

# Peripheral Guidance in Scenes: The Interaction of Scene Context and Object Content

Effie J. Pereira and Monica S. Castelhana  
Queen's University

In the present study, we examined how gaze guidance is affected by immediately available information in the periphery and investigated how search strategies differed across manipulations in the availability of scene context and object content information. Across 3 experiments, participants performed a visual search task in scenes while using a gaze-contingent moving-window paradigm. Extrafoveal information was manipulated across conditions to examine the contributions of object content, scene context, or some combination of the two. Experiment 1 demonstrated a possible interaction between scene context and object content information in improving guidance. Experiments 2 and 3 supported the notion that object content is selected for further scrutiny based on its position within scene context. These results suggest a prioritization of object information based on scene context, such that contextual information acts as a framework in the selection of relevant regions, and object information can then affect which specific locations in those regions are selected for further examination.

*Keywords:* gaze guidance, object content, scene context, visual search

Every person performs hundreds of searches every day; whether to find keys or a coffee mug, our visual and cognitive processes function with relative ease to guide us to the objects that we seek. Monitoring eye movements has allowed researchers to obtain unobtrusive insight into real-time visual and cognitive processing, thus allowing for the examination of information that is used to direct our gaze.

When describing eye movements, a distinction is drawn between two temporal phases of viewing—during *saccades* we orient our eyes toward regions of interest for further scrutiny, and during *fixations* our eyes remain relatively stable to allow for the processing of visual information at that location (Carpenter, 1988; Matin, 1974). As a function of the eye's optics, the gradient in retinal cone density, and the cortical sampling density of foveal photoreceptors, acuity is highest at the central point of fixation and drops off exponentially as eccentricity increases (Anstis, 1974; Strasburger, Rentschler, & Jüttner, 2011; Wilson, Levi, Maffei, Rovamo, & De Valois, 1990).

Although fine-detailed discrimination requires foveal processing (central  $\sim 2^\circ$ ), extrafoveal processing is used to process con-

textual information and to guide saccades to the next fixation point. The importance of peripheral information in processing contextual information has been demonstrated across a number of studies (Boucart, Moroni, Szaffarczyk, & Tran, 2013; Boucart, Moroni, Thibaut, Szaffarczyk, & Greene, 2013; Larson & Loschky, 2009). Further, many studies have emphasized the important role peripheral information plays in directing eye movements in scenes (Ehinger, Hidalgo-Sotelo, Torralba, & Oliva, 2009; Mannan, Ruddock, & Wooding, 1995; Neider & Zelinsky, 2006; Spotorno, Malcolm, & Tatler, 2014; Torralba, Oliva, Castelhana, & Henderson, 2006; Zelinsky & Schmidt, 2009). Thus, as an image is being viewed, fixations are not only directed by ongoing visual and cognitive processing of information at the point of fixation, but are also affected by available information in the periphery (Henderson & Hollingworth, 1999; Rayner, 2009). Recent research exploring peripheral contributions to guidance during search has demonstrated that when looking for a target within real-world scenes, eye movements are guided by two main information sources: object content and scene context (Castelhana & Henderson, 2007; Ehinger et al., 2009; Neider & Zelinsky, 2006; Spotorno et al., 2014; Torralba et al., 2006; van Diepen, Wampers, & d'Ydewalle, 1995; Vö & Schneider, 2010; Zelinsky & Schmidt, 2009).

When examining the influence of object information, early studies investigating gaze guidance during search have found that fixations tend to be directed toward objects that share similar target features (Findlay, 1997; Scialfa & Joffe, 1998; Williams & Reinhold, 2001). These results have also been demonstrated in nonhuman primates, showing that peripheral target features result in higher activation in single cell recordings than nontarget features in dimensions of luminance and color (Motter, 1994; Reynolds, Pasternak, & Desimone, 2000). Furthermore, studies that manipulate the availability of peripheral information have found that when target features are available in the periphery, saccades are

---

This article was published Online First August 4, 2014.

Effie J. Pereira and Monica S. Castelhana, Department of Psychology, Queen's University.

We thank Alvin Suen, Keziah Uychoode and Jacky Cheung for their help in creating the stimuli. This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada, Canadian Foundation for Innovation, Ontario Ministry of Research and Innovation, and Queen's University Advisory Research Committee to Monica S. Castelhana, and the Queen Elizabeth II Graduate Scholarship in Science and Technology to Effie J. Pereira.

Correspondence concerning this article should be addressed to Effie J. Pereira, Department of Psychology, Queen's University, Craine Hall 204, 62 Arch Street, Kingston, ON, K7L 3N6. E-mail: [effie.pereira@queensu.ca](mailto:effie.pereira@queensu.ca)

selectively made toward them (Eckstein, Beutter, & Stone, 2001; Pomplun, Reingold, & Shen, 2001). Consequently, recent research has also shown that visual information in the periphery can be strategically processed based on features, in order to guide eye movements toward likely targets (Zelinsky, 2008). However, while eye movements are directed toward target features, researchers have also found objects in general (rather than target features specifically) are selected for more detailed discrimination (Foulsham & Kingstone, 2013; Mannan et al., 1995; Nuthmann & Henderson, 2010). These findings suggest that observers may be parsing scenes into objects and using them as units to select for further scrutiny (i.e., fixating and processing). In the current study, we were interested in exploring the latter: the contribution of general, peripheral object information to the guidance of eye movements in scenes.

With scene images, studies examining extrafoveal information processing have focused on both object content and scene context. Early on, Buswell (1935) suggested that eye movements in scenes are related to the informativeness of selected regions, with viewers tending to fixate on people and objects rather than backgrounds. Mackworth and Morandi (1967) quantified this observation and found a correlation between where fixations occur on images and estimates of the relative importance of these specific regions by a separate group of individuals. Subsequent studies have provided support for this finding and suggested that peripheral object information in scenes does capture attention and guide fixation placement (De Graef, Christiaens, & d'Ydewalle, 1990; Friedman, 1979; Gordon, 2004; Henderson, Weeks, & Hollingworth, 1999; Loftus & Mackworth, 1978). However, most of the above-mentioned studies used line drawings as stimuli, which invariably simplifies how objects are parsed from their background and limits the generalizability to more complex images.

When using more complex, full-color, real-world scenes, findings typically reflect the greater intricacy of image characteristics, such as semantic richness and depth, which may alter search parameters and guidance (Foulsham & Underwood, 2008; Henderson, Brockmole, Castelhana, & Mack, 2007; Henderson & Ferreira, 2004; Land & Hayhoe, 2001). For instance, semantic knowledge can lead to expectations of where an object is likely to appear within a scene and can be used to limit fixation placement to relevant contextual areas of the scene (Castelhana & Henderson, 2007; Eckstein, Drescher, & Shimozaki, 2006; Neider & Zelinsky, 2006; Torralba et al., 2006; Zelinsky & Schmidt, 2009). Thus, the discussion of gaze guidance in complex real-world scenes is typically driven by the effects of scene context, and less so by object content and features (Ehinger et al., 2009; Wolfe, Vö, Evans, & Greene, 2011). However, it is clear that both contribute to gaze guidance to some extent.

As noted, there is ample evidence that attentional guidance is affected by both scene context and object content information, but it is unclear to what degree each of these factors contribute to guidance and how they potentially interact. In the past, studies have examined these respective contributions by applying Fourier filtering of spatial frequency on images, such that only high (high-pass filtering) or low (low-pass filtering) spatial frequency information is available. Selective removal of high spatial frequencies generates images that preserve coarse global information, such as large changes in luminance. When low spatial frequencies are removed, the image preserves fine details such as contours and

edges, but not contrasts across broader regions. In scene perception, high and low spatial frequencies are roughly associated with object content and scene context, respectively (Schyns & Oliva, 1994; Wampers & van Diepen, 1999).

The use of these transformed images has led to some interesting insights about scene processing. Generally, early researchers conjectured that only coarse peripheral information was used to guide eye movements because more peripheral parts of the retina are unable to resolve finer detail. However, Wampers and van Diepen (1999) examined the effect of low- and high-pass filtering of peripheral information on the guidance of gaze during a search task, and found a clear benefit of having high spatial frequency information in the periphery. They argued that even though high spatial frequency information in periphery could not be identified, it could be detected and could serve as a cue for object locations, thus guiding subsequent fixation placement. Similar patterns of results have been found in more recent studies using filtering and other techniques to degrade images, though the benefit of high versus low spatial frequencies in the periphery on eye movement guidance seems to vary across tasks (Foulsham, Teszka, & Kingstone, 2011; Laubrock, Cajar, & Engbert, 2013; Loschky & McConkie, 2002; Loschky, McConkie, Yang, & Miller, 2005; McConkie, Wolverson, & Loschky, 2001; Nuthmann, 2014). Although it is clear that the differing spatial frequency content in a filtered image does influence fixation placement, it is not clear how higher-level information maps on to these modifications.

Despite the association between higher spatial frequencies and objects and between lower spatial frequencies and scene context, there is an important caveat to note. When a Fourier filter is applied, both object and context availability is manipulated simultaneously, and there is no reason to believe that some object information may not still be available in lower spatial frequencies, and vice versa for context information in higher spatial frequencies (Oliva & Schyns, 1997; Oliva & Torralba, 2007; Schyns & Oliva, 1994). Thus, it remains unclear whether the peripheral availability of scene context and object content has differing effects on eye movement placement. To our knowledge, no one has examined the effects of scene context and object content availability independently in high-quality images.

Thus, in the present study, we sought to examine how eye movements were affected by immediately available information in the periphery, and how search strategies were differentially affected by the availability of scene context and object content information. Across three experiments, participants performed a visual search task in naturalistic scenes while using a gaze-contingent moving-window paradigm (Henderson, McClure, Pierce, & Schrock, 1997; van Diepen, 1997). Unlike previous studies, we directly manipulated the availability of extrafoveal information across conditions to examine the contributions of object content, scene context, or some combination of the two. These effects were examined by analyzing both general eye movement patterns as well as specific eye movement measures that reflected attentional guidance and target verification processes. Eye movements have been similarly segmented in past studies in order to establish a separation between the time taken to find the target and the time required to identify the target once it was fixated (Castelhana & Heaven, 2010; Castelhana, Pollatsek, & Cave, 2008; Malcolm & Henderson, 2009). Although research has shown that both object information and scene context affect gaze

guidance, it is not clear to what degree they affect the decision-making process needed to correctly identify the target, particularly when this information is manipulated extrafoveally.

Experiment 1 examined the individual role of object content and scene context in improving search strategies by presenting this information separately. If object information is prioritized for subsequent fixation placement, we would expect search guidance to be just as efficient with the full scene as when object content alone was available extrafoveally; however, if it is scene context that is prioritized, search will be just as efficient as the full scene when only contextual information is available extrafoveally. Experiments 2 and 3 further scrutinized the role of object content *above and beyond* scene context in guiding fixation placement by presenting scene context and object content information that either overlapped with the target region or did not. In doing so, if guidance was affected by object content information only when scene context was present, we would find slower and less effective searches when object content did not overlap with target regions (Experiments 2 and 3), and searches that were just as efficient as the full scene when object content did overlap (Experiment 3).

