

The time course of initial scene processing for eye movement guidance in natural scene search

Melissa L.-H. Võ

Visual Attention Lab, Harvard Medical School, USA, &
Psychology Department, University of Edinburgh, UK



John M. Henderson

Psychology Department, University of Edinburgh, UK



A brief glimpse of a scene is sufficient to comprehend its gist. Does information available from a brief glimpse also support further scene exploration? In five experiments, we investigated the role of initial scene processing on eye movement guidance for visual search in scenes. We used the flash-preview moving-window paradigm to separate the duration of the initial scene glimpse from subsequent search. By varying scene preview durations, we found that a 75-ms preview was sufficient to lead to increased search benefits compared to a no-preview control. Search efficiency was further increased by inserting additional scene-target integration time before search initiation: Reducing preview durations to as little as 50 ms led to search benefits only when combined with prolonged integration time. We therefore propose that both initial scene presentation duration and scene-target integration time are crucial for establishing contextual guidance in complex, naturalistic scenes. The present findings show that fast scene processing is not limited to activating gist. Instead, scene representations generated from a brief scene glimpse can also provide sufficient information to guide gaze during object search as long as enough time is available to integrate the initial scene representation.

Keywords: eye movement guidance, real-world scenes, visual search, flash-preview moving-window paradigm, preview duration, integration time, scene gist, cognitive knowledge structures

Citation: Võ, M. L.-H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision*, 10(3):14, 1–13, <http://journalofvision.org/10/3/14/>, doi:10.1167/10.3.14.

Introduction

We are able to extract the gist of a scene from a very brief glimpse (e.g., Biederman, 1981; Biederman, Mezzanotte, & Rabinowitz, 1982; Castelano & Henderson, 2008; Fei-Fei, Iyer, Koch, & Perona, 2007; Greene & Oliva, 2009; Intraub, 1980; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Oliva, 2005; Oliva & Torralba, 2006; Potter, 1975; Rousselet, Joubert, & Fabre-Thorpe, 2005; Schyns & Oliva, 1994; Thorpe, Fize, & Marlot, 1996; Van Rullen & Thorpe, 2001). However, we usually do not stop scene processing at the point of gist identification. Instead, we use the information extracted from the first scene glimpse to provide the context within which to further explore and plan subsequent actions over the scene.

Recent evidence suggests that the information from the initial glimpse contains sufficient structural and semantic information to support subsequent active exploration of the scene via eye movements (Castelano & Henderson, 2007). This finding requires that the initial scene representation is sufficiently detailed and survives long enough to support subsequent eye movement planning. Eye movements then allow further elaboration of the initial scene representation, which in turn can guide subsequent interaction (e.g., Castelano & Henderson, 2007; Friedman, 1979; Hollingworth, 2005; Hollingworth & Henderson, 2002; Tatler, Gilchrist, & Rusted, 2003). The aim of the

present study was to shed light on the time course of initial scene processing for guiding action. We were specifically interested in the time course between the initial glimpse of a scene and the initiation of object search via eye movements.

Our ability to rapidly recognize scenes within only a short glimpse has been demonstrated repeatedly over the years. In an early study, Potter (1975) found that a presentation duration of 125 ms was sufficient to allow for above-chance identification of target scenes embedded in a series of distractor scenes. Subsequent work showed that although semantic understanding of a scene was quickly extracted, additional time was needed to consolidate the scene in memory (Potter, 1976; see Intraub, 1980). To date, studies investigating the time course of initial scene processing have mainly focused on the speed at which visual information can be processed to allow for rapid scene categorization or object identification (e.g., Joubert et al., 2007; Rousselet et al., 2005; Thorpe et al., 1996; Van Rullen & Thorpe, 2001). Other studies investigating the temporal constraints of initial scene processing have focused on the minimum presentation duration (e.g., Greene & Oliva, 2009) or stimulus onset asynchrony between image and mask (e.g., Bacon-Macé, Macé, Fabre-Thorpe, & Thorpe, 2005) needed for above-chance scene categorization. Together, these studies have provided compelling evidence that sophisticated scene analysis can be accomplished with scene presentation

durations as brief as 50 ms. However, the time course of early scene processing with regard to its influence on eye movement planning has largely been neglected. While the rapid extraction of global low-level features may suffice for early scene identification (e.g., Greene & Oliva, 2009; Joubert et al., 2007; Schyns & Oliva, 1994; Tatler et al., 2003), the computation of goal-directed eye movements toward a probable target location might require more detailed conceptual analysis along with a greater degree of processing of the scene's spatial layout.

Kirchner and Thorpe (2006) investigated the speed at which initial scene processing can feed into saccadic programming using a forced-choice saccade task. Results showed that when two scenes were simultaneously flashed for 20 ms, participants were able to reliably make saccades to the side of the screen that contained an animal in as little as 120 ms, implying that the visual system only needs roughly 95–100 ms (allowing for saccade programming time) to provide an initial first pass analysis of scene images to produce a reliable eye movement response. While these effects speak to the capability of ultra-rapid scene processing and saccade programming, they do not provide information about temporal factors that determine the contextual guidance of eye movements within a scene based on semantic knowledge. More complex saccade programming to multiple possible target locations, such as when searching for objects embedded in scenes, might require a greater amount of scene analysis.

Rayner, Smith, Malcolm, and Henderson (2009) manipulated the amount of time a scene was visible during each fixation while participants performed an object search in naturalistic scenes. They found that participants needed to see the scene for at least 150 ms during each fixation to normally process a scene and plan eye movements. While these findings provide information on the minimum amount of time necessary for scene analysis to support saccadic programming during each fixation, the control of fixational eye movements during scene viewing is complex and probably involves a multitude of processes such as ongoing information processing of the previous fixation and foveal as well as extrafoveal processing during the current fixation. Thus, it is not clear from this study how quickly information relevant to eye movement planning is acquired in the initial glimpse of a scene.

In the study presented here, we focused on the time course of naturalistic scene processing from the initial glimpse of a scene until the initiation of object search. We were specifically interested in the minimum amount of scene presentation time needed to provide sufficient information to subsequently guide eye movements to probable target locations during search. We were also interested in whether establishing an initial scene representation might depend on the time available to consolidate and integrate the initial scene representation (e.g., Intraub, 1980; Potter, 1976). In this study, we therefore manipulated both scene presentation durations and subsequent integration time before the initiation of search.

