered by reducing luminance, critical fixation duration increases monotonically as luminance decreases (Walshe & Nuthmann, 2014; Henderson et al., 2013). These effects generalize to other image manipulations; for example, fixation duration also increases monotonically as the scene becomes more blurred by spatial frequency filtering (Henderson, Olejarczyk, Luke, & Schmidt, 2014; Glaholt, Rayner, & Reingold, 2013). Furthermore, the opposite effect is also seen: Critical fixations can be reduced in duration when scenes become less blurred during the critical fixation (Henderson, Olejarczyk, et al., 2014). Finally, results from survival analyses of fixation durations in scene viewing indicate a direct influence of scene on the duration of the first scene fixation (Glaholt & Reingold, 2012). Together these results indicate that the durations of fixations during scene viewing are under real-time control of current scene processing.

To account for the control of fixation duration during scene viewing, we developed a theory of active scene perception and implemented it as CRISP (Controlled Random-walk with Inhibition for Saccade Planning), a working computational model (Nuthmann et al., 2010). CRISP is predicated on two main assumptions (Figure 1). First, saccade programming is initiated by a random walk timing process. Second, consistent with direct control, ongoing perceptual and cognitive processes associated with current scene analysis can influence the random walk via moment-by-moment modulation of the transition rate. An increase in the random walk's transition rate



**Figure 1.** Schematic of CRISP model (Nuthmann et al., 2010). Fixation duration is generated by a random walk timing signal. The signal accumulates toward a threshold. Accumulation rate is affected by ongoing perceptual and cognitive processing. Once the threshold is reached a new saccade program is initiated. The saccade program enters a labile stage, which signals the engagement of the oculomotor system. Saccade programs can be cancelled in the labile stage by inhibition. At the end of the labile stage, a point of no return is reached. During the following nonlabile stage, the saccade can no longer be cancelled. Finally, the saccade is executed. Fixation durations are the time intervals between successive saccades.

leads to increased time before the next saccade program is initiated and therefore longer fixation duration. In addition, when a saccade program has already been initiated, processing difficulty can cancel that program if it is still in a labile state (Becker & Jurgens, 1979), a common assumption in gaze control models (Reichle, Rayner, & Pollatsek, 2003; Reichle, Pollatsek, Fisher, & Rayner, 1998). The CRISP model is able to account for fixation duration distributions observed during scene viewing, differences in fixation duration distributions across tasks, and the observed influence of the stimulus onset delay paradigm on fixation durations (Nuthmann et al., 2010).

In summary, there is good behavioral evidence for direct control of fixation duration during scene viewing, and direct control has been successfully modeled (Nuthmann et al., 2010). However, almost nothing is currently known about the neurocognitive systems associated with these processes. To investigate this topic, we simultaneously recorded and coregistered eye movements and fMRI while participants viewed digitized photographs of real-world scenes. We then conducted fixation-related (FIRE) analysis of the fMRI data (Richlan et al., 2014; Marsman, Renken, Velichkovsky, Hooymans, & Cornelissen, 2011). Our goal was to identify regions participating in the network supporting active scene viewing by localizing neural activation correlated with the durations of fixations.

During the study, participants viewed 22 scenes for 12 sec each in preparation for a later memory test. This viewing task has several advantages over unconstrained free viewing. First, in free viewing, participants tend to generate their own implicit task, with unknown variability in this task across participants. Second, the memory task has been used extensively in the eye movement literature, and so its relationship to eye movements is well known (Mills et al., 2011; Nuthmann et al., 2010; Castelhamo et al., 2009; Henderson et al., 1999). When participants actively view scene images in this task, they distribute fixations over the majority of the scene and they generate considerable variability in fixation durations. Furthermore, fixation time correlates with memory performance (Williams, Henderson, & Zacks, 2005; Hollingworth & Henderson, 2002). Third, fixation duration in this task has been successfully modeled, providing plausibility for the assumption that cognition and eye movement control interact during scene viewing (Nuthmann et al., 2010). Finally, scene analysis and encoding is associated with activity in medial-temporal brain regions including hippocampus and posterior parahippocampal gyrus (Staresina, Duncan, & Davachi, 2011; Henderson, Larson, & Zhu, 2007; Epstein & Kanwisher, 1998; Grady, McIntosh, Rajah, & Craik, 1998; Gabrieli, Brewer, Desmond, & Glover, 1997). We therefore have a priori reason to predict a correlation between fixation duration and activity in this region. The main issue we addressed, then, was the nature of the cortical network that is associated with fixation duration during active scene viewing.

