The effect of the first glimpse at a scene on eye movements during search

Anne P. Hillstrom · Helen Scholey ·
Simon P. Livversedge · Valerie Benson

Published online: 19 January 2012
© Psychonomic Society, Inc. 2012

Abstract Previewing scenes briefly makes finding target objects more efficient when viewing is through a gaze-contingent window (windowed viewing). In contrast, showing a preview of a randomly arranged search display does not benefit search efficiency when viewing during search is of the full display. Here, we tested whether a scene preview is beneficial when the scene is fully visible during search. Scene previews, when presented, were 250 ms in duration. During search, the scene was either fully visible or windowed. A preview always provided an advantage, in terms of decreasing the time to initially fixate and respond to targets and in terms of the total number of fixations. In windowed visibility, a preview reduced the distance of fixations from the target position until at least the fourth fixation. In full visibility, previewing reduced the distance of the second fixation but not of later fixations. The gist information derived from the initial glimpse of a scene allowed for placement of the first one or two fixations at information-rich locations, but when nonfoveal information was available, subsequent eye movements were only guided by online information.

Keywords Eye movements · Visual search · Scene gist · Gaze control · Scene perception

When searching through scenes, one encounters more guidance-relevant information than when searching through random arrays of objects, which results in behavior differences. For instance, a preview of a random search array before one knows the identity of the search target has been found to provide no benefit to search efficiency (Becker & Pashler, 2005; Wolfe, Klempen, & Dahlen, 2000). In contrast, a preview of a real scene does provide a search benefit (Hollingworth, 2009). Hollingworth found that a preview lasting from 500 ms to 10 s benefitted search because of both memory for the general scene layout and memory for the specific location of the target item.

Research about scene-preview benefits has often focussed on the benefit of brief previews, as this informs about what information can be extracted about the scene very rapidly, and on how useful that information is. The “gist” of a scene—a coarse understanding of its spatial and conceptual layout—can be determined with even a very brief exposure to the scene. After only 50 ms, people can categorize the kind of scene that was presented (Biederman, Rabinowitz, Glass, & Stacy, 1974; Greene & Oliva, 2009; Rousselet, Joubert, & Fabre-Thorpe, 2005; Schyns & Oliva, 1994) and perform better than chance on recognition memory for the scene (Potter & Levy, 1969). This ability no doubt contributes to efficient visual processing of new environments: If we turn a blind corner and see a new scene, we immediately gather enough visual information to assess whether to continue our progress without stopping and scanning the environment. Of course, we do not typically close our eyes after our first glimpse at a scene, so when studying natural behavior, it is difficult to distinguish the impact of the first glimpse from the impact of the inflow of information about the scene as we continue looking at it.

The term “gist” has been used to refer to both the conceptually driven (Underwood, Crundall, & Hodson, 2005)

A. P. Hillstrom (✉) · H. Scholey
Department of Psychology, University of Portsmouth,
King Henry Building, King Henry I St.,
Portsmouth PO1 2DY, UK
e-mail: anne.hillstrom@port.ac.uk

S. P. Livversedge · V. Benson
Centre for Visual Cognition, School of Psychology,
University of Southampton,
Southampton, UK

and perceptually driven global visual properties of the scene (Greene & Oliva, 2009; Torralba, Oliva, Castelhano & Henderson, 2006). Global visual properties include the spatial layout information and the statistical regularities of the visual properties. Conceptual information may derive from information about quickly recognized objects (e.g., Biederman et al., 1974), but the fact that sometimes scenes can be categorized faster than the objects in them suggests that the development of gist information need not require full identification of any object (Greene & Oliva 2009; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Oliva & Torralba, 2006).