To investigate the time course of initial scene processing, we used the flash-preview moving-window paradigm introduced by Castelhamo and Henderson (2007). This paradigm has been successfully applied to investigate the influence of the initial glimpse of a scene on subsequent eye movement control during search (Castelhamo & Henderson, 2007; Võ & Schneider, 2010). In this paradigm, participants are first presented with a brief preview of the search scene, followed by the presentation of a target word indicating which object they will be looking for. The scene is then presented again for search, but participants are able to view the scene through a gaze-contingent window that reveals only a small area of the scene tied to the current fixation location. Therefore, this paradigm allows the effect of the initial scene glimpse on subsequent eye movements to be isolated from the processing that takes place during later stages of scene viewing. It also allows the independent manipulation of the duration that a scene is initially presented and the delay between initial scene presentation and search initiation. These two factors—preview duration and integration time—were explored in the present study.

It has been shown that flashing a scene preview for 250 ms—a duration that is in the range of a typical fixation duration in scene perception (see Henderson, 2003; Rayner, 1998)—provides sufficient time to process the visual input to such a degree that scene knowledge can provide contextual guidance about where to find a certain target object in a scene (Castelhamo & Henderson, 2007; Võ & Schneider, 2010). In this study, we presented scene previews of 100 ms (Experiment 1a), 75 ms (Experiment 1b), and 50 ms (Experiment 1c) to investigate the minimum preview duration required for efficient object search. We also investigated whether providing additional integration time following the flashed scene preview and target word but before initiation of search can enhance the preview benefit (Experiments 2 and 3). Together, these experiments provide further insight into the time course of early scene processing for subsequent eye movement control.

General methods

Stimulus material

Forty-five full-color images of real-world scenes were presented in Experiments 1a–1c using a 1×3 design. For Experiments 2 and 3, 44 of the 45 scenes were used to accommodate a 2×2 design. Each scene was only presented once and experimental conditions were randomized across participants. Scenes were displayed on a 21-inch computer screen (resolution 800×600 pixels, 140 Hz) subtending visual angles of 25.66° (horizontal) and 19.23° (vertical) at a viewing distance of 90 cm. Targets

were positioned in highly context constraint scene locations. Previews of the search scene never included the target object.

Apparatus

Eye movements were recorded with an EyeLink1000 tower system (SR Research, Canada) at a sampling rate of 1000 Hz. The position of the right eye was tracked while viewing was binocular. Experimental sessions were carried out on a computer running OS Windows XP. Stimulus presentation and response recording was controlled by Experiment Builder (SR Research, Canada).

Procedure

The procedure of all experiments closely followed the procedure of the basic flash-preview moving-window paradigm (Castelhano & Henderson, 2007). Each participant was first informed that they would be shown a series of scenes in which they had to find predefined target objects as fast as possible. They were also informed that short previews of the scenes would precede the display of the search scene and that they should attend to these previews because they could provide helpful information.

After an initial fixation cross, the scene preview appeared and was subsequently masked. Following the mask, a target word was presented, which indicated the search target object. Following the target word, a refixation cross appeared to center gaze before presentation of the search scene.

In [Experiments 2](#) and [3](#), the duration of the refixation cross was manipulated to vary the delay before initiation of search. The search scene itself could only be explored through a 5° circular gaze-contingent window with the rest of the scene masked (see [Figure 1](#) for prototype trial sequence). The window size was set at 5° for both theoretical and practical reasons. From a theoretical perspective, 5° covers the fovea and parafovea, the regions where most high-resolution visual analysis takes place. Practically, we have found that 5° is about the smallest window region size that can be used without severely disrupting search.

Since we manipulated the timing of the trial sequences across experiments, these will be described in further detail within the appropriate sections of the experiments.

Eye movement data analysis

The interest area for each target object was defined as the rectangular box that was large enough to encompass that object. Fixation durations of less than 90 ms and more than 1000 ms were excluded as outliers. Raw data were subsequently filtered using SR Research Data Viewer. To investigate whether the temporal manipulations affected eye movements, a set of measures was calculated to analyze viewers' eye movement behavior. These measures were response time, latency to first target fixation, number of fixations to first target fixation, initial saccade latency, and initial saccade amplitude. While RT, latency, and number of fixations to target fixation provide information on the efficiency of target search, initial saccade latency and amplitude regardless of their direction reflect the general readiness to initiate search.

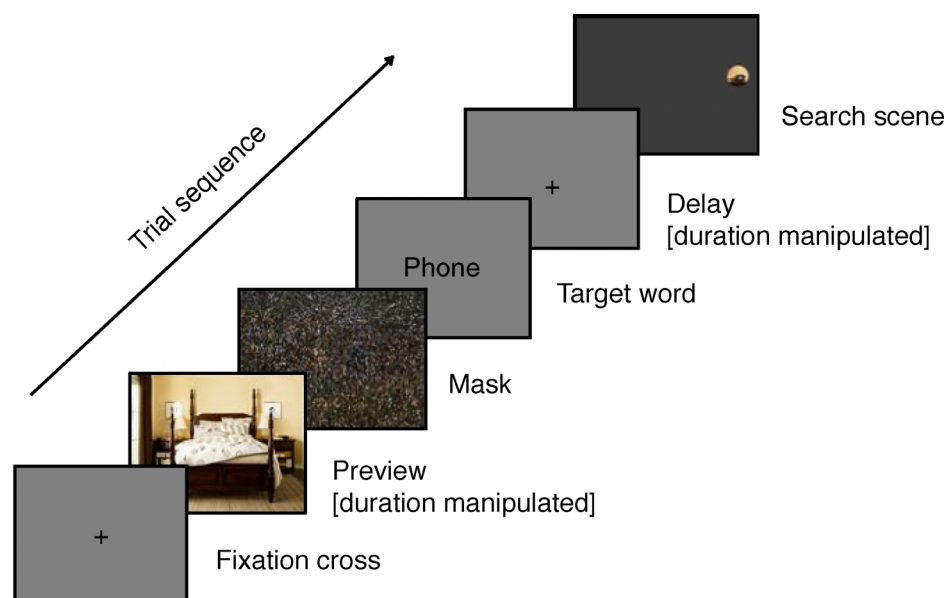


Figure 1. Trial sequence of the flash-preview moving-window paradigm.

Experiment 1a

Previous research using the flash-preview moving-window paradigm has shown that a 250-ms preview is sufficient to produce a preview benefit in object search. However, we do not yet know the lower limit on the duration of the preview. On the one hand, scene gist can be extracted from a glimpse that is much shorter than 250 ms (e.g., Biederman, 1981; Castelano & Henderson, 2008; Greene & Oliva, 2009; Intraub, 1980; Potter, 1975; Rousselet et al., 2005; Schyns & Oliva, 1994; Thorpe et al., 1996). On the other hand, the scene preview effect does not appear to be driven only by gist but also by abstract structural information (Castelano & Henderson, 2007). Can a preview benefit be observed for previews shorter than 250 ms? In [Experiment 1a](#), we tested whether a preview duration of 100 ms would suffice to provide scene preview benefits. We therefore compared scene preview benefits for preview durations of 100 ms with preview durations of 0 ms and 250 ms.