### Experiment 1

In Experiment 1, we explored how the availability of scene context and object content information in the periphery affects search performance. Participants searched for a target using a gaze-contingent moving-window. Information presented extrafoveally ( $>2^\circ$  from the center of fixation) was manipulated to include either scene context information only (Empty Scene), object information only (Object Array), both (Full Scene) or neither (No Scene control). If scene context provides the primary source of information in guiding search, then we would expect the Full and Empty Scene conditions to result in more efficient search guidance than the Object Array and No Scene condition. However, if object information represents the primary source of information for search guidance, then we would expect the Object Array and Full Scene conditions to show higher search efficiency than the Empty Scene and No Scene control conditions.

### Method

**Participants.** Thirty Queen's University undergraduate students, with normal or corrected-to-normal vision, participated for course credit or for \$10/hr.

**Apparatus and stimuli.** Eye movements were tracked using an EyeLink 1000 (SR Research; Mississauga, ON) at 2000Hz. The presentation of the stimuli was controlled by Experiment Builder (SR Research), and the stimuli were presented on a 21-in. CRT monitor with a refresh rate of 100 Hz. Participants sat 60 cm away from the display monitor, with their head stabilized by a head and chin rest. Although viewing was binocular, only the right eye was tracked.

The stimuli consisted of 52 computer-generated indoor and outdoor scenes created on Complete Home Design 5.0 (Data Becker; Düsseldorf, Germany). The scenes were displayed at a resolution of  $800 \times 600$  pixels, with images subtending a visual angle of  $38.1^\circ \times 28.6^\circ$ ; targets had an average size of  $3.2^\circ \times 2.8^\circ$ .

The search scene was presented with a  $2^\circ$  radius gaze-contingent moving-window (Henderson et al., 1997; van Diepen, Wampers,

& d'Ydewalle, 1998); the original search scene was presented foveally, while information presented extrafoveally was manipulated across four conditions. In two conditions, we manipulated the object content. For these manipulations, objects were defined (based on the definition proposed by Henderson & Hollingworth, 1999) as smaller-scale discrete entities that are easily moveable within scenes (e.g., books, mugs, trashcans, paintings). This is in contrast to background items that are large-scale and immovable or not-as-easily moveable structures or surfaces (e.g., cabinets, fridges, beds, couches). Targets were then defined as one of these objects in the scene, thus ensuring that the task-relevant object came from the same set as the manipulated objects.

The four extrafoveal conditions were (a) Full Scene: the original search scene excluding the target, (b) Empty Scene: the search scene with all of the objects removed, (c) Object Array: a gray screen displaying all of the objects within the scene, and (d) No Scene: a black screen control containing no information. None of the extrafoveal conditions contained the target object. Thus, the Empty Scene condition provided only scene context information, while the Object Array condition provided only object information. Figure 1 details example extrafoveal scene conditions for Experiment 1.

**Procedure.** Prior to the start of the experiment, participants were instructed to search the scene for the prespecified target and

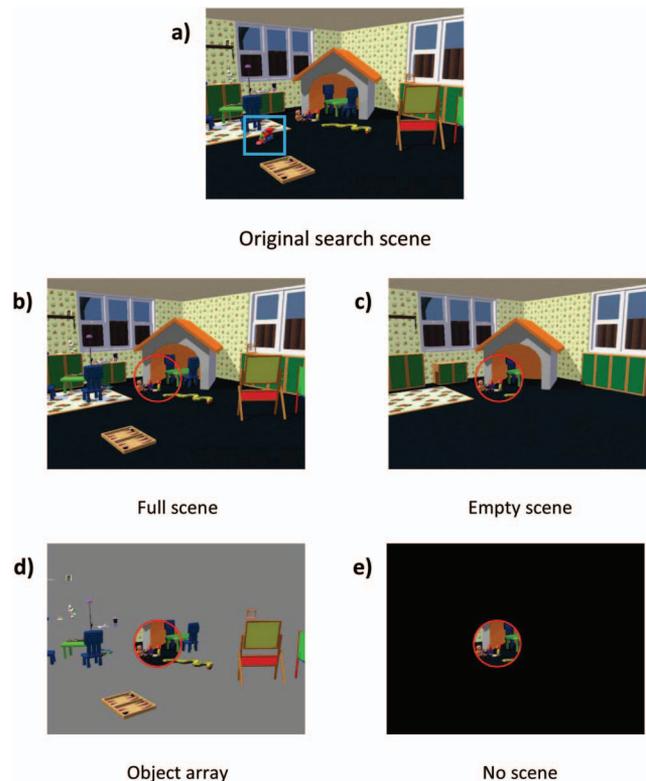


Figure 1. Example extrafoveal scene conditions for Experiment 1. The foveal scene was shown inside the gaze-contingent moving-window, and the extrafoveal scene was shown outside. For illustrative purposes, the target is highlighted in blue in the original search scene (a), and the moving-window is highlighted in red in the extrafoveal scenes: (b) Full Scene, (c) Empty Scene, (d) Object Array, and (e) No Scene.

to press a response button once they had found it. Participants were instructed that they would be viewing the scene through a moving-window and that this window was tied to their fixation. They were then calibrated on the eye tracker using a nine-point calibration screen to ensure high accuracy. The average spatial error was no greater than  $.4^\circ$ , and the maximum error never exceeded  $.7^\circ$ . Calibration was also checked prior to every trial using a five-point calibration screen to confirm accurate positioning of the gaze-contingent moving-window. For each trial, participants were first presented with a target word in the center of the screen for 2 s, followed by a fixation cross for 500 ms. The search scene was then displayed with a  $2^\circ$  radius gaze-contingent moving-window until a response was made or until 20 s had elapsed. The stimuli presented outside of the moving-window ( $>2^\circ$ ) varied with the extrafoveal condition but never included the target object. Participants were informed that the target would only be visible within the window. A typical trial sequence is illustrated in Figure 2. Participants completed four practice trials that depicted each of the extrafoveal conditions prior to the 48 experimental trials. Because we could use eye movements to verify that target objects had indeed been found, there were no target-absent trials. Conditions were counterbalanced across participants in a within-subjects design (12 trials/extrafoveal condition), appeared in a randomized order for each participant, and participants saw each search scene only once. The experiment lasted approximately 30 min.

## Results

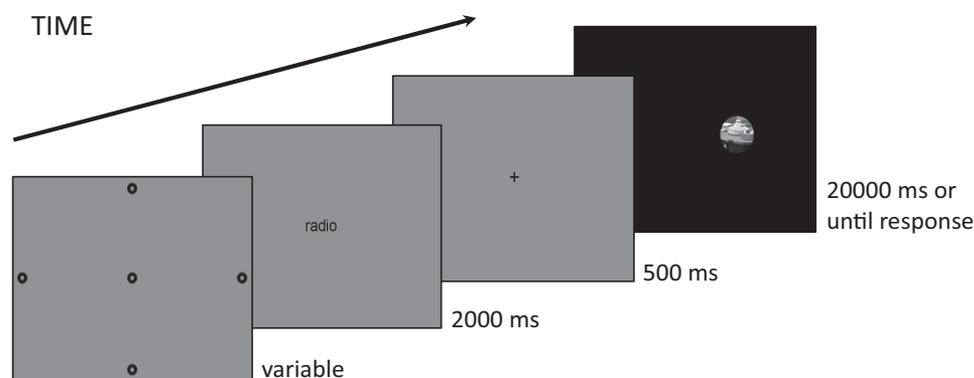
**Data analysis.** To investigate the relative effects of scene context and object content information availability extrafoveally, we examined both behavioral and eye movement measures. For the behavioral measures, we calculated accuracy and overall reaction time (RT). Because we were interested in processes involved in accurate search performance, all participants who performed at less than 80% accuracy were excluded from further analysis (two participants excluded, leaving 28 participants included in the analysis). For the eye movement analysis, targets were defined by a rectangular region approximately  $1^\circ$  from its outermost edge, and we examined three types of eye movement measures: (a) general,

(b) visual search, and (c) target verification. For the general eye movement measures, we examined fixation duration and saccade length across the entire viewing period. For the visual search eye movement measures, we examined patterns that reflected attentional guidance in visual search, which was defined as the time from the onset of the search scene to the beginning of the first fixation on the target. For the target verification measures, we examined eye movements that reflected the time required to identify the target, which was defined as the time from the beginning of the first fixation on the target to the response. For all eye movement measures, fixation durations less than 90 ms were removed from the analyses to eliminate artifacts from the eye tracker, and durations greater than 1,200 ms were excluded as outliers. Based on these criteria, 1,242 (3.6%) from a total of 34,869 fixations were dropped.

For each measure, we conducted six planned comparisons. To examine whether scene context provided a primary source of peripheral guidance during search, the Full Scene and Empty Scene were compared with the Object Array and No Scene. To examine whether extrafoveal object information aided in search, the Object Array condition was compared with the No Scene. Finally, to examine whether extrafoveal object information interacted with scene context and produced greater efficiency in search, the Full Scene was compared with the Empty Scene. To avoid possible Type I errors due to multiple planned comparisons, a Bonferroni correction was used ( $\alpha_{FW} = .05$ ;  $\alpha = .008$ ). Table 1 and Figure 3 summarize the behavioral and eye movement measures across the four extrafoveal scene conditions.

## Behavioral Measures

**Accuracy.** A trial was scored as accurate if the participant fixated on the target and pressed the button within three fixations of the target fixation; means are presented in Table 1. The overall accuracy rate was 90% and analysis showed that the No Scene condition had a significantly lower accuracy than the Full Scene,  $t(27) = 4.72$ ,  $p < .001$ ,  $d = 1.27$ , and Object Array conditions,  $t(27) = 4.17$ ,  $p < .001$ ,  $d = 1.16$ ; however, there were no differences between the Object Array and the Empty Scene con-



*Figure 2.* The trial sequence for all experiments (No Scene condition is depicted here). Participants would begin by fixating on the center point of the calibration screen. A word describing the target would be presented for 2 s, followed by a fixation cross for 500 ms. The search scene would then be shown with a  $2^\circ$  radius circular gaze-contingent moving-window, centered at fixation, for a maximum of 20 s or until the participant had made a response.