According to both the visual and conceptual definitions, recognition of gist allows the observer to draw on learned knowledge about where objects might belong in scenes. For instance, when searching for a knife, knowledge about kitchens suggests that the object is likely to be on a worktop and unlikely to be on the floor. In addition, on the basis of our previous visual experience with kitchens, Bayesian priors indicate that knives would likely appear more toward the middle of the room than at the bottom of the room. Castelhano and Henderson (2007) studied how scene gist affected target localization by presenting a 250-ms preview of the scene, followed by the target name and then by the scene to be searched. During search, displaying the scene via a small gaze-contingent window (2° of visual angle in diameter) removed the influence of peripheral visual information available during normal search. Thus, search was guided only by the information acquired from the preview. Results from the study showed that with a scene preview, search was faster, there were fewer fixations, and the summed length of all saccades was shorter. This was true when compared to conditions in which the search display was preceded by (1) a scene differing in identity and layout, (2) a scene with the same identity but a different layout, and (3) a scrambled picture containing parts of many scenes. In some of the experiments, the preview of the scene contained the target; in other experiments, the preview did not show the target. Removing the target from the preview reduced but did not eliminate the benefit of the preview. Finally, when the preview was smaller in scale than the search display, there was still a preview benefit, suggesting that the representation of the scene that the initial glimpse delivers is not metric. Together, these experiments indicated that the first glimpse of a scene allows the observer to learn the spatial layout of the scene and to use it to plan the oculomotor behavior of search. Võ and Henderson (2010) demonstrated qualitatively similar benefits of a preview on windowed search when the preview was shortened to as little as 75 ms.

Another study used the same paradigm, with previews that were identical to the search display (but without showing the target), that contained scene structure without objects, that contained objects without scene structure, or that were scrambled mosaics drawn from many scenes (Võ & Schneider,

2010). Võ and Schneider found that previewing only objects did not benefit search; that participants with relatively slow visual processing benefitted most from the identical previews; and that participants with relatively fast visual processing benefitted most from a preview of only the structure of the scene. Fast visual processing was measured both by whether the observer noticed and could describe the different kinds of previews used in the search trials and by obtaining an estimate of visual processing speed by finding the best-fit version of the theory-of-visual-attention model to a simple perceptual task for each participant (Bundesen, Habekost, & Kyllingsbaek, 2005). The researchers concluded that the fast processors drew local as well as global information from the preview and that the local information interfered with efficient processing of the global information in identifying likely target locations. They speculated that the local information acquired from the preview might have widened the range of previously encountered scenes that contributed to the Bayesian probabilities of where a target might be. On the other hand, slow processors drew only global information from the preview and benefitted from the structural information available in the global information, regardless of whether some local information changed.

These experiments point to universally fast processing of the global structure of a scene that enables better planning of eye movements as the scene is searched. However, it is important to note that the flash-preview windowed-search paradigm imposes a high demand on memory for the scene in order for search to be efficient. Although the results for this kind of search are important, everyday search through scenes is not limited in terms of how much of the scene is visible while the eyes move. Since peripherally available information influences saccade planning (e.g., Loschky & McConkie, 2002; van Diepen, De Graef, & d'Ydewalle, 1995; van Diepen & d'Ydewalle, 2003), it is important to determine whether the enduring influence of the first glimpse of the scene in windowed search generalizes to search in which observers can also be guided by peripheral visible information. In tasks in which goals develop over time, there is evidence that, rather than using working memory to retain as full an understanding of the visual environment as possible during an entire task, people use eye movements to gather “just in time” the information required for the next step in the task (Ballard, Hayhoe, & Pelz, 1995). The results of Becker and Pashler (2005) and Wolfe et al. (2000) are consistent with this understanding of the use of vision versus memory in visual search. If this just-in-time strategy applies to search through scenes, then there may be reason to believe that the planning of eye movements during search is not as strongly influenced by scene gist as the results of studies using windowed search might imply.

With this in mind, it is worth considering whether the preview benefit found by Hollingworth (2009) generalizes windowed-search results to search through fully visible

displays. However, the number of differences between the studies makes it difficult to compare Hollingworth's study to Castelhana and Henderson (2007) directly. Most importantly, the longer previews used by Hollingworth allowed the observer to make multiple eye movements during the preview, and thus to establish a firmer representation of the entire scene than a single fixation allows. Thus, it can be argued that Hollingworth's effects might not have been due to gist processing, but instead to more fully developed memory for a scene. Second, Hollingworth showed the target object itself, rather than its name, to indicate the goal of search. This modality change could have affected the use of the preview, in that the representation of the target might have competed with the representation of the scene in visual working memory. Third, the scenes he used were schematic rather than photographs, which might have had subtle effects on the processing choices. Finally, Hollingworth did not explore whether the preview provided a head start or influenced saccade planning throughout search. For these reasons, it was important to run a study in which observers searched both windowed and full-view displays, with and without a preview and with all other aspects consistent, in order to make a clear comparison of the effects of the information extracted from the first glance on subsequent eye movement sampling strategies under the two presentation conditions.