Methods

Participants

Fifteen native English-speaking students (10 females) from the University of Edinburgh ranging in age between 18 and 31 ($M = 21.9$, $SD = 3.14$) participated in [Experiment 1a](#) for £6/h. All participants reported normal or corrected-to-normal vision. One participant had to be replaced due to unstable recording of the eye.

Procedure

At the beginning of the experiment, the eye tracker was calibrated for each participant using a 9-point calibration and validation method. The participant's viewing position was fixed with a chin and forehead rest. Each trial sequence was preceded by a fixation check, i.e., in order to initiate the next trial, the participant had to fixate a cross centered on the screen for 200 ms. When the fixation check was deemed successful, the fixation cross was

replaced by the presentation of the scene's preview for either 0 ms, 100 ms, or 250 ms. After the presentation of a mask for 50 ms, a black target word that indicated the identity of the target object was displayed at the center of the gray screen for 1500 ms. A word rather than a picture of the target object was chosen to avoid the influence of a specific target template on subsequent search (Malcolm & Henderson, 2009). Following offset of the target word, a fixation cross was presented for 500 ms. The search scene was then shown through a 5° diameter circular window that moved with the participants' fixation location. The display beyond the window was a gray field. Thus, no peripheral vision was possible throughout the entire visual search. Participants had to search the scene for the target object and indicate its detection by holding fixation on the object and pressing a response button. The search scene was displayed until button press or for a maximum of 15 s. Five practice trials at the beginning of the experiment allowed participants to become accustomed to the experimental setup and the gaze-contingent window. The experiment lasted about 20 min.

Results

We compared scene preview benefits as a function of preview durations of 0 ms, 100 ms, and 250 ms using an analysis of variance (ANOVA) with preview duration as a within-subject factor. A summary of means can be seen in [Table 1](#).

Response time. RT was defined as the time elapsed from scene onset until button press and averaged 4151 ms across conditions. RT was significantly affected by the preview duration manipulation, $F(2,14) = 6.11$ in that RTs for both the 100-ms and 250-ms preview durations were diminished compared to the control condition, $t(14) = 4.01$, $p < 0.01$ and $t(14) = 2.74$, $p < 0.01$, respectively, while RTs for the 100-ms and 250-ms preview durations did not differ from each other, $t < 1$.

Latency to first target fixation. Latency was measured from search scene onset until the first fixation of the target object and averaged 3356 ms across all conditions. There

	Preview durations		
	0 ms	100 ms	250 ms
Response time in ms	4858 [263]	3970 [307]	3624 [338]
Latency to first target fixation in ms	3932 [250]	3197 [275]	2938 [337]
Number of fixations until target fixation	14.63 [1.01]	12.04 [1.07]	10.6 [1.12]
Initial saccade latency in ms	279 [12]	268 [12]	255 [13]
Initial saccade amplitude in degree visual angle	2.12 [0.17]	2.40 [0.14]	2.37 [0.11]

Table 1. Summary of mean values [standard errors] of [Experiment 1a](#) for the dependent variables (response time, latency to first target fixation, number of fixations to target fixation, initial saccade latency, and initial saccade amplitude) as a function of preview duration (0 ms vs. 100 ms vs. 250 ms).

was an effect of preview duration, $F(2,14) = 3.64$, $p < 0.05$. Latencies for both the 100-ms and 250-ms preview durations were significantly diminished compared to the control condition, $t(14) = 3.04$, $p < 0.01$ and $t(14) = 2.08$, $p < 0.05$, respectively, while the 100-ms and 250-ms preview conditions did not differ from each other, $t < 1$.

Number of fixations to first target fixation. This measure was defined as the number of discrete fixations until the target object was first fixated. The value does not include the initial scene fixation centered on the screen but does include the first fixation on the target object. On average, participants performed 12.42 fixations to the first fixation of the target object. There was again an effect of preview duration, $F(2,14) = 4.22$, $p < 0.05$. Preview durations of both 100 ms and 250 ms diminished the number of fixations to first target fixation compared to the control condition, $t(14) = 2.76$, $p < 0.01$ and $t(14) = 2.32$, $p < 0.05$, respectively, while the 100-ms and 250-ms preview durations did not differ from each other, $t(14) = 1.00$, $p > 0.05$.

Initial saccade latency. Initial saccade latency was measured from scene onset until the initiation of the first saccade and averaged 267 ms across conditions. There was a trend for an effect of preview duration, $F(2,14) = 2.75$, $p = 0.08$, in that initial saccade latency gradually declined from no preview over 100-ms to 250-ms preview.

Initial saccade amplitude. Initial saccade amplitude was measured as the length of the first saccade after search scene onset and averaged 2.30° visual angle across conditions and was significantly modulated by preview duration, $F(2,14) = 3.23$, $p < 0.05$. Participants made significantly longer initial saccade amplitudes following both the 100-ms and 250-ms preview conditions compared to the control condition, $t(14) = 2.71$, $p < 0.01$ and $t(14) = 1.69$, $p < 0.05$, respectively, while the 100-ms and 250-ms preview durations did not differ from each other, $t < 1$.

Summary

Experiment 1a shows that sufficient information can be acquired from a 100-ms glimpse of a scene to facilitate eye movement planning for visual search. With a 100-ms scene preview, participants were faster to subsequently move their eyes to the target object compared to the condition where no preview was shown. There was also a qualitative tendency for the preview benefit to be larger with a 250-ms than 100-ms preview, suggesting that visual information continues to accumulate beyond the first 100 ms of scene presentation, but these differences failed to reach statistical significance.

Experiment 1b

Since 100-ms previews of the search scene already led to significant scene preview benefits, we further decreased

preview durations to 75 ms in **Experiment 1b** to test whether search benefits would still be observable.

Methods

Participants

Fifteen native English-speaking students (8 females) from the University of Edinburgh ranging in age between 19 and 24 ($M = 22.1$, $SD = 1.71$) participated in **Experiment 1b** for £6/h. All participants reported normal or corrected-to-normal vision and none had taken part in **Experiment 1a**. One participant had to be replaced due to misunderstanding of target words.

Procedure

The procedure was identical to the one in **Experiment 1a** with the exception that we now compared preview durations of 0 ms, 75 ms, and 250 ms.