## METHODS

### Participants

Data from 31 right-handed participants (12 men), aged 18–35 years (mean = 21.45 years), were included in the analysis. Participants were recruited from the Columbia, South Carolina, community, were all native speakers of English, and reported normal or corrected-to-normal vision. All participants gave informed consent, were screened for MRI safety, and were given \$10 per hour for participation. Data from two additional participants were removed from analysis, one due to a technical problem with the scanner and the other due to falling asleep during the experiment.

### Stimuli

Twenty-two digitized photographs of real-world scenes taken from a large database of scene images compiled over several years were presented. All of these images depict tokens of common indoor and outdoor environments (Figure 1) and have been used in previous eye-tracking studies (Henderson et al., 2013; Nuthmann et al., 2010; Henderson & Pierce, 2008).

### Apparatus

Stimuli were presented using an Avotec (Stuart, FL) Silent Vision 6011 projector in its native resolution (1024 × 768) at a refresh rate of 60 Hz. Eye movements were recorded via an SR Research (Ontario, Canada) Eye-link 1000 long-range MRI eyetracker sampling at 1000 Hz. Viewing was binocular, and eye movements were recorded from one eye.

### Procedure

Participants were instructed to view each scene in preparation for a later memory test. Each functional run included 11 scene trials as well as 33 trials in which participants completed three tasks involving text that were not relevant to this study. Each scene (trial) was presented for 12 sec. Scenes were presented in random order across participants. Presentation of scene versus filler trials was randomized for each participant. An intertrial interval of 6 sec was inserted between each trial. Each functional run lasted about 14 min.

#### *Eye Movement Data Acquisition*

A 13-point calibration procedure was administered in the scanner before each functional run to map eye position to screen coordinates. Successful calibration required an average error of less than 0.49° and a maximum error of less than 0.99°. A fixation cross was presented in the center of the screen before each scene trial. Eye movements were recorded throughout each functional run.

#### *fMRI Data Acquisition*

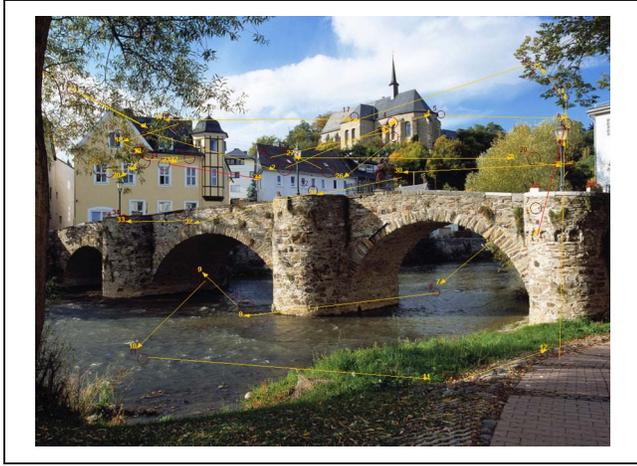
MR data were collected on a Siemens Medical Systems 3T Trio. A 3-D T1-weighted MP-RAGE RF-spoiled rapid flash scan in the sagittal plane and a T2/PD-weighted multi-slice axial 2-D dual Fast Turbo spin-echo scan in the axial plane were used. The multiecho whole-brain T1 scans had 1-mm isotropic voxel size and sufficient field of view to cover from the top of the head to the neck with the following protocol parameters: TR = 2530 msec, TE1 = 1.74 msec, TE2 = 3.6 msec, TE3 = 5.46 msec, TE4 = 7.32 msec, flip angle = 7°. All functional runs were acquired using gradient-echo, echo-planar images with the following protocol parameters: TR = 1850 msec, TE = 30 msec, flip angle = 75°. Volumes consisted of thirty-four 3-mm slices with transversal orientation. Each volume covered the whole brain with a field of view of 208 mm and 64 × 64 matrix, resulting in a 3.3 × 3.3 × 3.0 mm voxel size.

#### *fMRI Analysis*

The AFNI software package (Cox, 1996) was used for image analysis. Within-subject analysis involved slice timing correction, spatial coregistration (Cox & Jesmanowicz, 1999), and registration of functional images to the anatomy (Saad et al., 2009). Voxel-wise multiple linear regression was performed with the program 3dREMLfit, using reference functions representing each condition convolved with a standard hemodynamic response function. Reference functions representing the six motion parameters were included as covariates of no interest. In addition, the signal extracted from CSF was also included as noise covariates of no interest.

To examine the areas with activation correlated with fixation duration, an amplitude-modulated (parametric) regressor was used that contained the onset times (from the onset of each run) of each fixation and the duration of that fixation. A binary regressor coding the onset times of each fixation was also included in the regression model. There are multiple fixations within each TR. We take advantage of the fact that the timings of the fixations within each TR vary from TR to TR. This variation, combined with the large number of TRs, provides enough power to extract information from the low temporal resolution fMRI data based on the high-temporal resolution eye-tracking data. The ideal hemodynamic response resulting from this regressor was subsampled to match the time resolution of EPI images.