Table 1

Mean Accuracy and Target Verification Measures as a Function of Extrafoveal Scene Condition in Experiment 1

|                              | Full Scene |     | Empty Scene |     | Object Array |     | No Scene |      |
|------------------------------|------------|-----|-------------|-----|--------------|-----|----------|------|
|                              | M          | SD  | M           | SD  | M            | SD  | M        | SD   |
| Behavioral measures          |            |     |             |     |              |     |          |      |
| Accuracy (%)                 | 93.4       | 5.7 | 89.2        | 9.6 | 93.2         | 7.2 | 82.1     | 11.3 |
| Target verification measures |            |     |             |     |              |     |          |      |
| First fixation duration (ms) | 244        | 36  | 238         | 32  | 238          | 30  | 245      | 36   |
| First gaze duration (ms)     | 819        | 192 | 771         | 162 | 735          | 171 | 824      | 190  |
| Total time on target (ms)    | 972        | 285 | 954         | 182 | 998          | 233 | 993      | 247  |

ditions,  $t(27) = 1.66$ ,  $p = .11$ ,  $d = 0.46$ . Although participants performed better in certain extrafoveal conditions, the accuracy across all conditions remained high. For the remaining analyses, only correct trials were analyzed.

**Reaction time.** RT was defined as the elapsed time from onset of the search scene until the response button was pressed; means are presented in Figure 3a. On average, participants took approximately 6s to respond. We found the Full Scene condition was

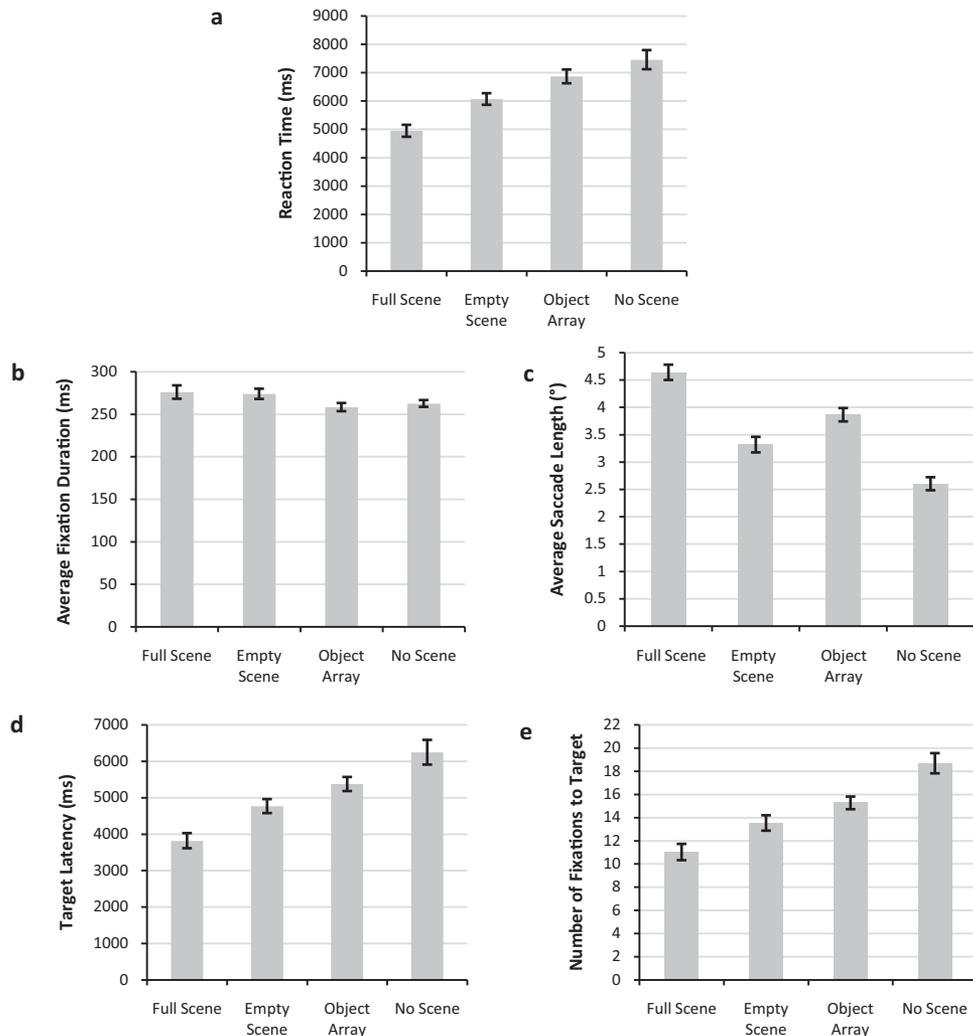


Figure 3. Mean  $\pm 1$  SE for (a) reaction time (ms), (b) average fixation duration (ms), (c) average saccade length ( $^{\circ}$ ), (d) target latency (ms), and (e) number of fixation to target as a function of extrafoveal scene condition in Experiment 1.

significantly faster than the Object Array,  $t(27) = -5.15, p < .001, d = 1.60$ , Empty Scene,  $t(27) = -3.78, p = .001, d = 1.02$ , and No Scene conditions,  $t(27) = -6.76, p < .001, d = 1.69$ , and the Empty Scene condition was significantly faster than the Object Array,  $t(27) = 2.99, p = .006, d = 0.67$ , and No Scene conditions,  $t(27) = -3.48, p = .002, d = 0.94$ . In addition, there was no significant difference between the No Scene and Object Array conditions,  $t(27) = -1.54, p = .14, d = 0.38$ . Thus, when scene context was available (Full and Empty Scene conditions), participants found the target object more efficiently than when only object information was presented (Object Array). Furthermore, there was no benefit in knowing the placement of objects in the periphery when compared to having no information available in the periphery.

### Eye Movement Measures: General

**Average fixation duration.** *Average fixation duration* was defined as the average duration of all fixations within a trial. The distribution of fixation durations for each extrafoveal condition is displayed in Figure 4a and means are presented in Figure 3b. The pattern from the distribution graph suggests that fixation duration is incrementally longer with more information available in the periphery. We found that the Object Array condition had significantly shorter average fixation durations than the Full Scene,  $t(27) = 3.09, p = .005, d = 0.51$ , and Empty Scene conditions,  $t(27) = -5.99, p < .001, d = 0.54$ . Further, there were no significant differences between the Full and Empty Scene conditions,  $t(27) = 0.41, p = .69, d = 0.06$ , nor between the Object Array and No Scene conditions,  $t(27) = -1.31, p = .20, d = 0.18$ . Hence, interestingly, the availability of scene context information (Full and Empty Scene conditions) resulted in longer processing of visual information than when only object content was available (Object Array).

**Average saccade length.** *Average saccade length* was defined as the average distance in degrees of visual angle between two consecutive fixations. The distribution of saccade lengths for each extrafoveal condition is displayed in Figure 4b and means are presented in Figure 3c. When comparing the average saccade lengths across conditions, we found a pattern of effects wherein

greater amounts of visual information available in the periphery resulted in longer average saccade lengths (Loschky & McConkie, 2002; Nuthmann, 2014; Shioiri & Ikeda, 1989): the Full Scene condition was significantly longer than all of the other conditions (Empty Scene,  $t(27) = 13.67, p < .001, d = 1.77$ , Object Array,  $t(27) = 7.51, p < .001, d = 1.12$ , and No Scene conditions,  $t(27) = 13.77, p < .001, d = 2.97$ ), followed by the Object Array condition, which was significantly longer than the remaining conditions (Empty Scene,  $t(27) = 6.26, p < .001, d = .78$ , and No Scene conditions,  $t(27) = 10.71, p < .001, d = 1.98$ ), and finally the Empty Scene condition, which had longer saccades than the No Scene condition,  $t(27) = 5.90, p < .001, d = 1.03$ .

### Eye Movement Measures: Visual Search

**Target latency.** *Target latency* was defined as the elapsed time from the onset of the search scene until the first fixation on the target (excluding this first fixation); means are presented in Figure 3d. This measure reveals how effectively participants searched through the scene before locating the target object. We found that the Full Scene condition had significantly shorter latencies than the Empty Scene,  $t(27) = -3.35, p = .002, d = 0.90$ , Object Array,  $t(27) = -4.73, p < .001, d = 1.47$ , and the No Scene conditions,  $t(27) = -6.25, p < .001, d = 1.63$ ; however, the Object Array condition did not significantly differ from the Empty Scene,  $t(27) = 2.74, p = .01, d = 0.60$ , and No Scene conditions,  $t(27) = -2.08, p = .05, d = 0.60$ . Although there was no difference in the time to the first fixation on target when either scene context alone (Empty Scene) or object content alone (Object Array) was available, the presence of both (Full Scene) led participants to the target faster than when only scene context or only object information was available. In addition, knowing the placement of objects in the periphery (in the Object Array) did not result in shorter latencies than having no information at all.

**Number of fixations to target.** *Number of fixations to target* was defined as the number of individual fixations made until the first fixation on the target (excluding the first fixation); means are presented in Figure 3e. Although related to target latency, this measure reveals whether participants were effectively selecting likely target locations for fixations. We found the Full Scene

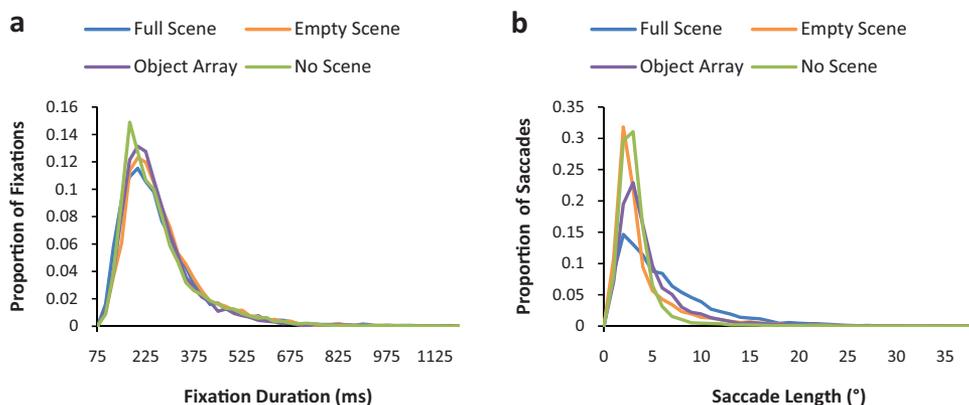


Figure 4. Distributions of (a) fixation durations (ms) and (b) saccade lengths ( $^{\circ}$ ) as a function of extrafoveal scene condition in Experiment 1.

condition had significantly fewer fixations than the Object Array,  $t(27) = -4.59, p < .001, d = 1.28$ , and the No Scene conditions,  $t(27) = -6.59, p < .001, d = 1.83$ ; however, there was no significant difference between the Full Scene and Empty Scene conditions,  $t(27) = -2.44, p = .02, d = 0.69$ . We found in the No Scene control condition, participants made significantly more fixations than either the Empty Scene,  $t(27) = -4.44, p < .001, d = 1.26$ , and Object Array conditions,  $t(27) = -3.24, p = .003, d = 0.89$ . As well, there were no differences found between the Empty Scene and Object Array conditions,  $t(27) = 2.44, p = .02, d = 0.54$ . Thus, fixation selection was highly efficient when scene context was available in the periphery (Full and Empty Scene), regardless of whether object content was available. Further, object content (Object Array) led to significantly fewer fixations to the target than no scene information, suggesting that although not as effective as scene context, the availability of object content had some effect on fixation placement during search.