Results

A summary of mean values can be seen in **Table 2**.

Response time. RT averaged 4734 ms across conditions and was significantly affected by the preview duration manipulation, $F(2,14) = 3.80$, in that RTs diminished with increasing preview duration [0 ms vs. 75 ms: $t(14) = 3.67$, $p < 0.01$; 75 ms vs. 250 ms: $t(14) = 1.63$, $p = 0.06$; 0 ms vs. 250 ms: $t(14) = 3.72$, $p < 0.01$].

Latency to first target fixation. Latency averaged 3936 ms with a main effect of preview duration, $F(2,14) = 7.81$, $p < 0.01$, in that the latency to first target fixation gradually diminished from 0 ms over 75 ms to 250 ms [0 ms vs. 75 ms: $t(14) = 2.03$, $p < 0.05$; 75 ms vs. 250 ms: $t(14) = 1.88$, $p < 0.05$; 0 ms vs. 250 ms: $t(14) = 4.04$, $p < 0.01$].

Number of fixations to first target fixation. On average, participants performed 13.76 fixations to the first fixation of the target object. There was an effect of preview duration, $F(2,14) = 9.24$, $p < 0.01$. Again we observed a gradual decline in the number of fixations to first target fixation from 0 ms over 75 ms to 250 ms [0 ms vs. 75 ms: $t(14) = 2.14$, $p < 0.05$; 75 ms vs. 250 ms: $t(14) = 1.99$, $p < 0.05$; 0 ms vs. 250 ms: $t(14) = 4.66$, $p < 0.01$].

Initial saccade latency. Initial saccade latency averaged 279 ms across conditions. There was a tendency for a decrease in initial saccade latency with longer preview durations, $F(2,14) = 2.73$, $p = 0.08$.

Initial saccade amplitude. Initial saccade amplitude averaged 2.34° visual angle across conditions. There was an effect of preview duration, $F(2,14) = 5.87$, $p < 0.01$, in that both the 75-ms and 250-ms preview durations were followed by significantly longer initial saccade amplitudes compared to the control condition, $t(14) = 2.57$, $p < 0.05$.

	Preview durations		
	0 ms	75 ms	250 ms
Response time in ms	5286 [255]	4809 [256]	4107 [392]
Latency to first target fixation in ms	4658 [247]	3923 [244]	3228 [366]
Number of fixations until target fixation	16.35 [0.86]	13.68 [1.19]	11.24 [0.89]
Initial saccade latency in ms	293 [11]	277 [8]	267 [14]
Initial saccade amplitude in degree visual angle	2.04 [0.15]	2.33 [0.15]	2.65 [0.29]

Table 2. Summary of mean values [standard errors] of [Experiment 1b](#) for the dependent variables (response time, latency to first target fixation, number of fixations to target fixation, initial saccade latency, and initial saccade amplitude) as a function of preview duration (0 ms vs. 75 ms vs. 250 ms).

and $t(14) = 2.97$, $p < 0.01$, respectively, while the 75 ms showed a tendency for shorter initial saccade amplitudes compared to the 250-ms preview, $t(14) = 1.61$, $p = 0.06$.

Summary

In [Experiment 1b](#), RT as well as eye movement data showed that a preview duration of 75 ms was sufficient to produce scene preview benefits compared to the control condition. At the same time, a 75-ms glimpse of the preview did not reach the degree of preview benefit that a 250-ms glimpse did. Thus, it seems that some information can be acquired in 75 ms, while additional information continues to accrue after that time. This follows the general pattern already observed in [Experiment 1a](#) according to which search benefits decrease for decreasing preview durations.

Experiment 1c

Since 75-ms presentations of scene previews resulted in scene preview benefits, we further decreased preview durations to 50 ms in [Experiment 1c](#) to test whether search benefits would be observable for such short scene presentations.

Methods

Participants

Fifteen native English-speaking students (9 females) from the University of Edinburgh ranging in age between 19 and 31 ($M = 22.93$, $SD = 4.28$) participated in [Experiment 1a](#) for course credit or for £6/h. All participants reported normal or corrected-to-normal vision and none had taken part in [Experiment 1a](#) or [1b](#).

Procedure

The procedure was identical to the procedures of [Experiments 1a](#) and [1b](#). However, in [Experiment 1c](#) we compared preview durations of 0 ms, 50 ms, and 250 ms.

Results

A summary of mean values can be seen in [Table 3](#).

Response time. RT averaged 4808 ms across conditions and was significantly affected by the preview duration manipulation, $F(2,14) = 6.52$, in that RTs for the 250-ms preview durations was diminished compared to the control condition, $t(14) = 3.69$, $p < 0.01$, as well as to the 50-ms preview condition, $t(14) = 2.20$, $p < 0.05$. RTs for the 50-ms and control conditions did not differ from each other, $t(14) = 1.22$, $p > 0.05$.

	Preview durations		
	0 ms	50 ms	250 ms
Response time in ms	5429 [342]	4970 [351]	4025 [268]
Latency to first target fixation in ms	4427 [295]	4282 [273]	3213 [361]
Number of fixations until target fixation	16.51 [1.1]	15.56 [1.12]	11.41 [0.67]
Initial saccade latency in ms	324 [19]	280 [17]	273 [11]
Initial saccade amplitude in degree visual angle	2.07 [0.1]	2.38 [0.23]	2.69 [0.15]

Table 3. Summary of mean values [standard errors] of [Experiment 1c](#) for the dependent variables (response time, latency to first target fixation, number of fixations to target fixation, initial saccade latency, and initial saccade amplitude) as a function of preview duration (0 ms vs. 50 ms vs. 250 ms).

Latency to first target fixation. Latency averaged 3974 ms and was significantly affected by preview duration, $F(2,14) = 7.27$, $p < 0.01$, in that the latency following the 250-ms preview was significantly diminished compared to the control condition, $t(14) = 3.83$, $p < 0.01$. However, the 50-ms preview duration did not show significant scene preview benefits compared to the control, $t < 1$, while showing longer latencies than the 250-ms duration condition, $t(14) = 2.77$, $p < 0.01$.

Number of fixations to first target fixation. On average, participants performed 14.49 fixations to the first fixation of the target object. We observed an effect of preview duration, $F(2,14) = 7.98$, $p < 0.01$, in that a preview duration of 250 ms decreased the number of fixations to first target fixation compared to the control condition, $t(14) = 2.73.906$, $p < 0.01$, as well as compared to the 50-ms preview duration, $t(14) = 3.06$, $p < 0.01$, while the 50-ms and control conditions did not differ from each other, $t < 0.01$.