The goal of this experiment, then, was to explore whether eye movements when searching scenes are influenced by the gist extracted in the first fixation on the scene. The flash-preview windowed-search paradigm was broken down into a manipulation of two factors: scene *visibility* (whether the scene was fully visible during search or visible only through a peephole) and scene *preview* (whether or not a preview of the scene was presented before the target name was given to the participants). If planning of eye movements is strongly influenced by the scene gist extracted from preview, then one should see a similar influence of preview for an extended time, regardless of scene visibility. However, if people rely more on information currently being seen than on the gist of the scene remembered from the first fixation, then the effect of the preview should be weak, or even nonexistent, when the scene is fully visible during search.

Method

Design

Scene preview (present vs. absent) and scene visibility (fully visible vs. gaze-contingent window) were manipulated within subjects. Each scene was seen once by each participant. The assignment of scenes to experimental conditions was counterbalanced across participants.

Participants

A group of 20 participants (3 male, 17 female) participated in the study either in partial fulfilment of course credit or for £5. Their mean age was 31.35 years (range: 19–57). All reported having normal or corrected-to-normal vision.

Stimuli and equipment

A computer controlled by the Experiment Builder software (SR Research, Ltd, Osgoode Canada) presented stimuli on a 19-in. ViewSonic G90b CRT monitor running at a 120-Hz screen refresh rate. A second, linked computer controlled the eye-tracker. Manual responses were made on a game pad. An EyeLink 1000 eyetracker (SR Research, Ltd, Osgoode Canada) running at a 500-Hz rate tracked eye movements. A chinrest stabilized the eyes 50 cm from the display.

A total of 99 scenes (internal and external) were taken from a variety of sources; 32 had been used in the original study by Castelhana and Henderson (2007). The primary criterion for choosing the scenes was the presence of an object in the image that could serve as an unambiguous target that people would recognize by name. The pictures, 800×600 pixels, were presented to fill the screen. Figure 1 shows two scenes that were used in the study; finding the target was relatively easy in Fig. 1a and relatively difficult in Fig. 1b.

The 800×600 pixel postpreview mask was composed of 50×50 pixel regions clipped from the original images and arranged randomly in an approximate grid, with some overlap between regions.

The target name was presented as a one- to four-word description centered on the monitor in a large, clearly visible font. Over 90% of the target names were one or two words in length. The background of the target name and of all message screens presented during the experiment was set at a medium grey (RGB = [117, 117, 117]) to minimize abrupt luminance changes.

The mean distance from the center of the display to the center of a target was 7.45° , ranging from 2.58° to 12.32° . Targets were on average 3.3° to a side, ranging from 0.9° to 7° . In all, 21 targets were in the top left quadrant of the scene, 24 were in the top right quadrant, 29 were in the bottom left quadrant, and 25 were in the bottom right quadrant.

Procedure

After the participant was introduced to the task, the eye-tracker was calibrated to less than 0.5° error. The participant then searched through all 99 scenes, ordered randomly. Calibration accuracy was checked after every search display,

Fig. 1 Two representative sample scenes. In panel **a**, the target was the bench. In panel **b**, the target was the flag. Search was easier with the scene in panel A than the scene in panel **b**

and recalibrations were carried out whenever necessary. Trials for all four viewing conditions (with and without preview, windowed vs. full view) were presented in a single block in random order.

Each trial sequence was as follows. First, a spot was presented at the center of the display until the participant looked at it steadily. Either a preview of the scene (identical to the searched scene, and so including the target object) or a grey screen was presented for 250 ms, followed by the mask for 50 ms. Then the name of the target was presented for 2,000 ms, followed by the scene until the participant pressed a button while looking at where he or she believed the target to be. During search, either the scene was fully visible or a circular region was visible, centered at the point of gaze with a diameter of 2.1° . The gaze-contingent window followed the point of gaze as the participant searched the display.

Results

Due to tracker loss, 4% of the data were excluded from the analyses.