### Eye Movement Measures: Target Verification

**First fixation duration.** *First fixation duration* was defined as the duration of the initial fixation on the target, and is typically used as a measure of early processing; means are presented in Table 1. There were no significant differences found between any of the extrafoveal conditions,  $ps > .34, ds < 0.20$ . Although peripheral information had an effect on guidance to the target, once the target was fixated, the type of extrafoveal information did not influence the fixation time on the target.

**First gaze duration.** *First gaze duration* was defined as the sum of all fixations on the target from first entry to first exit; means are presented in Table 1. Although related to first fixation duration, this measure provides us with information on later stages of processing. As with the first fixation duration, there were no significant differences detected between any of the extrafoveal conditions,  $ps > .01, ds < 0.50$ .

**Total time on target.** *Total time on target* was defined as the total amount of time spent fixating on the target before the response button was pressed, and is a measure of the total time spent processing the target; means are presented in Table 1. As with the previous verification measure, there were no significant differences found between any of the extrafoveal conditions,  $ps > .11, ds < 0.21$ .

### Discussion

In general, eye movements did differ in their characteristics depending on the availability of peripheral information. Average fixation duration was higher when scene context was available compared to only object information, suggesting that more information was being processed on each fixation when context was present. Furthermore, we found that average saccade length was longer when there was more information available in the periphery. This is consistent with previous research showing that greater visual information has an effect on saccadic length (Loschky & McConkie, 2002; Nuthmann, 2014; Saida & Ikeda, 1979; Shioiri & Ikeda, 1989). In previous gaze-contingent moving-window studies, the amount of visual information was diminished by either low-pass filtering (Loschky & McConkie, 2002; Nuthmann, 2014) or adding random pixel noise (Shioiri & Ikeda, 1989). Across

studies, researchers found that diminished visual information led to shorter saccade lengths as more saccades were made within the high resolution window. In the current study, we manipulated high-level properties of the scene and find a similar pattern of results: as the amount of visual information available increases, so do average saccade lengths. Thus, even across differing methods of varying visual information using a moving-window, there is a correspondence across results.

With the visual search eye movement measures, we demonstrated that the availability of scene context information in the periphery resulted in more effective selection of target locations. This can be seen in comparing the Full Scene and Empty Scene conditions to the No Scene control condition, where target latency was faster and the number of fixations to the target was smaller when scene context was available. This finding is consistent with previous studies showing the benefits of scene context information during visual search (Castelhano & Henderson, 2007; Neider & Zelinsky, 2006; Torralba et al., 2006; Vö & Schneider, 2010). We also found that object content itself had some effect on gaze guidance in scenes, as there were fewer fixations during search when some object information was available as compared to no information. In addition, although object information did not result in shorter overall search times (RT) nor in shorter latency to the target, the effect sizes were still medium to large (according to Cohen, 1998;  $ds$  for RT and target latency were 0.38 and 0.60, respectively). This suggests that there may have been some effect of object content, but that these effects were not as strong as those found in the other extrafoveal conditions.

We did not, however, find any significant differences in the target verification measures. Although there is some evidence in previous studies that extrafoveal visual information may have an influence on the verification of a target object, it was not a complete surprise that there were no differences in the current study. Previous studies have shown that objects in unexpected places take longer to identify than those in expected ones (Castelhano & Heaven, 2011; Malcolm & Henderson, 2010). In the present study, unlike past studies, there were no semantic or location manipulations in target placement, and once the target was fixated, participants would have access to full scene context information through the moving-window.

In summary, the pattern of results from Experiment 1 suggests that scene context may play an especially important role during the initial stages of visual search in guiding eye movements, but the contribution of object information is not entirely negligible. Overall, the results also provide support for a possible interaction between scene context and object content information in improving guidance, as shown by the differences between the Full and Empty Scenes—the additional object information in the former resulted in numerically faster target latencies as compared to when no object information was available. This pattern suggests that scene context information alone does not automatically lead to more efficient searches and that object information may provide a way of prioritizing fixation placement within specific regions of the scene context. If both scene context and object information interact to guide attention more effectively, we should be able to manipulate how fast the target is found depending on where objects are placed within the scene. In Experiment 2, we sought to determine how scene context and object information interact by

manipulating the placement of object content relative to target-relevant areas.

## Experiment 2

In Experiment 2, we examined whether search strategies would be equally affected when the placement of object content was manipulated in the periphery. We introduced a Sparse-Cluster Scene extrafoveal condition that contained a subset of object clusters from the original scene (see Figure 5). The Sparse-Cluster Scene condition provided a limited set of object content information (compared to the Full Scene) that did not overlap with the region in which the target object appeared. If the placement of object content has an influence on attentional guidance during search, we would expect that the Sparse-Cluster Scene condition would slow search efficiency in comparison to the Empty Scene and Full Scene conditions, as eye movements would be directed toward irrelevant object clusters, thus impeding search performance.

## Method

**Participants.** Thirty Queen's University undergraduates, with normal or corrected-to-normal vision, participated for course credit

or for \$10/hr. None of the participants had taken part in Experiment 1.

**Apparatus and stimuli.** The apparatus was identical to Experiment 1. The stimuli were identical to Experiment 1, except for one extrafoveal condition. The Object Array condition in Experiment 1 was replaced by a Sparse-Cluster Scene condition in which the search scene was presented with a subset of object clusters that did not overlap with the target object region. Figure 5 displays the four extrafoveal scene conditions used in Experiment 2.

**Procedure.** The procedure was identical to Experiment 1.

## Results

**Data analysis.** As with Experiment 1, we examined behavioral and eye movement measures across extrafoveal scene conditions. Fixation durations less than 90 ms and greater than 1,200 ms were excluded as outliers; out of 31,567 fixations, 1,085 were dropped (3.4%). Six planned comparisons were conducted for each measure, and participants whose accuracy was less than 80% were excluded from the analysis (two participants excluded, leaving 28 participants included in the analysis). As with the previous experiment, we used a Bonferroni correction to avoid possible Type I errors ( $\alpha_{FW} = .05$ ;  $\alpha = .008$ ). Table 2 and Figure 6 summarize the behavioral and eye movement measures across the four extrafoveal scene conditions.

## Behavioral Measures

**Accuracy.** Overall, the average accuracy rate was 90%. We found that the No Scene condition had significantly lower accuracy than the Full Scene,  $t(27) = 5.26$ ,  $p < .001$ ,  $d = 1.43$ , Sparse-Cluster Scene,  $t(27) = 4.15$ ,  $p < .001$ ,  $d = 1.10$ , and Empty Scene conditions,  $t(27) = 5.41$ ,  $p < .001$ ,  $d = 1.36$ . Although participants performed worse when no information was presented in the periphery, the accuracy for all of the conditions remained high. As with Experiment 1, only correct trials were included in the RT and eye movement measures. Table 2 summarizes the results across the four extrafoveal scene conditions.

**Reaction time.** On average, participants took approximately 6 s to respond and means are presented in Figure 6a. We found that participants were significantly faster in the Full Scene condition than the Sparse-Cluster Scene,  $t(27) = -3.76$ ,  $p = .001$ ,  $d = 1.00$ , Empty Scene,  $t(27) = -2.95$ ,  $p = .007$ ,  $d = 0.87$ , and No Scene conditions,  $t(27) = -7.70$ ,  $p < .001$ ,  $d = 1.92$ . There were no significant differences found between the Sparse-Cluster Scene and Empty Scene conditions,  $t(27) = -0.29$ ,  $p = .78$ ,  $d = 0.07$ ; however, both these extrafoveal conditions resulted in faster searches as compared to the No Scene condition, (Sparse-Cluster vs. No Scene,  $t(27) = -4.65$ ,  $p < .001$ ,  $d = 1.24$ , Empty vs. No Scene,  $t(27) = -4.26$ ,  $p < .001$ ,  $d = 1.05$ ).

## Eye Movement Measures: General

**Average fixation duration.** The distribution of fixation durations for each extrafoveal condition is displayed in Figure 7a and means are presented in Figure 6b. Unlike in Experiment 1, we found no significant differences for the average fixation duration between any of the extrafoveal conditions,  $ps > .03$ ,  $ds < 0.26$ .

**Average saccade length.** The distribution of saccade lengths for each extrafoveal condition is displayed in Figure 7b and means

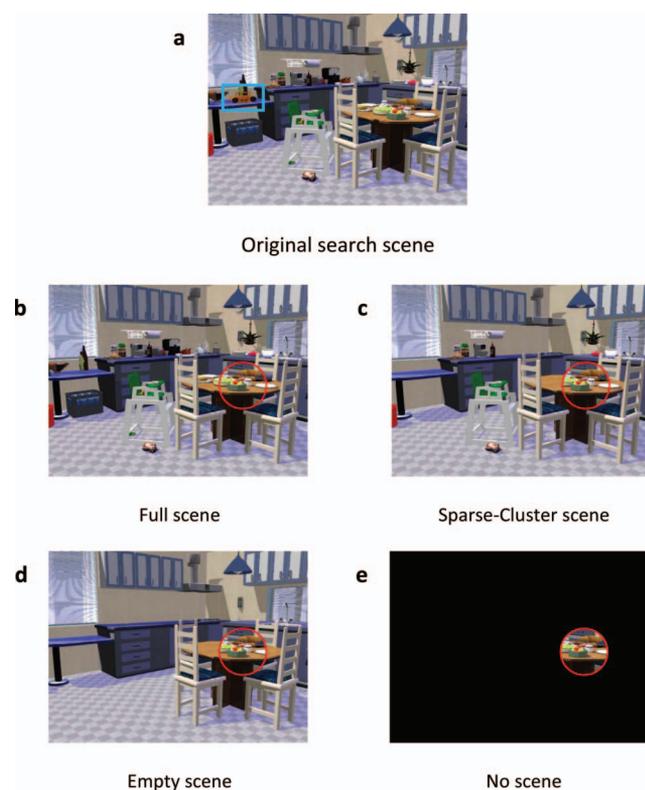


Figure 5. Example extrafoveal scene conditions for Experiment 2. The foveal scene was shown inside the gaze-contingent moving-window, and the extrafoveal scene was shown outside. For illustrative purposes, the target is highlighted in blue in the original search scene (a), and the moving-window is highlighted in red in the extrafoveal scenes: (b) Full Scene, (c) Sparse-Cluster Scene, (d) Empty Scene, and (e) No Scene.