Initial saccade latency. Initial saccade latency averaged 292 ms across conditions. There was a significant effect of preview duration, $F(2,14) = 6.54$, $p < 0.01$, in that initial saccade latency was significantly diminished for both the 50-ms and 250-ms conditions compared to the control, $t(14) = 2.35$, $p < 0.05$ and $t(14) = 3.33$, $p < 0.01$, respectively. The 50-ms and 250-ms preview conditions did not differ in their effect on initial saccade latency, $t < 1$.

Initial saccade amplitude. Initial saccade amplitude averaged 2.38° visual angle across conditions. There was an effect of preview duration, $F(2,14) = 5.93$, $p < 0.01$. Initial saccade amplitudes were significantly longer for both the 50-ms and 250-ms preview durations compared to the control condition, $t(14) = 3.64$, $p < 0.01$ and $t(14) = 2.91$, $p < 0.01$, respectively, while the 50-ms and 250-ms preview durations did not differ from each other, $t(14) = 1.44$, $p > 0.05$.

Summary

Overall, the results of [Experiment 1c](#) showed that a preview duration of 50 ms did not lead to scene preview benefits compared to a no-preview control. A significant scene preview benefit was only evident in the 250-ms condition. In the one exception, initial eye movements were modulated by preview presentation times of 50 ms: the amplitude of the first saccade was marginally increased and initial saccade latency was significantly diminished following a 50-ms preview compared to the control condition.

Discussion

The first set of experiments aimed at investigating the minimum scene presentation time needed to allow for subsequent facilitation of eye movement guidance in

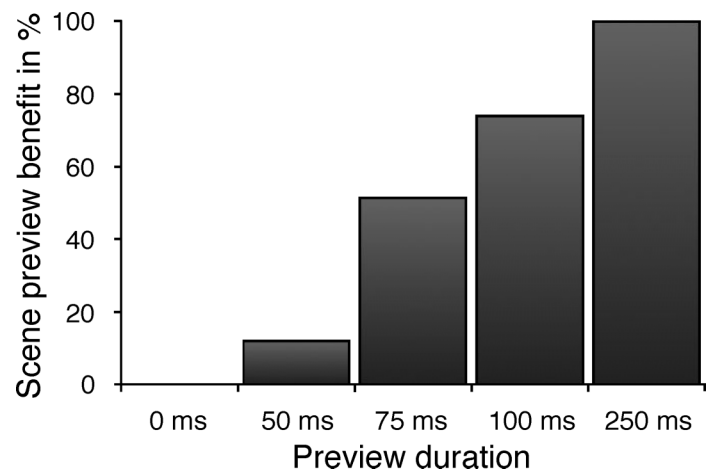


Figure 2. Overview of scene preview benefits for the latency to first target fixation, here plotted as a percentage of baseline condition in each experiment, across [Experiments 1a–1c](#). RTs and number of fixations showed the same pattern of effects.

scene search. We varied preview durations from 100 ms ([Experiment 1a](#)) to 75 ms ([Experiment 1b](#)) to 50 ms ([Experiment 1c](#)) across the three experiments. For response times, latencies, and number of fixations to first target fixations, we observed a decrease in scene preview benefit with decreasing preview duration (see [Figure 2](#) for a visualization of this relationship for the latency measure).

Initial eye movements also showed a general pattern of preview duration effects in that initial saccades seemed to be executed faster and further into the scene with increasing preview duration. Interestingly, while a 50-ms preview did not lead to overall search benefits, initial eye movements were nevertheless modulated by the short presentation times. This might be due to the fact that these initial eye movement measures—which include all initiating saccades—provide global information on the general readiness of the participants to initiate search rather than their ability to direct their first saccade toward the target. Thus, from these findings we can surmise that scene presentations as short as 50 ms might provide enough information to trigger initial saccades, but longer presentation times might be necessary for search benefits to arise.

Experiment 2

[Experiments 1a–1c](#) provided evidence that information useful for planning search-related eye movements can be acquired from a scene in as little as 75 ms. However, scene processing does not cease once the visual input has disappeared. In order to achieve scene preview benefits

from the scene preview, information about the search target also has to be processed in relation to the scene input to activate knowledge and expectations regarding probable target locations within the scene (Torralba, Oliva, Castelhana, & Henderson, 2006). We term this phase of processing scene-target integration time, to distinguish it from scene consolidation time alone. An increase in the time available to set up scene priors on the basis of combined visual scene and abstract target information might therefore further increase scene preview benefit. In [Experiment 2](#), we tested this hypothesis by inserting additional integration time after the presentation of the target word and before the start of search.

Methods

Participants

Sixteen native English-speaking students (9 females) from the University of Edinburgh ranging in age between 18 and 30 ($M = 20.25$, $SD = 3.21$) participated in [Experiment 2](#) for £6/h. All participants reported normal or corrected-to-normal vision and none had taken part in any of the prior experiments.

Procedure

The procedure of [Experiment 2](#) closely followed the procedures of [Experiments 1a–1c](#). However, instead of varying the preview duration, we kept preview presentation time constant at 250 ms, while varying the time before search scene onset between 500 ms and 3000 ms. This time was filled with a fixation cross. The trial sequence started with a fixation check. When the fixation check was deemed successful, the fixation cross was replaced by the presentation of the scene's preview or a gray screen for 250 ms. After the presentation of a mask for 50 ms, a black target word was displayed at the center of the gray screen for 500 ms, which indicated the identity of the target object. Before the search scene was displayed, a fixation cross was presented for either 500 ms or 3000 ms. Participants then again viewed the search scene through a gaze-contingent window.

Results

Data were submitted to a 2×2 ANOVA with preview (scene vs. control) and delay (500 ms vs. 3000 ms) as within-subject factors. Significant interactions were followed up by planned contrasts. A summary of means can be seen in [Table 4](#).

Response time. RT averaged 4381 ms across conditions. RT was significantly affected by the preview manipulation in that RTs following a scene preview were shorter than following a gray control screen, $F(1,15) = 21.40$, $p < 0.01$. There was no main effect of delay, $F < 1$. However, there was a significant interaction of preview and delay, $F(1,15) = 5.64$, $p < 0.05$, characterized by a greater preview effect for the longer delay. While the scene preview benefit—measured as the difference in search performance between preview presentation and control—reached 568 ms with a 500-ms delay, $t(15) = 2.25$, $p < 0.05$, the preview effect was more pronounced for the 3000-ms delay reaching 1231 ms, $t(15) = 5.44$, $p < 0.01$.