The individual statistical maps and the anatomical scans were projected into standard stereotaxic space (Talairach & Tournoux, 1988) and smoothed with a Gaussian filter of 5 mm FWHM. In a random effects analysis, group maps were created by comparing activations against a constant value of 0. The group maps were thresholded at voxelwise  $p < .01$  and corrected for multiple comparisons by removing clusters with below-threshold



**Figure 2.** Example scan pattern of one participant viewing one scene during the experiment. Purple lines represent saccades and circles represent fixations, with size of circles representing fixation durations. Fixation durations (in msec) are also shown. Yellow lines represent blinks.

size to achieve a mapwise corrected  $\alpha < 0.05$ . Using the 3dClustSim program with 10,000 iterations, the cluster threshold was determined through Monte Carlo simulations that estimate the chance probability of spatially contiguous voxels exceeding the voxelwise  $p$  threshold, that is, of false-positive noise clusters. The smoothness of the data was estimated with the AFNI program 3dFWHMx using regression residuals as input. The analysis was restricted to a mask that excluded areas outside the brain, as well as deep white matter areas and the ventricles.

#### *FIRE fMRI Analysis*

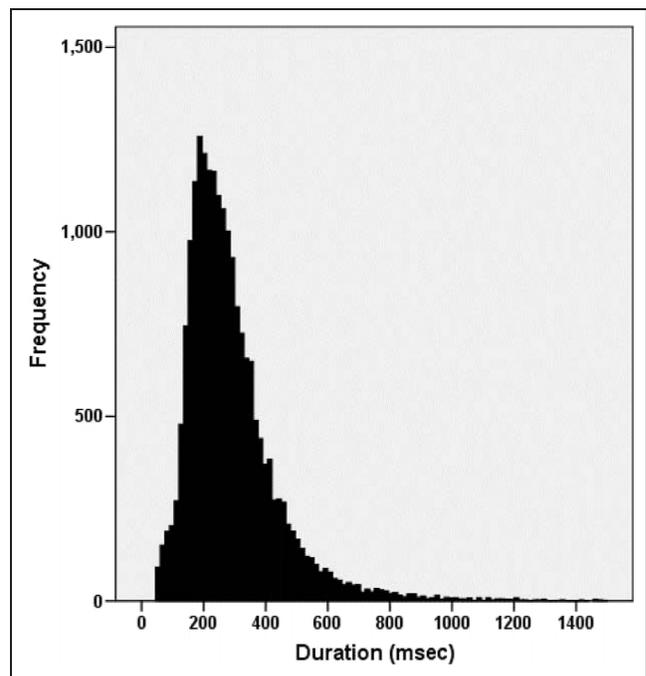
The eye movement data were analyzed offline to identify fixations and saccades using DataViewer (SR Research Ltd., version 1.11.1). All fixations meeting the following criteria were included in the analyses: A fixation could not be immediately preceded or followed by a blink, had to have a duration between 50 and 1500 msec, and could not follow a saccade greater than  $14^\circ$ . A total of 16.8% of fixations were excluded by these criteria (14% due to track losses and blinks). The fMRI and eye-tracking data were synchronized so that fixation onset from the eye-tracker could be aligned with the fMRI data. This was accomplished by aligning the onset of the trial run with the onset of the functional scan. Times of experiment onset, block onsets, and fixation onsets were saved in the eye movement record by the Experiment Builder program controlling the experiment. In addition, both the scanner time and eye-tracker time were recorded via a dedicated TCP/IP port to a separate data logger. This made it possible to coregister eye movement and fMRI events.

## RESULTS

Participants produced 20,527 fixations that met the inclusion criteria. Eye movement characteristics were similar to those typically observed in scene viewing studies outside the scanner (Henderson, 2011, 2013). Mean fixation duration was 290 msec ( $SD = 155$  msec), and mean saccade amplitude was  $2.73^\circ$  ( $SD = 2.18^\circ$ ). Figure 2 shows an example scan pattern of one participant viewing one scene, and Figure 3 shows the distribution of all included fixation durations.

The FIRE fMRI analysis produced activation correlated with fixation onset (Table 1 and Figure 4) and with fixation duration (Table 2 and Figure 5). The onset results demonstrate general increased activation in visual, attentional, and eye movement networks in occipital, temporal, parietal, and frontal areas, as expected. Specifically, there was bilateral occipital activation extending ventrally to lingual (LING), fusiform, and parahippocampal gyri (PHG) and to the hippocampus. Activation extended dorsally to right precuneus and parieto-occipital sulcus, bilateral superior parietal gyrus, intraparietal sulcus, postcentral sulcus, and right postcentral gyrus. Frontal activation was seen in left precentral sulcus, right precentral gyrus, left orbital gyrus, bilateral middle frontal gyrus including the FEFs, and bilateral superior frontal gyrus (SFG) including the supplementary eye field. Subcortical activation was observed in bilateral diencephalon, putamen, caudate, superior colliculus, and cerebellum.