Responses were considered accurate if the participant looked at the target when pressing the button or immediately before pressing the button. Participants responded accurately on 81.8% of trials (ranging from 61% to 97%). Participants' accuracy in the four conditions, averaged across scenes, was submitted to an ANOVA with Scene Visibility and Preview as factors. This analysis showed that responses were more accurate when searching through a fully visible scene (91.2%, $SD = 7.2$) than when searching through a window (72.2%, $SD = 14.3$), $F(1, 19) = 61.86$, $p < .001$, $\eta^2 = .77$. Neither preview nor the interaction between scene visibility and preview affected accuracy, $F_s < 1$ for both.

Response time and oculomotor behavior were analyzed only for trials that ended with correct responses. The period analyzed began with the (re)appearance of the scene after the target description was shown. Fixations shorter than 50 ms were excluded from the analyses. For each analysis of a measure, a trial was omitted if the measure for that trial was more than two standard deviations above or below the mean of that measure. Figure 2 shows average response times, average numbers of fixations, times until the first fixation on the target, average fixation durations, decision times (response time minus time of first target fixation), and first-saccade amplitudes in the four experimental conditions. Table 1 presents the means and standard deviations for the same measures.

An ANOVA was conducted on each dependent variable to look for effects of preview and scene visibility. To summarize the results, visibility affected all measures significantly, but the preview main effect and the interaction between preview and visibility were significant for some, but not all, of the measures. For all significant interactions, preview had a significant effect for both full visibility and windowed visibility, so the interactions reflected a difference in magnitude of the preview effect.

Details of the ANOVA results follow. With full scene visibility versus windowed viewing, responses to full scenes were considerably faster, $F(1, 19) = 235.89$, $p < .001$, $\eta^2 = .93$; there were fewer fixations, $F(1, 19) = 233.87$, $p < .001$, $\eta^2 = .93$; the target was fixated sooner, $F(1, 19) = 218.85$, $p < .001$, $\eta^2 = .92$; fixation durations were shorter, $F(1, 19) = 59.03$, $p < .001$, $\eta^2 = .76$; decisions to respond to seen targets were made more quickly, $F(1, 19) = 47.00$, $p < .001$, $\eta^2 = .71$; and the first saccade was longer in amplitude, $F(1, 19) = 244.04$, $p < .001$, $\eta^2 = .93$. With a preview, responses were faster, $F(1, 19) = 68.20$, $p < .001$, $\eta^2 = .78$; there were fewer fixations, $F(1, 19) = 55.24$, $p < .001$, $\eta^2 = .74$; and the time to first fixate the target was shorter, $F(1, 19) = 45.96$, $p < .001$, $\eta^2 = .71$, than without a preview. There

Fig. 2 Data for the four display conditions: (a) Response times. (b) Numbers of fixations. (c) Times until the target was first fixated. (d) Average fixation durations. (e) Decision times (time between the first fixation of the target and the response). (f) First-saccade amplitudes. Error bars represent one standard error of the mean

was no significant effect of preview on fixation durations, $F(1, 19) = 3.54$, $p = .075$; on decision times, $F(1, 19) = 2.68$, $p = .118$; or on the amplitudes of first saccades, $F(1, 19) = 1.70$, $p = .208$. There was a significant interaction between visibility and preview for response times, numbers of fixations, and latencies to first target fixations, $F(1, 19) = 40.10$, $p < .001$, $\eta^2 = .68$; $F(1, 19) = 31.44$, $p < .001$, $\eta^2 = .62$; and $F(1, 19) = 25.97$, $p < .001$, $\eta^2 = .58$, respectively. Follow-up t tests showed that information gathered from a preview speeded responses for both windowed viewing, $t(19) = 7.54$, $p < .001$, and full-scene viewing, $t(19) = 2.65$, $p = .016$; reduced the number of fixations for both windowed viewing, $t(19) = 6.72$, $p < .001$, and full-scene viewing, $t(19) = 3.40$, $p = .003$; and speeded the first fixation of the target for both windowed viewing, $t(19) = 6.04$, $p < .001$, and full-scene viewing, $t(19) = 4.59$, $p < .001$. For first-saccade amplitudes, decision times, and fixation durations, there were no significant interactions, $F(1, 19) = 2.81$, $p = .110$; $F(1, 19) = 3.19$, $p = .090$; and $F(1, 19) < 1$, respectively.