Latency to first target fixation. Latency averaged 3495 ms across all conditions. Similar to RT, there was a strong effect of preview, $F(1,15) = 17.94$, $p < 0.01$, but no effect of delay, $F < 1$. The significant interaction, $F(1,15) = 5.63$, $p < 0.05$, was also characterized by a greater degree of preview benefit for longer delays. The scene preview benefit amounted to 524 ms with a 500-ms delay and failed to reach significance, $t(15) = 1.77$, $p > 0.05$, while the preview effect was more pronounced for the 3000-ms delay amounting to 1230 ms, $t(15) = 5.94$, $p < 0.01$.

Number of fixations to first target fixation. On average, participants performed 12.06 fixations to the first fixation of the target object. While a scene preview significantly decreased the number of fixations, $F(1,15) = 19.22$, $p < 0.01$, the delay manipulation showed no significant main effect, $F < 1$. However, preview and delay significantly interacted, $F(1,15) = 5.81$, $p < 0.05$. While the preview effect (1.76 fixations) did not reach significance with a 500-ms delay, $t(15) = 1.94$, $p > 0.05$, the preview effect was significant for the 3000-ms delay reaching 4.26 fixations, $t(15) = 5.94$, $p < 0.01$.

Initial saccade latency. Initial saccade latency averaged 314 ms across conditions. There was a significant main

	Scene		Control	
	500 ms	3000 ms	500 ms	3000 ms
Response time in ms	4026 [278]	3838 [239]	4594 [260]	5069 [238]
Latency to first target fixation in ms	3208 [300]	2904 [248]	3732 [233]	4134 [230]
Number of fixations until target fixation	11 [0.92]	10.1 [0.77]	12.76 [0.64]	14.36 [0.79]
Initial saccade latency in ms	294 [12]	302 [12]	332 [17]	326 [15]
Initial saccade amplitude in degree visual angle	2.71 [0.15]	2.73 [0.22]	2.02 [0.11]	1.89 [0.10]

Table 4. Summary of mean values [standard errors] of [Experiment 2](#) for the dependent variables (response time, latency to first target fixation, number of fixations to target fixation, initial saccade latency, and initial saccade amplitude) as a function of preview (scene vs. control) and delay (500 ms vs. 3000 ms).

effect of preview, $F(1,15) = 9.72$, $p < 0.01$, with less initial saccade latency following a preview than a control, while both delay and its interaction with preview failed to reach significance, both $F_s < 1$.

Initial saccade amplitude. Initial saccade amplitude averaged 2.33° visual angle across conditions. There was again a significant main effect of preview, $F(1,15) = 24.14$, $p < 0.01$, with greater saccade amplitudes following a preview than a control, while both delay and its interaction with preview failed to reach significance, both $F_s < 1$.

Summary

A 250-ms preview resulted in scene preview benefits across all measures. Most importantly, prolonging the delay between target word offset and search scene onset influenced search performance by increasing the preview benefit as seen in RTs, latency, and number of fixations until target fixation. Similar to the first set of experiments, initial eye movements were modulated by presenting participants with a scene preview compared to no scene preview, but integration time did not affect initial eye movements.

Discussion

[Experiment 2](#) showed that inserting a longer delay prior to initiation of object search further increased scene preview benefits. Thus, scene preview benefit following a 250-ms preview presentation was enhanced when participants were given more time to combine the visual information from the preview with information elicited by the target word.

Experiment 3

The findings so far raise the question whether a reduction in preview duration can be further offset by providing additional target-scene integration time in generating scene preview benefits. In [Experiment 3](#), we therefore tested whether a 50-ms preview—which had previously not led to significant scene preview benefits—would provide sufficient scene information to support efficient eye movement guidance when paired with a 3000-ms delay for target-scene integration before the initiation of search.

Methods

Participants

Sixteen native English-speaking students (9 females) from the University of Edinburgh ranging in age between

18 and 27 ($M = 21.63$, $SD = 2.87$) participated in [Experiment 3](#) for £6/h. All participants reported normal or corrected-to-normal vision and none had taken part in any of the prior experiments.

Procedure

The procedure of [Experiment 3](#) was identical to the procedure of [Experiment 2](#) apart from the preview duration that was reduced from 250 ms to 50 ms.

Results

Response time. RT averaged 4446 ms across conditions. RT was significantly affected by the preview manipulation in that RTs following a scene preview were shorter than following a gray control screen, $F(1,15) = 11.95$, $p < 0.01$. There was no main effect of delay, $F(1,15) = 1.11$, $p > 0.05$. However, there was a significant interaction of preview and delay, $F(1,15) = 6.23$, $p < 0.05$, characterized by a greater scene preview benefit for the longer delay. While the scene preview benefit amounted to 246 ms for a delay of 500 ms without reaching significance, $t < 1$, participants profited significantly more from a 50-ms preview when combined with a 3000-ms delay, $t(15) = 4.33$, $p < 0.01$, with a benefit amounting to 1249 ms.

Latency to first target fixation. Latency averaged 3506 ms across all conditions. Again, there was a strong effect of preview, $F(1,15) = 12.25$, $p < 0.01$, but no main effect of delay, $F < 1$. However, delay significantly interacted with preview, $F(1,15) = 5.82$, $p < 0.05$. Scene preview benefit was not significant for the 500-ms delay condition with a benefit of only 372 ms, $t(15) 1.23$, $p > 0.05$, but there was a significant scene preview benefit for the 50-ms preview condition compared to the control for a delay of 3000 ms, $t(15) = 4.31$, $p < 0.01$, which amounted to 1267 ms.

Number of fixations to first target fixation. On average, participants performed 12.65 fixations to the first fixation of the target object. The main effect of scene preview was significant, $F(1,15) = 12.98$, $p < 0.01$, but the delay manipulation showed no significant main effect, $F(1,15) < 1$. The significant interaction of preview and delay, $F(1,15) = 4.72$, $p < 0.05$, was characterized by a lack of scene preview benefit for the 500-ms delay that measured 1.45 fixations, $t(15) = 1.4$, $p > 0.05$, while participants significantly benefited from the 50-ms preview in the 3000-ms delay condition, $t(15) = 3.98$, $p < 0.01$, with a benefit of 4.41 fixations.

Initial saccade latency. Initial saccade latency averaged 309 ms across conditions. There was a significant main effect of delay, $F(1,15) = 5.55$, $p < 0.05$, with greater initial saccade latencies following a 3000-ms delay compared to a 500-ms delay, while both preview and its interaction with delay failed to reach significance, $F(1,15) = 2.26$, $p > 0.05$ and $F < 1$, respectively.