Of primary theoretical interest was activation correlated with the durations of fixations (Figure 5). Activation was



**Figure 3.** Histogram showing the fixation duration distribution in the experiment collapsed over participants and scene images for all fixations included in the analyses.

**Table 1.** Activation Associated with Fixation Onset

<i>Fixation Onset</i>						
<i>Volume</i>	<i>Mean</i>	<i>Max</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>Cluster Coverage</i>
259578	4.38	8.612	28	-34	-9	R parahippocampal gyrus, R hippocampus, R lingual gyrus, R fusiform gyrus
		8.338	-34	-88	8	L middle occipital gyrus, L occipital pole
		8.22	-19	-25	0	L ventral diencephalon, L superior colliculus, L putamen
		7.982	31	-88	5	R middle occipital gyrus, R occipital pole
		7.898	-28	-49	-12	L fusiform gyrus, L parahippocampal gyrus, L hippocampus, L lingual gyrus
		7.668	-13	-82	-12	L lateral occipital gyrus
		7.265	-10	-58	2	L lingual gyrus
		7.047	-22	-73	41	L superior parietal gyrus, L intraparietal sulcus
		7.027	19	-55	14	R precuneus, R parieto-occipital sulcus
		6.552	22	-70	47	R superior parietal gyrus, R intraparietal sulcus
		5.882	31	-7	-3	R putamen, R ventral diencephalon, R superior colliculus
		5.035	-19	-34	-36	L cerebellum
		4.873	22	-31	-36	R cerebellum
		4.452	13	4	14	R/L caudate
		3.98	-37	-43	47	L postcentral sulcus, L superior parietal gyrus
9612	3.18	4.756	55	22	29	R middle frontal gyrus, R precentral gyrus
		4.074	61	-7	29	R postcentral gyrus
7209	3.6	5.191	-28	-4	50	L medial frontal eye field
6696	3.22	5.095	-43	7	32	L precentral sulcus
4320	3.31	4.798	-1	4	59	Supplementary eye field, R/L superior frontal gyrus
3429	3.42	4.809	25	-1	53	R medial frontal eye field
2700	3.27	4.163	37	-34	41	R postcentral sulcus
1512	3.5	5.559	-43	43	0	L orbital gyrus
972	3.36	4.513	13	-40	-39	R cerebellum

Brain regions with activation associated with fixation onset during scene viewing. Locations of peak activation are shown for each cluster with significant activity ( $p < .01$  corrected for multiple comparisons). Multiple peaks required separation by a minimum of 25 voxels. The volume of the cluster ( $\mu$ l), peak  $z$  score, Talairach coordinates, and anatomical structures are shown. L = left hemisphere; R = right hemisphere.

positively correlated with fixation duration in bilateral occipital pole, middle occipital gyrus, cuneus, LING, right dorsolateral pFC (DLPFC), and right middle frontal gyrus. Activation was negatively correlated with fixation duration in right hippocampus, PHG, and paracentral lobule. Additional negatively correlated activation was observed in bilateral cerebellum and right amygdala.