Shorter fixations during full-scene viewing than during windowed viewing are consistent with using parafoveal and peripheral vision to partially preprocess information at the next saccade target location. The lack of an effect of preview on average fixation durations suggests that the information from the preview was insufficient to facilitate object identification robustly across all fixations. The lack of an effect of preview on first-saccade amplitudes suggests that the information extracted from the preview—the scene gist—was insufficient to influence saccade targeting. To assess the impact of preview on the initial approach to the target, we considered whether the first few saccades were directed towards the target to a greater extent with than without a preview. Student's t tests were carried out to examine the effect of preview for each fixation in each of the Guidance $\times$ Visibility conditions. A Bonferroni correction for multiple tests established $p = .006$ as the criterion for significance. As can be observed in Fig. 3 and Table 1, for windowed visibility, fixations were closer to the target in the second, third, and fourth fixations with a preview than without a preview, $t(19) = 3.21, 4.33$, and 4.23 ; $ps = .005, < .001$, and $< .001$, respectively. The first fixation, however, was not closer, $t(19) = 2.01$, $p = .059$. For fully visible scenes, the second fixation was closer to the target with a preview than without, $t(19) = 4.23$, $p < .001$, but the first, third, and fourth fixations were unaffected by a preview, $t(19) = 2.49, 2.77$, and -0.35 ; $ps = .022, .012$, and $.728$, respectively. Thus, the influence of the information gleaned from the preview guided eye movements through the fourth fixation when no peripheral information was available, but the same guidance occurred to a far lesser extent when nonfoveal information was available.

Discussion

This experiment compared the influence of a brief scene preview, roughly the duration of a short fixation during scene perception (Rayner, 1998), on search behavior when the scene was fully visible as compared to when only a small part of the scene was visible at any one time during search. The goal was to determine whether gist derived from a preview of a scene influenced search efficiency and strategy equivalently in naturalistic viewing and in windowed viewing conditions. In both viewing conditions, previews led to fewer fixations, faster response times, and shorter times until the target was first fixated, with an effect of greater magnitude in the windowed condition than in the full-visibility condition. In neither viewing condition did a preview affect decision times, saccade durations, or first-saccade amplitudes. Where we used the same or similar measures, these results are consistent with what has been reported by Castelhamo and Henderson (2007), Hollingworth

Table 1 Performance measures in the four display conditions

Measure	Full Scene Visible				Windowed View			
	With Preview		Without Preview		With Preview		Without Preview	
	M	SD	M	SD	M	SD	M	SD
Response time (ms)	1,217.2	294.9	1,364.4	272.3	3,812.8	1,225.5	5,299.6	1,175.5
Number of fixations	5.2	1.0	5.8	0.9	13.5	3.9	18.2	4.1
Time until target fixation (ms)	592.7	134.3	771.5	137.5	2,648.0	946.7	3,928.3	963.8
Average fixation duration (ms)	211.8	38.4	214.5	32.1	256.2	35.0	260.7	36.2
Decision time (ms)	624.5	232.7	592.9	164.3	1,164.9	486.6	1,371.3	678.5
First saccade amplitude (°)	4.2	1.0	4.2	0.7	1.9	0.5	1.5	0.4
Distance from target: Fixation 1 (°)	6.6	1.0	7.6	0.9	6.9	0.9	7.6	0.9
Distance from target: Fixation 2 (°)	5.2	0.9	6.5	0.9	6.4	1.0	7.7	0.9
Distance from target: Fixation 3 (°)	4.1	1.1	4.9	0.9	6.2	1.0	7.7	0.8
Distance from target: Fixation 4 (°)	3.6	1.0	3.5	0.9	6.4	1.2	7.8	0.7

(2009), and Võ and Henderson (2010). The one inconsistency is that Võ and Henderson found an effect of preview on first-saccade amplitudes in windowed visibility. To see whether this inconsistency of results was due to our choice of scene/target combinations (perhaps our scenes provided less information about potential target locations), we reran analyses on the half of the scenes that had fewer than the median number of fixations in the previewed, windowed condition. This removed images in which the preview gave very little information about where the target might be. The same pattern of results was obtained. Therefore, we speculate that the lack of an effect of preview on first-saccade amplitudes might be due to mixing windowed-visibility trials with full-visibility trials, thereby weakening participants' motivation to try to visually process the preview in the windowed condition.