	Scene		Control	
	500 ms	3000 ms	500 ms	3000 ms
Response time in ms	4446 [245]	3698 [172]	4692 [275]	4947 [243]
Latency to first target fixation in ms	3434 [265]	2758 [186]	3806 [257]	4025 [243]
Number of fixations until target fixation	12.08 [0.76]	10.29 [0.67]	13.53 [0.81]	14.70 [0.92]
Initial saccade latency in ms	286 [13]	317 [20]	305 [16]	327 [19]
Initial saccade amplitude in degree visual angle	2.44 [0.18]	2.35 [0.10]	2.25 [0.12]	2.27 [0.12]

Table 5. Summary of mean values [standard errors] of [Experiment 3](#) for the dependent variables (response time, latency to first target fixation, number of fixations to target fixation, initial saccade latency, and initial saccade amplitude) as a function of preview (scene vs. control) and delay (500 ms vs. 3000 ms).

Initial saccade amplitude. Initial saccade amplitude averaged 2.33° visual angle across conditions. There were no main effects for either preview or delay, $F(1,15) = 1.75$, $p > 0.05$ and $F < 1$, respectively, and also no interaction, $F < 1$.

A summary of mean values can be seen in [Table 5](#).

Summary

A 50-ms preview led to search benefits as seen in RT, latency, and number of fixations to target fixation. These scene preview benefits were modulated by the delay manipulation in that a 50-ms preview failed to produce scene preview benefits with a short delay but did benefit search with a delay of 3000 ms. Initial eye movements were not modulated by the short preview except for initial saccade latency, which increased with increasing integration time.

Discussion

[Experiment 3](#) showed that combining a 50-ms preview—which had failed to provide significant scene preview benefits in [Experiment 1c](#)—with a delay of 3000 ms led to more pronounced scene preview benefits. Thus, providing participants with more time to integrate the initial scene representation with scene knowledge facilitated eye movement guidance when searching naturalistic scenes (see [Figures 3a](#) and [3b](#)).

These results suggest that search guidance can be enhanced not only by prolonging the availability of the visual image of a scene but also by providing more time to integrate the information extracted from an initial glimpse of a scene with knowledge about how that scene information is to be used.

General discussion

The aim of the present study was to investigate the time course of information accumulation from the initial glimpse of a scene until the initiation of active object search via eye movements. To investigate this issue, we used the flash-preview moving-window paradigm in which a scene is briefly previewed prior to visual search

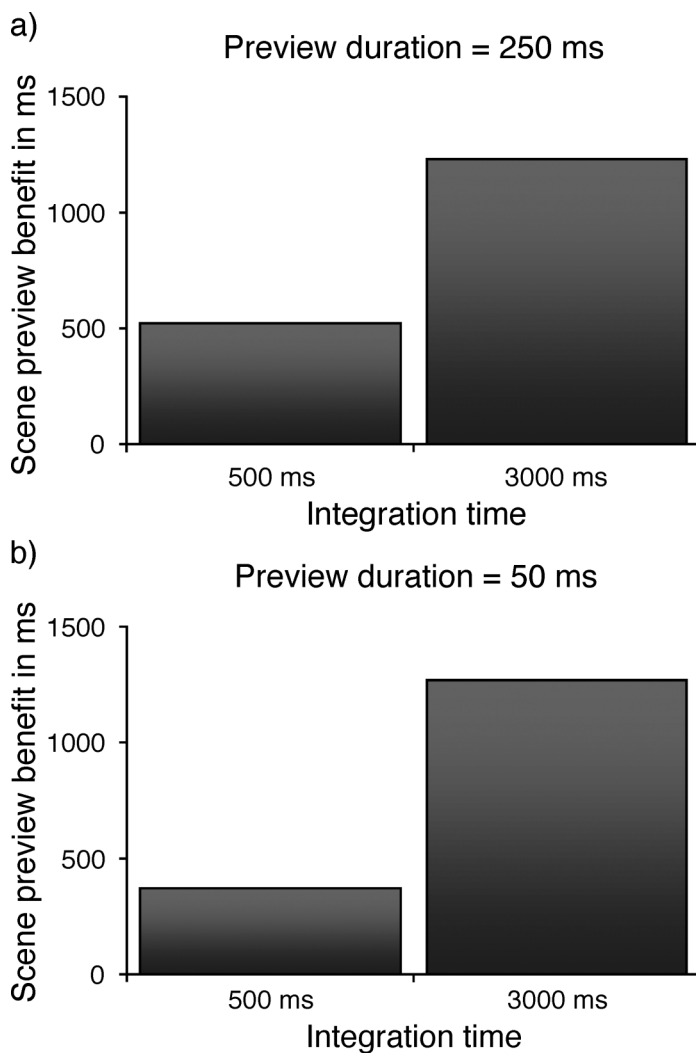


Figure 3. Scene preview benefits as a function of short (500 ms) vs. long (3000 ms) integration times observed for preview durations of (a) 250 ms in [Experiment 2](#) and (b) 50 ms in [Experiment 3](#).

for a target object (Castelhamo & Henderson, 2007). The search then takes place through a moving window tied to the current fixation location that reveals only a small area of the scene. Experiments 1a–1c showed that a scene preview duration of 75 ms—and therefore within the time range normally attributed to gist identification (e.g., Biederman et al., 1982; Castelhamo & Henderson, 2008; Greene & Oliva, 2009; Potter, 1975; Rousselet et al., 2005; Schyns & Oliva, 1994; Thorpe et al., 1996)—provided sufficient information to subsequently benefit search. Experiment 2 showed that inserting additional integration time before the initiation of search led to increased scene preview benefits when combined with a preview duration of 250 ms. Finally, Experiment 3 demonstrated that the presentation of a scene preview for as little as 50 ms produced scene preview benefits when combined with longer integration time. Following the brief preview and by the time the search scene appeared, the computation of highly probable target locations—on the basis of the visual percept of the scene as well as activated structural and semantic knowledge-facilitated eye movement guidance during subsequent object search.

While RT as well as the latency and number of fixations to the search target were consistently modulated by both scene presentation duration and scene-target integration time, the effects on initial saccade latency and amplitude seemed less sensitive to these manipulations. This might be due to the relatively small amount of variability of these eye movement measures compared to the other measures. Nevertheless, we consistently observed prolonged initial saccade amplitudes and decreased saccade latencies for conditions in which a scene preview was visible as compared to a control. Thus, these measures might indicate a general readiness to initiate search by means of quickly executed and long saccades into the search display.