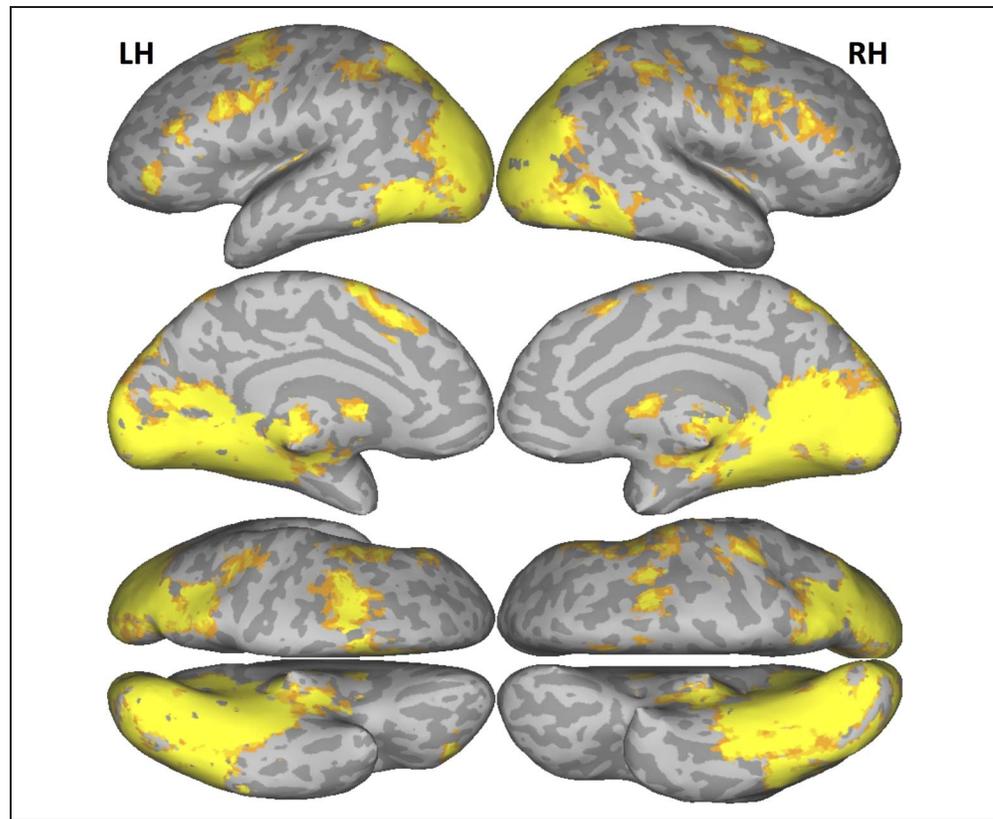
## DISCUSSION

Prior behavioral and computational research has provided strong evidence for direct control of fixation duration during scene viewing. Fixation duration increases in real

time when perceptual and cognitive processing is more difficult and decreases when these processes are easier. Yet the neural correlates of the direct control of eye movements during scene viewing have not been investigated. In this study, we coregistered eye tracking and fMRI to investigate the neurocognitive systems associated with fixation duration in scene viewing. We simultaneously recorded and coregistered high-resolution eye movements and fMRI while participants viewed digitized photographs of real-world scenes. We then used FIRE analysis of the fMRI data to identify neural activation associated with the duration of fixation.

Our main theoretical question concerned the relationship between neural activation and fixation duration.

**Figure 4.** Areas of activation in a whole-brain analysis significantly correlated with fixation onset during scene viewing. The maps are displayed using Caret (Van Essen et al., 2001) on an inflated cortical surface of a representative participant, with gyri shown as light gray and sulci shown as dark gray. Hot regions show positive correlation, and cool regions show negative correlation (threshold  $p < .01$  corrected for multiple comparisons).



Correlations were observed in several cortical regions. First, positive correlations were seen in early visual areas and in frontal areas. The positive correlation in visual areas (bilateral occipital pole, middle occipital gyrus, cuneus, and LING) is consistent with the well-described relationship between increased fixation duration and greater levels of visual processing (Henderson, 2011;

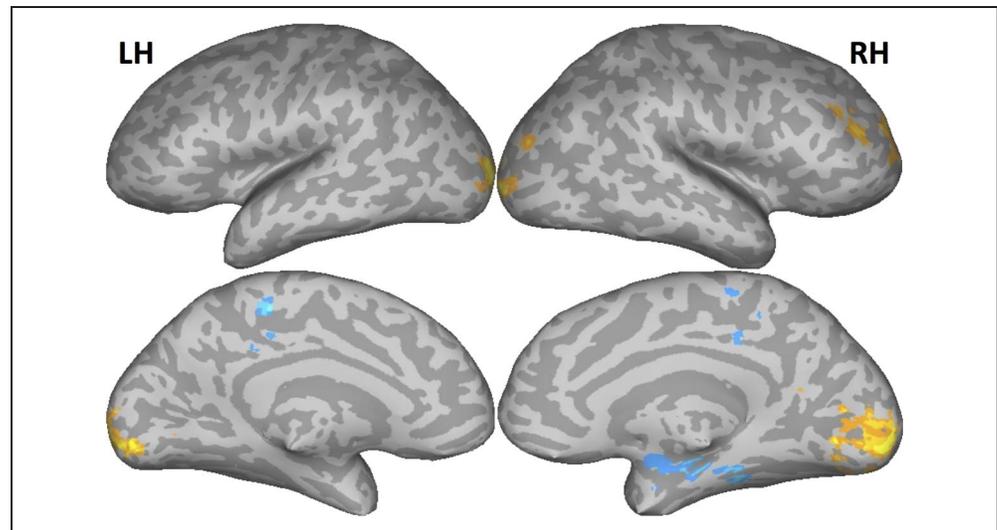
Rayner, 2009). This general relationship is observed across a range of tasks including scene viewing, reading, and visual search. As noted in the Introduction, effects of visual encoding difficulty on fixation duration during scene viewing are also observed in real time, with manipulations that remove or degrade the image producing immediate fixation duration effects. The present results