Table 2

Mean Accuracy and Target Verification Measures as a Function of Extrafoveal Scene Condition in Experiment 2

|                              | Full Scene |     | Sparse-Cluster Scene |     | Empty Scene |     | No Scene |      |
|------------------------------|------------|-----|----------------------|-----|-------------|-----|----------|------|
|                              | M          | SD  | M                    | SD  | M           | SD  | M        | SD   |
| Behavioral measures          |            |     |                      |     |             |     |          |      |
| Accuracy (%)                 | 94.3       | 7.5 | 91.1                 | 7.1 | 93.8        | 7.7 | 81.0     | 10.8 |
| Target verification measures |            |     |                      |     |             |     |          |      |
| First fixation duration (ms) | 230        | 29  | 225                  | 27  | 242         | 36  | 231      | 26   |
| First gaze duration (ms)     | 709        | 136 | 687                  | 158 | 697         | 153 | 742      | 174  |
| Total time on target (ms)    | 779        | 173 | 805                  | 212 | 821         | 178 | 848      | 176  |

are presented in Figure 6c. The pattern of average saccade length across conditions was identical to those found in Experiment 1. We found that the amount of information available in the periphery influenced average saccade lengths across conditions. The Full

Scene condition was significantly longer than all of the other conditions: Sparse-Cluster Scene,  $t(27) = 7.00, p < .001, d = 1.13$ , Empty Scene,  $t(27) = 11.96, p < .001, d = 1.82$ , and No Scene conditions,  $t(27) = 20.96, p < .001, d = 4.20$ , followed by

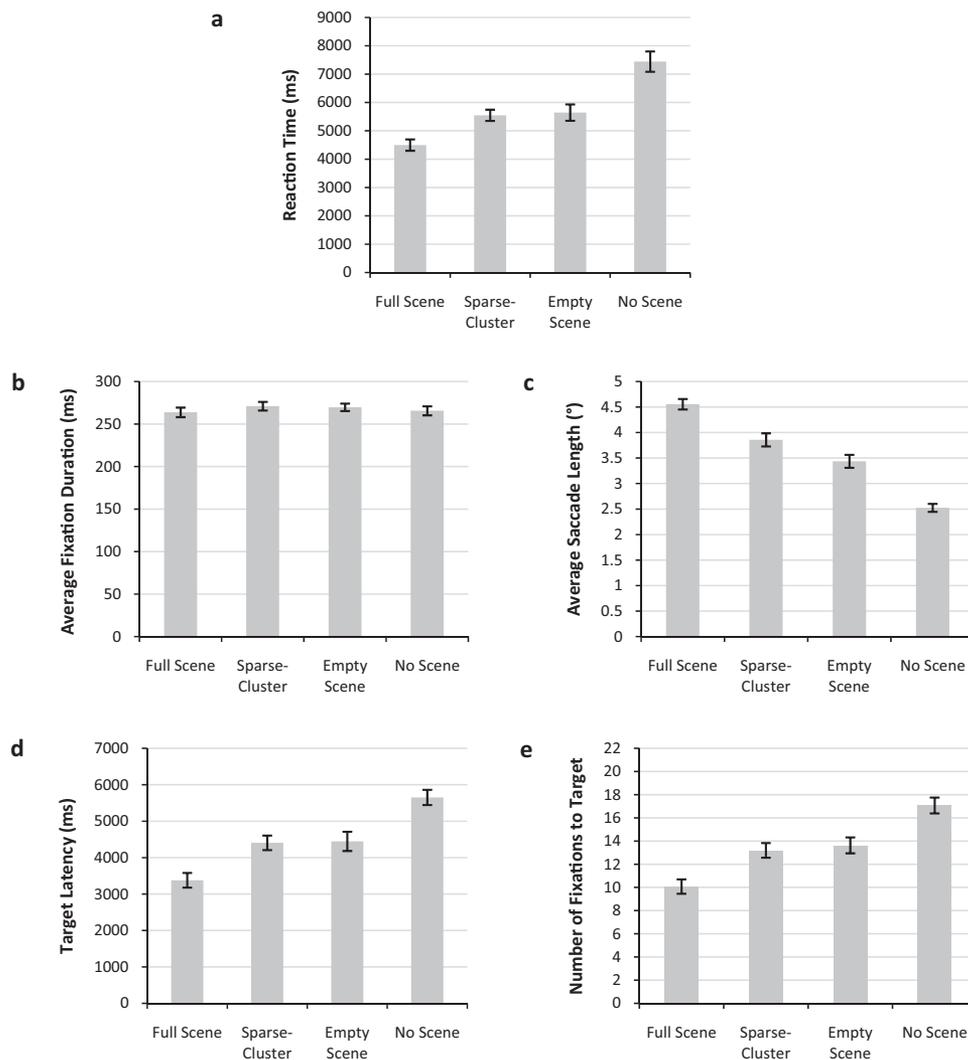


Figure 6. Mean  $\pm$  1 SE for (a) reaction time (ms), (b) average fixation duration (ms), (c) average saccade length ( $^{\circ}$ ), (d) target latency (ms), and (e) number of fixation to target as a function of extrafoveal scene condition in Experiment 2.

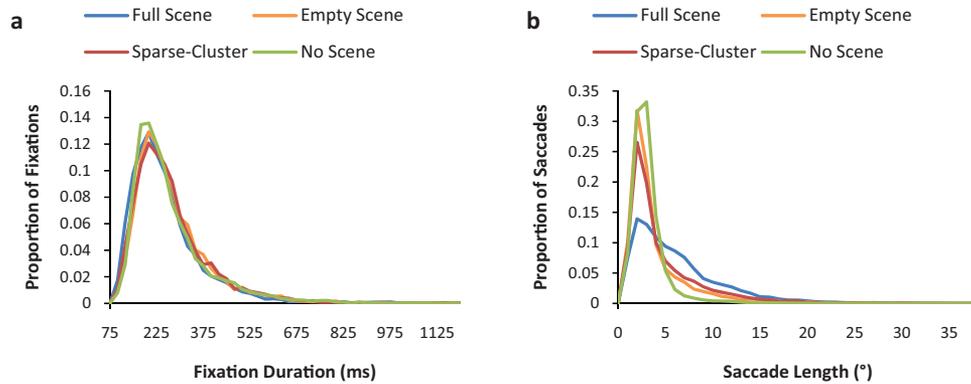


Figure 7. Distributions of (a) fixation durations (ms) and (b) saccade lengths (°) as a function of extrafoveal scene condition in Experiment 2.

the Sparse-Cluster Scene, which was significantly longer than the remaining conditions: Empty Scene,  $t(27) = 8.30$ ,  $p < .001$ ,  $d = 0.62$ , and No Scene conditions,  $t(27) = 13.18$ ,  $p < .001$ ,  $d = 2.37$ , and finally, the Empty Scene had longer saccades than the No Scene condition,  $t(27) = 9.52$ ,  $p < .001$ ,  $d = 1.63$ .

### Eye Movement Measures: Visual Search

**Target latency.** Again, target latency was defined as the elapsed time from the search scene onset to the first fixation on the target (excluding this first fixation); means are presented in Figure 6d. We found that the Sparse-Cluster Scene had significantly longer latencies than the Full Scene condition,  $t(27) = -3.85$ ,  $p = .001$ ,  $d = 0.97$ , but there was no significant difference found between the Sparse-Cluster Scene and Empty Scene conditions,  $t(27) = -0.13$ ,  $p = .90$ ,  $d = 0.03$ . Additionally, there was no difference between the Empty and Full Scene conditions,  $t(27) = -2.70$ ,  $p = .01$ ,  $d = 0.86$ . We also found that the No Scene condition had significantly longer latencies than all the other conditions: Full Scene,  $t(27) = -7.48$ ,  $p < .001$ ,  $d = 2.09$ , Sparse-Cluster Scene,  $t(27) = -4.78$ ,  $p < .001$ ,  $d = 1.15$ , and Empty Scene conditions,  $t(27) = -4.47$ ,  $p < .001$ ,  $d = 0.96$ . Therefore, having object information in the Sparse-Cluster Scene condition resulted in slower latency to the target than the Full Scene condition. Thus, there is some indication that being guided by peripheral object information that does not overlap with the target resulted in poorer search performance.

**Number of fixations to target.** Means are presented in Figure 6e. We found that the Sparse-Cluster Scene had a significantly greater number of fixations than the Full Scene condition,  $t(27) = -3.55$ ,  $p = .001$ ,  $d = 0.94$ , but there was no significant difference between the Sparse-Cluster Scene and Empty Scene conditions,  $t(27) = -0.46$ ,  $p = .65$ ,  $d = 0.12$ . Additionally, we found that the Empty Scene resulted in greater number of fixations than the Full Scene conditions,  $t(27) = -3.13$ ,  $p = .004$ ,  $d = 1.02$ . Finally, we found that the No Scene condition required a significantly higher number of fixations to the target than all other conditions: Full Scene,  $t(27) = -6.81$ ,  $p < .001$ ,  $d = 2.03$ , Sparse-Cluster Scene,  $t(27) = -4.11$ ,  $p < .001$ ,  $d = 1.11$ , and Empty Scene conditions,  $t(27) = -4.25$ ,  $p < .001$ ,  $d = 0.95$ . Hence, similar to the target latency measure, being guided by

object content information in the Sparse-Cluster Scene condition resulted in greater number of fixations and less efficient searches than the Full Scene condition. In addition, the difference between the Full Scene and Empty Scene conditions suggests that object content and scene context information together led to more effective fixation placement than scene context information alone.

### Eye Movement Measures: Target Verification

**First fixation duration.** Means are presented in Table 2. As was found in Experiment 1, there were no significant differences found between any of the extrafoveal conditions,  $ps > .02$ ,  $ds < 0.52$ .

**First gaze duration.** Means are presented in Table 2. As with the previous verification measure, there were no significant differences detected between any of the extrafoveal conditions,  $ps > .08$ ,  $ds < 0.34$ .

**Total time on target.** Means are presented in Table 2. There were no significant differences found across extrafoveal conditions,  $ps > .05$ ,  $ds < 0.40$ .

### Discussion

Similar to Experiment 1, we found that general eye movement parameters differed according to the availability of peripheral information, such that when there was more information available, average saccade lengths increased. However, there were no differences in average fixation duration across extrafoveal conditions, suggesting that processing times were consistent regardless of the availability of peripheral information. Also similarly to Experiment 1, we did not find any target verification differences across extrafoveal conditions, suggesting that scene context and object content information primarily affected attentional guidance to the target.

Experiment 2 replicated the findings of the previous experiment, which found that participants performed better across all measures when scene context was available than when no extrafoveal information was available. However, we also showed that adding information about object content that did not overlap with the target object region (in the Sparse-Cluster condition) negatively affected attentional guidance to the target, as found in the number

of fixations to the target. We posit that because the Sparse-Cluster Scenes did not contain clusters of objects that overlapped with the target's placement in the scene, the object information was "misleading," directing gaze away from the target's location. These results are consistent with an interaction between scene context and object content: whereas scene context provides general information about overall optimal fixation placement, object content appears to provide specific information about where to aim fixations. In the following experiment, we test this interaction more thoroughly by including an extrafoveal condition containing target clusters that overlapped with the target region.