Fig. 3 Fixation distances from the target during early stages of search for the four display conditions. Error bars represent one standard error of the mean

There was, however, a difference between the windowed- and full-visibility conditions in the effect of preview on the proximity of the first four fixations to the target. Windowed viewing without a preview was aimless. There was no evidence that early successive fixations moved progressively closer to the target. Adding a preview to windowed viewing led to fixations being closer to the target overall and to the first few fixations moving progressively closer to the target.

Searching fully visible scenes was efficient even without a preview. Search was rapid, and fixations moved progressively closer to the target. Adding a preview benefitted only the second fixation. Thus, the effect of the preview under normal scene viewing conditions had a very limited duration. When peripheral information was available during search, participants relied much more on the visible characteristics of the scene than on gist from the preview when planning their eye movements. It is also possible that the effect of the preview was larger in this study than it would be in real life, for two reasons. First, the target was present in all previews, and although it is unlikely that the full preview effect was due to seeing the target, this might have played a role in some trials. Second, the mixture of full-visibility and windowed-viewing conditions might have led participants to pay more attention to the preview than they would have if they did not anticipate the possibility that search would be difficult because of windowing. Without running a condition in which viewing type during search was blocked, we cannot be sure whether this was the case.

Castelhano and Henderson (2007) showed that when nonfoveal information is not available during fixations, gist information extracted from a preview of a scene guides eye movements. Hollingworth (2009) showed that in search through a fully visible scene, scene memory from a long-duration preview (memory that was possibly richer than the gist of the scene) makes search more efficient. We have

extended this work by showing that in search through a fully visible scene, scene gist information, too, can guide eye movements for a short time and make search more efficient, by placing the first one or two fixations in a more information-rich location of the display. After that, however, search is guided by the nonfoveal information obtained online from those and subsequent fixations.

Although the source of guidance may differ according to visibility, search through a scene is made more efficiency by a brief scene preview, regardless of whether visibility is windowed or full. In contrast, foreknowledge from a brief, or even a longer, preview of a randomly ordered search array does not improve search efficiency when the display is fully visible during search (Becker & Pashler, 2005; Wolfe et al., 2000). This adds indirectly to the volume of evidence accumulating showing that scanning of meaningful environments is driven more by prior knowledge than by stimulus salience.

References

- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, 7, 66–80.
- Becker, M. W., & Pashler, H. (2005). Awareness of the continuously visible: Information acquisition during preview. *Perception & Psychophysics*, 67, 1391–1403. doi:10.3758/BF03193644
- Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, 103, 597–600. doi:10.1037/h0037158
- Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, 112, 291–328. doi:10.1037/0033-295X.112.2.291
- Castelhano, M. S., & 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.
- Greene, M. R., & Oliva, A. (2009). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology*, 58, 137–176. doi:10.1016/j.cogpsych.2008.06.001
- Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*, 17, 273–291. doi:10.1080/13506280802193367
- Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, 47, 3286–3297.
- Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology: Applied*, 8, 99–117.
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research: Visual Perception*, 155, 23–36.
- Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, 81, 10–15. doi:10.1037/h0027470
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. doi:10.1037/0033-2909.124.3.372
- 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.
- 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. doi:10.1037/0033-295X.113.4.766
- Underwood, G., Crundall, D., & Hodson, K. (2005). Confirming statements about pictures of natural scenes: Evidence of the processing of gist from eye movements. *Perception*, 34, 1069–1082.
- van Diepen, P. M. J., De Graef, P., & d’Ydewalle, G. (1995). Chronometry of foveal information extraction during scene perception. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 349–362). Amsterdam: Elsevier.
- van Diepen, P. M. J., & d’Ydewalle, G. (2003). Early peripheral and foveal processing in fixations during scene perception. *Visual Cognition*, 10, 79–100.
- Vö, 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. doi:10.1167/10.3.14
- Vö, 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. doi:10.1080/13506280802547901
- Wolfe, J. M., Klempe, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 693–716. doi:10.1037/0096-1523.26.2.693