**Table 2.** Activation Associated with Fixation Duration

<i>Fixation Duration</i>						
<i>Volume</i>	<i>Mean</i>	<i>Max</i>	<i>x</i>	<i>y</i>	<i>z</i>	
<i>Positive Correlation</i>						
12582	3.17	4.688	-16	-97	2	L/R occipital pole, L/R middle occipital gyrus, L/R cuneus, L/R lingual gyrus
2214	3.01	3.966	28	55	23	R middle frontal gyrus/sulcus, R dorsolateral prefrontal cortex
<i>Negative Correlation</i>						
1647	-2.93	-4.242	10	-28	-33	Brain stem, R cerebellum
		-3.959	-4	-46	-42	L cerebellum
1431	-2.99	-3.834	25	-22	-12	R hippocampus, R parahippocampal gyrus, R amygdala
918	-2.88	-3.887	-4	-31	44	L/R paracentral lobule

Brain regions with activation correlated with fixation duration during scene viewing. Locations of peak activation are shown for each cluster with significant activity ( $p < .01$  corrected for multiple comparisons). Multiple peaks required separation by a minimum of 25 voxels. The volume of the cluster ( $\mu$ l), peak  $z$  score, Talairach coordinates, and anatomical structures are shown. L = left hemisphere; R = right hemisphere.

**Figure 5.** Areas of activation in a whole-brain analysis significantly correlated with fixation duration during scene viewing. Hot regions show positive correlation, and cool regions show negative correlation ( $\alpha < .01$  FWE corrected).



suggest that this behavioral relationship between visual encoding and fixation duration is reflected in increased activation in visual areas, with increased activation associated with longer durations. This result is also consistent with the recent finding that fixation duration during reading is associated with the structure of early visual cortex (Henderson, Choi, & Luke, 2014).

Positive correlation between fixation duration and neural activity was also observed in right DLPFC, a region typically associated with working memory and executive functions such as task-switching and inhibition. Activation of right DLPFC is also often observed during oculomotor control tasks that require inhibiting and modifying saccade programs (Pierrot-Deseilligny, Müri, Nyffeler, & Milea, 2005; Munoz & Everling, 2004; Pierrot-Deseilligny, Milea, & Müri, 2004). Increased activation of DLPFC associated with increased fixation duration is consistent with real-time control of fixation duration during scene viewing based on executive and inhibitory processes associated with monitoring ongoing visual and cognitive processing difficulty.

In contrast to the positive correlations with visual and prefrontal activation, negative correlations were observed in regions typically associated with higher-level visual scene processing and encoding, including PHG and the hippocampus. It is well known that the hippocampus plays a role in spatial processing (Burgess, Maguire, & O'Keefe, 2002; O'Keefe & Nadel, 1978). There is also a large body of evidence suggesting that PHG is associated with computations that support high-level visual scene analysis and encoding (Staresina et al., 2011; Henderson et al., 2007; Epstein & Kanwisher, 1998; Grady et al., 1998; Gabrieli et al., 1997). Perhaps more controversial is the hypothesis that the hippocampus plays a direct role in visual scene processing (Graham, Barense, & Lee, 2010), with evidence suggesting that it may be involved in visual analysis particularly when complex conjunctions and relational visual information are encoded

(Aly, Ranganath, & Yonelinas, 2013; Lee, Yeung, & Barense, 2012).

The present results suggest that the known pairwise relationships between scene analysis and fixation duration, and scene analysis and PHG/hippocampal activation extend to fixation duration and PHG/hippocampal activation. Although we cannot draw causal inferences from the present data, it is tempting to speculate that less scene-associated activation was associated with poorer scene encoding, leading to a signal to delay saccade programming and extend the current fixation as a way to boost encoding. Note that this general scheme is generic to computational models that account for fixation durations in other domains like reading. For example, E-Z Reader posits that reduced activation in lexical processing (the L1 module in the model) results in a delay in the initiation of saccade programming and therefore to increased fixation duration (Reichle, Pollatsek, & Rayner, 2006; Reichle et al., 1998).

An alternative account for the negative correlation between fixation duration and activation in medial-temporal structures is that this activation reflected the operation of the default mode network (Buckner, Andrews-Hanna, & Schacter, 2008; Mason et al., 2007; Raichle & Snyder, 2007; Raichle et al., 2001). On this explanation, increased fixation durations and the associated increases in visual and attentional processes caused decreased activity in the default mode. Although we cannot rule out this possibility, it is interesting to note that other regions associated with the default mode network did not produce negatively correlated activation. Also, in a related study in which we examined fMRI activation correlated with fixation duration in natural reading, we did not observe negatively correlated activation in medial-temporal structures, although we did observe positive correlations in occipital areas similar to this study. In light of these considerations, it seems more parsimonious to suppose that the medial-temporal activation negatively correlated with

fixation duration observed here was associated with scene-specific computations.