### Experiment 3

In Experiment 3, we further explored the role of object content in the periphery in guiding search by introducing a Target-Cluster Scene extrafoveal condition that contained object clusters which included a cluster surrounding the target object (again, the target itself was not presented in the periphery). As such, the Target-Cluster Scenes contained object content information that could potentially benefit search performance. This was contrasted with the Full Scene and Empty Scene conditions along with the misleading object content information contained in the Sparse-Cluster Scene condition. If object content information does have an influence on search guidance and it interacts with scene context information, we expect that the object clusters in the Target-Cluster Scenes would direct eye movements toward the target more efficiently than in either the Sparse-Cluster and Empty Scene conditions, and that guidance in Target-Cluster condition may be just as effective as the Full Scene condition.

### Method

**Participants.** Sixty<sup>1</sup> Queen's University undergraduates, with normal or corrected-to-normal vision, participated for course credit or for \$10/hr. None of the participants had participated in the previous experiments.

**Apparatus and stimuli.** The apparatus was identical to previous experiments. The stimuli were similar to Experiment 2 with the exception that the No Scene condition was excluded, and we included a Target-Cluster Scene condition, in which the extrafoveal scene had a few clusters of objects, including a cluster overlapping with the target region. Figure 8 displays the four extrafoveal scene conditions used in Experiment 3.

**Procedure.** The procedure was identical to that of the previous experiments.

### Results

**Data analysis.** Again, we examined behavioral and eye movement measurements. As with the previous experiments, fixation durations less than 90 ms and greater than 1,200 ms were excluded as outliers (from a total of 70,523 fixations, 3,488 were dropped [4.9%]). Participants with an accuracy score less than 80% were excluded from the analysis (none of the participants were excluded). For each measure, six planned comparisons were conducted using a Bonferroni correction ( $\alpha_{FW} = .05$ ;  $\alpha = .008$ ). To examine whether beneficial object content information influences search guidance, the Full Scene and Target-Cluster Scene were each compared with the Sparse-Cluster and Empty Scene conditions. To examine how beneficial

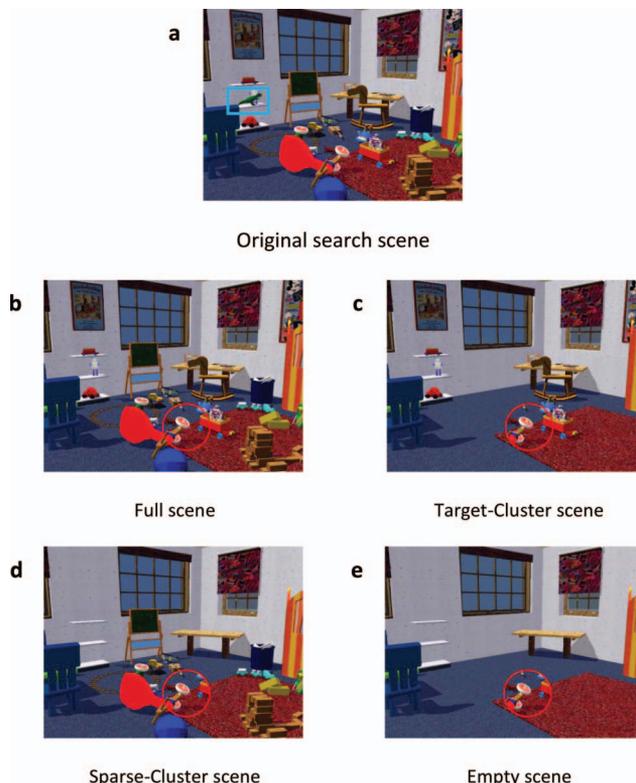


Figure 8. Example extrafoveal scene conditions for Experiment 3. The foveal scene was shown inside the gaze-contingent moving-window, and the extrafoveal scene was shown outside. For illustrative purposes, the target is highlighted in blue in the original search scene (a), and the moving-window is highlighted in red in the extrafoveal scenes: (b) Full Scene, (c) Target-Cluster Scene, (d) Sparse-Cluster Scene, and (e) Empty Scene.

object information influenced search guidance, the Full Scene was compared with the Target-Cluster Scene. Finally, to examine whether misleading object information interacted with scene context to produce lower efficiency in search, the Sparse-Cluster Scene was compared with the Empty Scene. Table 3 and Figure 9 summarize the behavioral and eye movement measures across the four extrafoveal scene conditions.

### Behavioral Measures

**Accuracy.** Accuracy was calculated the same way as with the previous experiments, with trials scored as correct if participants

<sup>1</sup> According to the guideline set by Cohen (1998), the effect size of the visual search measures in Experiment 2 were quite large between the Full and Empty Scenes ( $d > 1$ ), but quite small between the Empty and Sparse-Cluster Scenes ( $d < .1$ ). We hypothesized that the difference between the Target-Cluster and the Empty condition would be slightly larger than the Sparse-Cluster, but not as large as that between the Full and Empty condition. Thus, we performed an a priori power analysis to estimate sample size based on a medium effect size (.3 to .5). This led to an estimated sample size of 66 to 176 participants. We elected to run 60 participants based on the lower end of this estimate and due to counterbalancing constraints. Having such a large sample size for an eye tracking study is unusual, but the factors we were looking at required more power to detect a difference, if one existed.

Table 3

Mean Accuracy and Target Verification Measures as a Function of Extrafoveal Scene Condition in Experiment 3

|                              | Full Scene |     | Target-Cluster Scene |     | Sparse-Cluster Scene |     | Empty Scene |     |
|------------------------------|------------|-----|----------------------|-----|----------------------|-----|-------------|-----|
|                              | M          | SD  | M                    | SD  | M                    | SD  | M           | SD  |
| Behavioral measures          |            |     |                      |     |                      |     |             |     |
| Accuracy (%)                 | 93.6       | 6.6 | 90.6                 | 7.7 | 91.4                 | 7.7 | 90.2        | 9.1 |
| Target verification measures |            |     |                      |     |                      |     |             |     |
| First fixation duration (ms) | 237        | 27  | 242                  | 34  | 234                  | 30  | 233         | 35  |
| First gaze duration (ms)     | 697        | 145 | 676                  | 118 | 700                  | 122 | 706         | 131 |
| Total time on target (ms)    | 785        | 168 | 796                  | 177 | 807                  | 163 | 826         | 163 |

fixated on the target within three fixations of the button press. The average accuracy rate was 92% and did not significantly differ across extrafoveal scene condition,  $ps > .02$ ,  $ds < 0.42$ . RT and eye movement measures included only correct trials. Table 3 summarizes these results across the four extrafoveal scene conditions.

**Reaction time.** On average, RTs were approximately 5 s and means are presented in Figure 9a. Search times were faster in the Full Scene compared to both the Sparse-Cluster Scene,  $t(59) = -5.30$ ,  $p < .001$ ,  $d = 0.89$  and Empty Scene condition,  $t(59) = -5.98$ ,  $p < .001$ ,  $d = 1.01$ . There were no differences found between the Full Scene and Target-Cluster Scene condi-

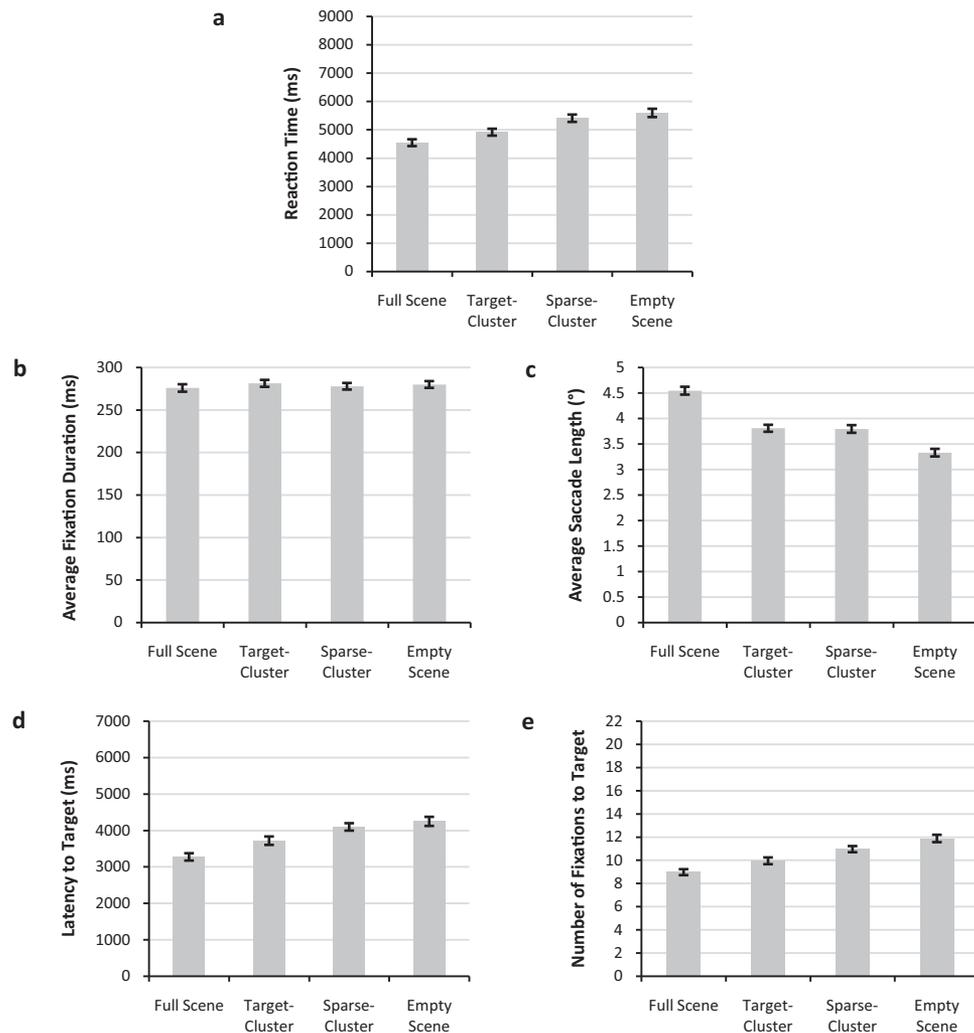


Figure 9. Mean  $\pm$  1 SE for (a) reaction time (ms), (b) average fixation duration (ms), (c) average saccade length ( $^{\circ}$ ), (d) target latency (ms), and (e) number of fixation to target as a function of extrafoveal scene condition in Experiment 3.

tions,  $t(59) = -2.14, p = .04, d = 0.40$ , nor between the Sparse-Cluster Scene and Empty Scene conditions,  $t(59) = -0.91, p = .37, d = 0.18$ . There was also no significant difference between the Target-Cluster Scene and Sparse-Cluster Scene conditions,  $t(59) = -2.65, p = .01, d = 0.51$ ; however, the Target-Cluster Scene was faster than the Empty Scene condition,  $t(59) = -3.88, p < .001, d = 0.65$ . Thus, providing target-related object content information (Target-Cluster Scene) resulted in faster search times than when only scene context was available (Empty Scene), and it equalled the benefit of the information provided by the Full Scene condition.