The ability of viewers to extract information useful for planning search-related eye movements from a scene in as little as 50 ms adds to the findings of Rayner et al. (2009), who reported that participants needed to see a scene for at least 150 ms during each fixation to allow for normal scene processing. However, a reason for the different minimum time estimates in Rayner et al.'s study and the present study lies in the paradigms used. While the flash-preview moving-window paradigm used here manipulated only the duration of the initial glimpse of the scene, the mask onset delay paradigm used by Rayner et al. limits scene viewing to a brief amount of time during each fixation. Thus, the findings of Rayner et al. relate to the time course of ongoing rather than initial scene processing. An interesting set of open questions concerns the relationship between the amount of visual information acquired from an initial scene glimpse and the amount of visual information needed during subsequent fixations.

Observing search benefits with a minimum preview duration of 50 ms also appears inconsistent with findings

of Underwood and Green (2003), who used a sentence verification paradigm and found that a 750-ms scene preview did not reliably facilitate performance relative to a no-preview condition. The authors concluded that early knowledge of the gist of a scene might provide little help in the process of goal-directed search. However, the authors themselves raised the possibility that when an extended view of a scene is available, extraction of gist might not be a priority, reducing the impact of a 750-ms preview. The flash-preview moving-window paradigm used in our study probably encouraged participants to extract as much information from the preview as possible since viewing during search was restricted. In addition, the limited access to visual input during search might have motivated participants to further elaborate the initial scene representation once the preview had disappeared, i.e., in the additional delay inserted before initiation of search. It remains to be determined whether a brief scene preview prior to search in the absence of a moving window would facilitate search.

The findings of this study are in line with the contextual guidance model (Torralba et al., 2006), which proposes that bottom-up saliency as well as top-down mechanisms based on scene context interact during initial scene processing. In addition to computing a saliency map on the basis of local scene processing, global processing of a scene's context enables search to be guided by previous experience and expectations about where certain objects are most commonly found in scenes. Thus, eye movements can be successfully guided to the target object when the scene gist and spatial structure extracted from a briefly flashed scene are integrated with target knowledge. Our data clearly show that both the duration a scene is visible for inspection (i.e., scene presentation time) and the time available after the search target has been specified (i.e., scene-target integration time) modulate eye movement control in search. We therefore propose that both scene presentation duration and integration time are crucial for generating contextual guidance for object search in complex, naturalistic scenes.

A similar distinction between an initial scene representation initially established on the basis of visual features on the one hand and the influence of knowledge structures on the other has been made in the cognitive relevance framework proposed by Henderson, Malcolm, and Schandl (2009). Here objects are prioritized for fixation primarily on the basis of cognitive knowledge structures interacting with task goals rather than the objects' bottom-up saliency. However, the image plays an important role in two ways: First, the scene image is necessary to establish the initial scene representation. Second, on the basis of these initial scene representations cognitive knowledge structures can be activated. Our data imply that the presentation duration of a scene determines the first, whereas integration time supports the activation and application of knowledge structures to the initially established scene representation.

To date, the role of integrating initial scene information to promote subsequent contextual guidance has received little attention. The results of the present study confirm the speed at which the visual-cognitive system is able to extract useful scene information. The results further imply that by providing sufficient time to apply top-down scene knowledge to an initially crude scene representation, expectations regarding the scene's composition can be rapidly activated, facilitating guidance of eye movements. This speaks to the great flexibility of initial scene representations, which are continuously modulated by incoming visual input as well as by a viewer's prior experience and the current agenda. The integration of these different sorts of information is essential when maximizing the use of briefly presented visual scene input for the benefit of subsequent eye guidance.

Conclusions

Scene representations generated from a brief scene glimpse can provide sufficient information to guide subsequent behavior. We have found that a scene presentation time of 50 ms is sufficient to establish scene representations based on which eye movements can be efficiently controlled to benefit object search. However, the briefest scene durations are only useful when sufficient time for target-scene integration is available.

Acknowledgments

This project was supported by Grant RES-00-22-2721 from the Economic and Social Research Council of the UK to JMH.

Commercial relationships: none.

Corresponding author: Melissa L.-H. Võ.

Email: mlvo@search.bwh.harvard.edu.

Address: 64 Sidney Street, Suite 170, Cambridge, MA 02139, USA.

References

- Bacon-Macé, N., Macé, M., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: Backward masking and natural scene categorisation. *Vision Research*, *45*, 1459–1469. [PubMed]
- Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organisation* (pp. 213–253). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Biederman, I., Mezzanote, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177. [PubMed]
- Castelhano, M., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 753–763. [PubMed]
- Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on the activation of scene gist. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 660–675. [PubMed]
- Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision*, *7*(1):10, 1–29, <http://journalofvision.org/7/1/10/>, doi:10.1167/7.1.10. [PubMed] [Article]
- Friedman, A. (1979). Framing pictures: The roles of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355. [PubMed]
- Greene, M. R., & Oliva, A. (2009). The briefest of glances. The time course of natural scene understanding. *Psychological Science*, *40*, 464–472. [PubMed] [Article]
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504. [PubMed]
- Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review*, *16*, 850–856. [PubMed]
- Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 396–411. [PubMed]
- 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.
- Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 1–12. [PubMed]
- Joubert, O., Rousselet, G., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*, 3286–3297. [PubMed]
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*, 1762–1776. [PubMed]

- Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, 9(11):8, 1–13, <http://journalofvision.org/9/11/8/>, doi:10.1167/9.11.8. [PubMed] [Article]
- Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *The encyclopedia of neurobiology of attention* (pp. 251–256). San Diego, CA: Elsevier.
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, 155, 23–36. [PubMed]
- Potter, M. C. (1975). Meaning in visual scenes. *Science*, 187, 965–966.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 509–522. [PubMed]
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. [PubMed] [Article]
- Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2009). Eye movements and visual encoding during scene perception. *Psychological Science*, 20, 6–10. [PubMed] [Article]
- Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the “gist” of real-world natural scenes? *Visual Cognition*, 12, 852–877.
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5, 195–200.
- Tatler, B. W., Gilchrist, I. D., & Rusted, J. (2003). The time course of abstract visual representation. *Perception*, 32, 579–592. [PubMed]
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522. [PubMed]
- 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. [PubMed]
- Underwood, G., & Green, A. (2003). Processing the gist of natural scenes: Sentence verification with intrinsic and extrinsic configurations of objects. *Cognitive Processing*, 4, 119–136.
- Van Rullen, R., & Thorpe, S. (2001). The time course of visual processing: From early perception to decision making. *Journal of Cognitive Neuroscience*, 13, 454–461. [PubMed]
- Võ, M. L.-H., & Schneider, W. X. (2010). A glimpse is not a glimpse: Differential processing of flashed scene previews leads to differential target search benefits. *Visual Cognition*, 18, 171–200.