Finally, the negative correlations between activation and fixation duration in bilateral medial SFG and paracentral lobule is generally consistent with Hillen and colleagues, who observed SFG activation in an analysis of a common gaze control network for reading and several pseudoreading tasks (Hillen et al., 2013). Medial SFG is also associated with motor task-switching and response inhibition, and switching between action plans (Rushworth, Buckley, Behrens, Walton, & Bannerman, 2007; Taylor, Nobre, & Rushworth, 2007). These processes are likely to be involved in the type of sequential eye movement control that is needed during real-world scene viewing, consistent with the CRISP model in which saccade programs can be modified and even cancelled based on real-time scene interpretation (Nuthmann et al., 2010).

The overall pattern of results can be accommodated by an account that is consistent with CRISP: Fixations in which scene analysis and encoding are more difficult are associated with less higher-level scene interpretation and memory encoding (less activation) and require greater involvement of executive and inhibitory functions to control fixation durations (greater activation), which increases fixation duration and therefore increases visual encoding (greater activation). In the vocabulary of CRISP, increased difficulty in scene analysis generates inhibitory control signals that feed back to reduce the accumulation rate of the random walk timer and inhibit completed saccade programs, which increases fixation duration and therefore visual encoding time within that fixation. This interpretation is clearly underconstrained by the present data, but it is consistent with the large body of existing behavioral evidence, an existing computational theory designed to account for those data, and the current pattern of imaging data.

In summary, eye-tracking research has demonstrated that much of the variability in fixation duration during active scene viewing is accounted for by attentional, perceptual, and cognitive processes associated with the scene (Henderson, 2011; Rayner, 2009; Henderson et al., 1999). We have demonstrated that fixation duration during scene viewing correlates with activity in cortical structures associated with attention and with scene analysis and encoding. These results support the view that fixation duration in scene perception reflects underlying real-time scene processing, consistent with the assumptions of current computational models of active vision for scene perception such as CRISP (Nuthmann et al., 2010). The results also demonstrate that FIRE fMRI can provide a unique source of converging evidence constraining theories of active scene viewing and for grounding these theories in neurobiology. This study sets the stage for using FIRE fMRI to investigate other topics related to the neurocognition of natural scene viewing as well as active vision tasks more generally.

## Acknowledgments

This research was supported by the National Science Foundation (BCS-1151358). We thank Simon Liversedge, Eyal Reingold, and Andy Yonelinas for their comments on and discussion of this research and William Brixius, Jennifer Olejarczyk, and Joseph Schmidt for their help with stimulus preparation and data collection.

Reprint requests should be sent to John M. Henderson, Institute for Mind and Brain, University of South Carolina, 1800 Gervais Street, Columbia, SC 29208, or via e-mail: Jhenderson.psy@gmail.com.

## REFERENCES

- Aly, M., Ranganath, C., & Yonelinas, A. P. (2013). Detecting changes in scenes: The hippocampus is critical for strength-based perception. *Neuron*, *78*, 1127–1137.
- Becker, W., & Jurgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, *19*, 967–983.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38.
- Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*, 625–641.
- Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Castelano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*, 6.1–6.15.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Cox, R. W., & Jesmanowicz, A. (1999). Real-time 3D image registration of functional MRI. *Magnetic Resonance in Medicine*, *42*, 1014–1018.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Gabrieli, J. D., Brewer, J. B., Desmond, J. E., & Glover, G. H. (1997). Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science*, *276*, 264–266.
- Glaholt, M. G., Rayner, K., & Reingold, E. M. (2013). Spatial frequency filtering and the direct control of fixation durations during scene viewing. *Attention, Perception, & Psychophysics*, *75*, 1761–1773.
- Glaholt, M. G., & Reingold, E. M. (2012). Direct control of fixation times in scene viewing: Evidence from analysis of the distribution of first fixation duration. *Visual Cognition*, *20*, 605–626.
- Grady, C. L., McIntosh, A. R., Rajah, M. N., & Craik, F. I. (1998). Neural correlates of the episodic encoding of pictures and words. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 2703–2708.
- Graham, K. S., Barense, M. D., & Lee, A. C. H. (2010). Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia*, *48*, 831–853.
- Henderson, J. M. (2011). Eye movements and scene perception. In S. P. Liversedge, I. D. Gilchrist, & S. Everling, *The Oxford handbook of eye movements* (pp. 593–606). New York: Oxford University Press.
- Henderson, J. M. (2013). Eye movements. In D. Reisberg, *The Oxford handbook of cognitive psychology* (pp. 69–82). New York: Oxford University Press.