### Eye Movement Measures: General

**Average fixation duration.** The distribution of fixation duration for each extrafoveal condition is displayed in Figure 10a and means are presented in Figure 9b. There were no significant differences between any extrafoveal conditions,  $ps > .07, ds < 0.17$ .

**Average saccade length.** The distribution of saccade lengths for each extrafoveal condition is displayed in Figure 10b and means are presented in Figure 9c. The pattern of average saccade lengths across conditions were similar to those found in Experiment 2. Saccade lengths in the Full Scene condition were significantly longer than in all of the other conditions (Target-Cluster Scene,  $t(59) = 12.33, p < .001, d = 1.31$ , Sparse-Cluster Scene,  $t(59) = 11.91, p < .001, d = 1.27$ , and Empty Scene conditions,  $t(59) = 16.16, p < .001, d = 2.07$ ). In turn, the Empty Scene condition had significantly shorter saccades than the Target-Cluster Scene,  $t(59) = 8.12, p < .001, d = 0.86$ , and Sparse-Cluster Scene conditions,  $t(59) = 8.39, p < .001, d = 0.80$ . However, there was no significant difference found between the Target-Cluster Scene and Sparse-Cluster Scene conditions,  $t(59) = 0.24, p = .81, d = 0.02$ .

### Eye Movement Measures: Visual Search

**Target latency.** Means are presented in Figure 9d. The Full Scene condition had a shorter target latency than all other extrafoveal conditions: Target-Cluster Scene,  $t(59) = -3.05, p = .003, d = 0.53$ , Sparse-Cluster Scene,  $t(59) = -6.73, p < .001, d = 1.05$ , and Empty Scene conditions,  $t(59) = -6.39, p < .001, d =$

1.11. We found no differences between the Sparse-Cluster Scene and the Target-Cluster Scene conditions,  $t(59) = -2.44, p = .02, d = 0.45$ , nor between the Sparse-Cluster Scene and Empty Scene conditions,  $t(59) = -0.95, p = .34, d = 0.17$ . However, importantly, the Target-Cluster Scene latency was significantly shorter than the Empty Scene condition,  $t(59) = -3.37, p < .001, d = 0.57$ .

**Number of fixations to target.** Means are presented in Figure 9e. Participants made fewer fixations in the Target-Cluster Scene than the Empty Scene condition,  $t(59) = -4.56, p < .001, d = 0.82$ , and fewer fixations in the Full Scene than the Sparse-Cluster Scene condition,  $t(59) = -5.73, p < .001, d = 0.98$ . There were no differences between the Full Scene and Target-Cluster Scene conditions,  $t(59) = -2.61, p = .01, d = 0.46$ , nor between the Sparse-Cluster and Empty Scene conditions,  $t(59) = -2.37, p = .02, d = 0.41$ . In addition, there was no significant difference found between the Target-Cluster and Sparse-Cluster Scene conditions,  $t(59) = -2.57, p = .01, d = 0.47$ . As such, participants had a similar number of fixations to the target with beneficial object information (Target-Cluster Scene) as with the entire scene (Full Scene), and a similar number of fixations with misleading object information (Sparse-Cluster Scenes) as with only scene context information (Empty Scene). This suggests that misleading object information resulted in less effective search and having some beneficial object information led to more efficient guidance toward the target, more so than when scene context information alone was available in the periphery.

### Eye Movement Measures: Target Verification

**First fixation duration.** Means are presented in Table 3. No significant differences were found between the extrafoveal conditions,  $ps > .02, ds < 0.27$ .

**First gaze duration.** Means are presented in Table 3. There were no significant differences found across extrafoveal conditions,  $ps > .04, ds < 0.24$ .

**Total time on target.** Means are presented in Table 3. No significant differences were found across extrafoveal conditions,  $ps > .02, ds < 0.25$ .

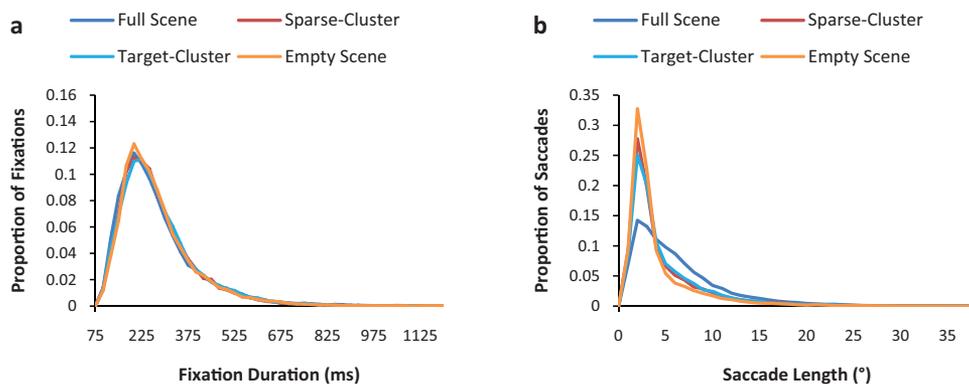


Figure 10. Distributions of (a) fixation durations (ms) and (b) saccade lengths ( $^{\circ}$ ) as a function of extrafoveal scene condition in Experiment 3.

## Discussion

Similar to Experiment 2, we found no differences in average fixation duration across the extrafoveal conditions and as more information was available in the periphery, participants tended to make longer saccades. However, there were no differences in average saccade length based on whether misleading or beneficial object information was present in the periphery. This pattern seems intuitive as the Target-Cluster and Sparse-Cluster Scene conditions would be the most similar to one another in terms of information density (as compared to other extrafoveal conditions). Similarly to the previous two experiments, no differences were found across any of the target verification measures, suggesting that extrafoveal information was not beneficial to processing the target after it had already been located.

Experiment 3 further investigated the role of object information in guiding eye movements by showing that object content influences gaze above and beyond scene context information. The difference between the Target-Cluster and Empty Scene conditions in the number of fixations to the target showed more effective fixation placement when target-relevant information was available, and demonstrated that object information does guide fixation placement during visual search. Furthermore, the presence of this target-relevant object information also resulted in shorter target latency when compared to the Empty Scene condition, suggesting that guidance by object information can be used to improve search performance. These results suggest that although scene context information plays a central and beneficial role in directing eye movements, object content can provide more refined guidance when available in the periphery.

## General Discussion

In the present study, we examined how eye movements were guided by information in the periphery and the degree to which search strategies were modulated by two key sources of peripheral information: scene context and object content. In Experiment 1, we directly compared scene context and object content information by having each type of information presented in the periphery alone. When compared to the Full Scene and No Scene control, we found that search performance (as reflected in behavioral and eye movement measures) was superior for contextual information than object information. The eye movement patterns also provided support for an interaction between the two sources, as some additional search benefit was attained from object information when scene context information was available.

In Experiment 2, we attempted to determine how scene context and object information interacted by manipulating object placement in the periphery. Our findings demonstrated that adding information about object content that was not beneficial to search led to less effective search strategies than having either the entire scene with objects available or scene context information alone. These results lend further support to the notion that object information also plays a significant role in guidance.

In Experiment 3, we further investigated this interaction by altering object information to either be beneficial or disadvantageous to search performance. Object clusters were placed in the scene so that they either overlapped with the target region (Target-Cluster condition) or did not overlap (Sparse-Cluster condition).

Consistent with Experiment 2, we found that object information did affect performance as reflected in the eye movement measures. Search performance was more efficient when objects overlapped with the target region in the periphery. As well, across all experiments, extrafoveal information did not appear to affect how long it took to process and verify the target object once it had been fixated. Taken as a whole, this study suggests that scene context plays a substantial role in guiding search toward probable target locations. However, when object information was available within scene context, it guided eye movements to specific locations within the relevant region. We believe these results are consistent with previous research that posits that contextual information acts as a framework in the selection of relevant regions (Neider & Zelinsky, 2006; Pomplun et al., 2001; Torralba et al., 2006), but the availability of object information affects which specific locations in those regions are selected for further scrutiny.

Much like previous moving-window studies, across all experiments, we found that there was a relationship between the amount of visual information available extrafoveally and the average saccade length. We found that as the amount of visual information increased (or as more potential targets were available), saccade lengths also increased. This pattern is consistent with previous studies that degraded extrafoveal information by either blurring (Loschky & McConkie, 2002; Nuthmann, 2014) or adding visual noise (Shioiri & Ikeda, 1989). These studies found that average saccade length greatly decreased as the degradation of visual information increased. Together, these studies show that regardless of the type of manipulation used, the less available visual information there is in extrafoveal regions, the less likely that those regions will be targeted by fixations.

## The Interplay of Scene Context and Object Content

Scene context information is known to have a strong effect on search strategies by determining where fixations are placed (Castelhano & Heaven, 2010; Neider & Zelinsky, 2006; Torralba et al., 2006; Wolfe, Alvarez, Rosenholtz, Kuzmova, & Sherman, 2011; Zelinsky & Schmidt, 2009). Our findings add to previous research in the visual search literature on the effects of peripherally available contextual information, and more importantly, emphasizes the role of scene context as a primary framework in directing attentional guidance (Neider & Zelinsky, 2006; Pomplun et al., 2001; Torralba et al., 2006).

One interesting question that arises from the current study is what would occur if we put target relevant information (i.e., objects that overlap in target features) outside the expected scene context region? We know from previous studies that target feature information attracts fixations as it very likely corresponds to a high target-likelihood (Eckstein et al., 2001; Pomplun et al., 2001). However, we also know from past research that target objects that are misplaced in a scene or placed in an incompatible scene context take much longer to find and process (Davenport & Potter, 2004; Malcolm & Henderson, 2010). Additionally, in a recent study, Castelhano and Heaven (2011) found that even when scene context information is incompatible with a target object, participants still tend to look for targets in places that are equivalent to its likely placement within its normally compatible scene. For instance, participants searching for an open cookbook in a bathroom would look on the bathroom counter, akin to its expected place-

ment in a kitchen (i.e., counter). These results suggest that we are strongly biased toward using scene context information. In a more recent study, [Spotorno et al. \(2014\)](#) manipulated the quality of the target template (picture vs. word) and the placement of the target object (expected vs. unexpected). Interestingly, when the target template was highly specified (picture), then the placement of the target within the scene had less of a detrimental effect on search than if the target was specified as a word. Taken together, these studies suggest that while we are strongly biased toward using scene context information, object features may also play a significant role in guiding search in scenes.