- Henderson, J. M., Choi, W., & Luke, S. G. (2014). Morphology of primary visual cortex predicts individual differences in fixation duration during text reading. *Journal of Cognitive Neuroscience*, *26*, 2880–2888.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.
- Henderson, J. M., Larson, C. L., & Zhu, D. C. (2007). Cortical activation to indoor versus outdoor scenes: An fMRI study. *Experimental Brain Research*, *179*, 75–84.
- Henderson, J. M., Nuthmann, A., & Luke, S. G. (2013). Eye movement control during scene viewing: Immediate effects of scene luminance on fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 318–322.
- Henderson, J. M., Olejarczyk, J., Luke, S. G., & Schmidt, J. (2014). Eye movement control during scene viewing: Immediate degradation and enhancement effects of spatial frequency filtering. *Visual Cognition*, *22*, 486–502.
- Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin & Review*, *15*, 566–573.
- Henderson, J. M., & Smith, T. J. (2009). How are eye fixation durations controlled during scene viewing? Further evidence from a scene onset delay paradigm. *Visual Cognition*, *17*, 1055–1082.
- Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228.
- Hillen, R., Günther, T., Kohlen, C., Eckers, C., van Ermingen-Marbach, M., Sass, K., et al. (2013). Identifying brain systems for gaze orienting during reading: fMRI investigation of the Landolt paradigm. *Frontiers in Human Neuroscience*, *7*, article 384.
- 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.
- Land, M., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*, 3559–3565.
- Lee, A. C., Yeung, L. K., & Barse, M. D. (2012, April 17). The hippocampus and visual perception. *Frontiers in Human Neuroscience*, *6*, 91.
- Loftus, G. (1985). Picture perception: Effects of luminance on available information and information-extraction rate. *Journal of Experimental Psychology General*, *114*, 342–356.
- Luke, S. G., Nuthmann, A., & Henderson, J. M. (2013). Eye movement control in scene viewing and reading: Evidence from the stimulus onset delay paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 10–15.
- Marsman, J. B. C., Renken, R., Velichkovsky, B. M., Hooymans, J. M. M., & Cornelissen, F. W. (2011). Fixation based event-related fMRI analysis: Using eye fixations as events in functional magnetic resonance imaging to reveal cortical processing during the free exploration of visual images. *Human Brain Mapping*, *33*, 307–318.
- Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science (New York, NY)*, *315*, 393–395.
- Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision*, *11*, 1–15.
- Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 667–682.
- Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*, 218–228.
- Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review*, *117*, 382–405.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map* (Vol. 3). Oxford: Clarendon Press.
- Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Current Opinion in Neurology*, *17*, 17–25.
- Pierrot-Deseilligny, C., Müri, R. M., Nyffeler, T., & Milea, D. (2005). The role of the human dorsolateral prefrontal cortex in ocular motor behavior. *Annals of the New York Academy of Sciences*, *1039*, 239–251.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 676–682.
- Raichle, M. E., & Snyder, A. Z. (2007). A default mode of brain function: A brief history of an evolving idea. *Neuroimage*, *37*, 1083–1090.
- Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157.
- Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7*, 4–22.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *The Behavioral and Brain Sciences*, *26*, 445–476.
- Richlan, F., Gagl, B., Hawelka, S., Braun, M., Schurz, M., Kronbichler, M., et al. (2014). Fixation-related fMRI analysis in the domain of reading research: Using self-paced eye movements as markers for hemodynamic brain responses during visual letter string processing. *Cerebral Cortex*, *24*, 2647–2656.
- Rushworth, M. F., Buckley, M. J., Behrens, T. E., Walton, M. E., & Bannerman, D. M. (2007). Functional organization of the medial frontal cortex. *Current Opinion in Neurobiology*, *17*, 220–227.
- Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., & Cox, R. W. (2009). A new method for improving functional-structural MRI alignment using local Pearson correlation. *Neuroimage*, *4*, 839–848.
- Staresina, B. P., Duncan, K. D., & Davachi, L. (2011). Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. *Journal of Neuroscience*, *31*, 8739–8747.
- Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. New York: Thieme.
- Taylor, P. C. J., Nobre, A. C., & Rushworth, M. F. S. (2007). Subsecond changes in top-down control exerted by human medial frontal cortex during conflict and action selection: A combined transcranial magnetic stimulation electroencephalography study. *Journal of Neuroscience*, *27*, 11343–11353.
- Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, *8*, 443–459.
- Walshe, R. C., & Nuthmann, A. (2014). Asymmetrical control of fixation durations in scene viewing. *Vision Research*, *100*, 38–46.
- Williams, C. C., Henderson, J. M., & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception*, *67*, 816–827.