### Set Size and Visual Search in Scenes

Although we did not explicitly manipulate it, the results of the present study also show an interesting effect of set size within scenes. Strangely, we found that across a number of eye movement measures within all three experiments, search performance in the Full Scene condition was better than conditions in which fewer objects in the scene were present (e.g., Sparse-Cluster). Past studies examining set size in search arrays have repeatedly shown that increasing set size will proportionally increase search time ([Enns, 1990](#); [Treisman, 1993](#)). However, compared to traditional search arrays, object counts in scenes seem to operate with different constraints. Instead, search performance seems to be based on a subset of the available object information ([Bravo & Farid, 2004](#); [Henderson, Chanceaux, & Smith, 2009](#); [Rosenholtz, Li, & Nakano, 2007](#); [Wolfe et al., 2011](#)). For instance, [Wolfe and colleagues \(2011\)](#) found that scenes produced more efficient searches than expected when search efficiency slopes (denoting ms/item) were examined. These findings are consistent with the notion that scene context aids search by greatly reducing the effective attentional set of the search space ([Torralba et al., 2006](#); [Wolfe et al., 2011](#)). However, this does not fully explain the patterns found here. In the current study, we believe that the Full Scene may be more advantageous than the other conditions in Experiments 1 and 2 because both scene context information *and* object information at the target region was available. This is supported by the results in Experiment 3, where the Full Scene differed in its visual search measure from the Empty and Sparse-Cluster Scene conditions, but did not differ from the Target-Cluster Scene condition, presumably because the latter condition also provides the same target location information.

### Guidance From High and Low Spatial Frequency Information

As stated in the introduction, many studies have looked at the influence of peripheral information on fixation placement by filtering out different spatial frequency bands from an image. From these studies, researchers have theorized that high spatial frequency information is associated with objects within scenes and low spatial frequency information is associated with global scene characteristics. Many researchers have posited that fixation placement is mostly influenced by low spatial frequency information because of the decreased perceptual acuity in the periphery ([Groner & Groner, 1996](#); [Groner, Groner, & von Mühlhelen, 2008](#); [Shioiri & Ikeda, 1989](#)). For example, in a picture naming task, [Groner et al. \(2008\)](#) found that saccade lengths were similar

between undegraded and low frequency images, but much shorter with high spatial frequency images. The authors thus concluded that fixation placement was based upon low frequency information. However, other researchers have shown that task constraints seem to have an influence on the type of information used for eye movement guidance. For instance, using a gaze-contingent moving-window paradigm, [Wampers and van Diepen \(1999\)](#) demonstrated that high-spatial frequency information in the periphery is preferentially selected over coarse, low frequency information during visual search tasks. In the present study, we manipulated high-level information directly and did not use filtered images; yet, we found a similar pattern of results as [van Diepen et al. \(1995\)](#), in that fixations were directed to object clusters. Further, we found that these object clusters seem to be selected according to their placement within the scene. Thus, although we did not use the same technique to examine the role of peripheral information, the data falls in line with that reported in previous studies.

### Saliency and Search Guidance

The notion that salient features attract attention in the periphery likely plays a role in the current study, though it was not a factor that was directly manipulated. Computational models of visual saliency define likely fixation points by highlighting areas that differ greatly from surrounding regions ([Bruce & Tsotsos, 2009](#); [Itti & Koch, 2000](#); [Itti, Koch, & Niebur, 1998](#); [Koch & Ullman, 1985](#)). As such, regions that are uniform are considered uninformative, whereas those that are different from neighboring regions are considered highly informative. Many behavioral studies have also shown that peripherally “informative” regions are likely to predict where fixations occur ([Mackworth & Morandi, 1967](#); [Reinagel & Zador, 1999](#)). For instance, [Reinagel and Zador \(1999\)](#) found that regions that correlate with high spatial contrast (as found in object content) tend to capture interest. Thus, information that is highly associated with object content seems to be a main driving force in guiding eye movements. In the present study, the placement of high density information was manipulated by using object clusters. In both Experiment 2 and 3, we found that the placement of this information affected fixation placement and search efficiency. In the case of a search task, this finding is intuitive, as the primary goal is linked to being able to select these basic features for further scrutiny. Although highly ecological, it is not clear if visual search tasks bias fixations to object clusters, and whether the same findings would be expected if a different task were used.

### Conclusion

In summary, the present study provides evidence that search strategies and eye movement guidance is modulated by an interaction between the information obtained from scene context and from object information. We found that although both have an effect, they do so to different extents: scene context acts as a framework for guiding eye movements to generally relevant regions, while object content provides information about specific areas to be targeted. Thus, our findings suggest that when discussing influential factors on processing of information in the periphery, it is important to consider the level of information being investigated.

## References

- Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14*, 589–592. doi:10.1016/0042-6989(74)90049-2
- Boucart, M., Moroni, C., Szaffarczyk, S., & Tran, T. H. C. (2013). Implicit processing of scene context in macular degeneration. *Investigative Ophthalmology & Visual Science*, *54*, 1950–1957. doi:10.1167/iov.12-9680
- Boucart, M., Moroni, C., Thibaut, M., Szaffarczyk, S., & Greene, M. (2013). Scene categorization at large visual eccentricities. *Vision Research*, *86*, 35–42. doi:10.1016/j.visres.2013.04.006
- Bravo, M. J., & Farid, H. (2004). Recognizing and segmenting objects in clutter. *Vision Research*, *44*, 385–396. doi:10.1016/j.visres.2003.09.031
- Bruce, N. D. B., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, *9*, 5. doi:10.1167/9.3.5
- Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Carpenter, R. H. S. (1988). *Movements of the eyes*. London: Pion.
- Castelhano, M. S., & Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Attention, Perception, & Psychophysics*, *72*, 1283–1297. doi:10.3758/APP.72.5.1283
- Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review*, *18*, 890–896. doi:10.3758/s13423-011-0107-8
- 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. doi:10.1037/0096-1523.33.4.753
- Castelhano, M. S., Pollatsek, A., & Cave, K. (2008). Typicality aids search for an unspecified target, but only in identification and not in attentional guidance. *Psychonomic Bulletin & Review*, *15*, 795–801. doi:10.3758/PBR.15.4.795
- Cohen, J. (1998). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Erlbaum.
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*, 559–564. doi:10.1111/j.0956-7976.2004.00719.x
- De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, *52*, 317–329. doi:10.1007/BF00868064
- Eckstein, M. P., Beutter, B. R., & Stone, L. S. (2001). Quantifying the performance limits of human saccadic targeting during visual search. *Perception*, *30*, 1389–1401. doi:10.1068/p3128
- Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science*, *17*, 973–980. doi:10.1111/j.1467-9280.2006.01815.x
- Ehinger, K. A., Hidalgo-Sotelo, B., Torralba, A., & Oliva, A. (2009). Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition*, *17*, 945–978. doi:10.1080/13506280902834720
- Enns, J. T. (1990). Three-dimensional features that pop-out in visual search. In D. Brogan (Ed.), *Visual search* (Vol. 1) (pp. 37–45). London, UK: Taylor & Francis.
- Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research*, *37*, 617–631. doi:10.1016/S0042-6989(96)00218-0
- Foulsham, T., & Kingstone, A. (2013). Optimal and preferred eye landing positions in objects and scenes. *Quarterly Journal of Experimental Psychology*, *66*, 1707–1728. doi:10.1080/17470218.2012.762798
- Foulsham, T., Teszka, R., & Kingstone, A. (2011). Saccade control in natural images is shaped by the information visible at fixation: Evidence from asymmetric gaze-contingent windows. *Attention, Perception, & Psychophysics*, *73*, 266–283. doi:10.3758/s13414-010-0014-5
- Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*, 6. doi:10.1167/8.2.6
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355. doi:10.1037/0096-3445.108.3.316
- Gordon, R. D. (2004). Attentional allocation during the perception of scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 760–777. doi:10.1037/0096-1523.30.4.760
- Groner, M. T., Groner, R., & von Mühlhelen, A. (2008). The effect of spatial frequency content on parameters of eye movements. *Psychological Research*, *72*, 601–608. doi:10.1007/s00426-008-0167-1
- Groner, R., & Groner, M. T. (1996). *The effect of spatial frequency filtering on eye movement parameters*. Paper presented at the Workshop on Spatial Scale Interactions, Durham, UK.
- Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford, UK: Elsevier. doi:10.1016/B978-008044980-7/50027-6
- Henderson, J. M., Chanceaux, M., & Smith, T. J. (2009). The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision*, *9*, 32.
- Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York, NY: Psychology Press.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271. doi:10.1146/annurev.psych.50.1.243
- Henderson, J. M., McClure, K. K., Pierce, S., & Schrock, G. (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception & Psychophysics*, *59*, 323–346. doi:10.3758/BF03211901
- Henderson, J. M., Weeks, P. A. Jr., & Hollingworth, A. (1999). Effects of semantic consistency on eye movements during scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228. doi:10.1037/0096-1523.25.1.210
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506. doi:10.1016/S0042-6989(99)00163-7
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*, 1254–1259. doi:10.1109/34.730558
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*, 3559–3565. doi:10.1016/S0042-6989(01)00102-X
- Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*, 6. doi:10.1167/9.10.6
- Laubrock, J., Cajar, A., & Engbert, R. (2013). Control of fixation duration during scene viewing by interaction of foveal and peripheral processing. *Journal of Vision*, *13*, 11. doi:10.1167/13.12.11
- Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 565–572. doi:10.1037/0096-1523.4.4.565
- Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multi-resolutional display. *Journal of Experimental Psychology: Applied*, *8*, 99–117. doi:10.1037/1076-898X.8.2.99

- Loschky, L. C., McConkie, G. W., Yang, J., & Miller, M. E. (2005). The limits of visual resolution in natural scene viewing. *Visual Cognition, 12*, 1057–1092. doi:10.1080/13506280444000652
- Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics, 2*, 547–552. doi:10.3758/BF03210264
- 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*, 8. doi:10.1167/9.11.8
- Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision, 10*, 4.
- Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision, 9*, 363–386. doi:10.1163/156856895X00052
- Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81*, 899–917. doi:10.1037/h0037368
- McConkie, G. W., Wolverson, G. S., & Loschky, L. C. (2001). An environment for studying gaze-contingent multi-resolutional displays. In M. S. Vassiliou & T. S. Huang (Eds.), *Computer-science handbook for displays: Summary of findings from the army research lab's advanced displays and interactive displays federated laboratory* (pp. 55–61). Thousand Oaks, CA: Rockwell Scientific Company.
- Motter, B. C. (1994). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *The Journal of Neuroscience, 14*, 2190–2199.
- Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research, 46*, 614–621. doi:10.1016/j.visres.2005.08.025
- Nuthmann, A. (2014). How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 342–360. doi:10.1037/a0033854
- Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision, 10*, 20. doi:10.1167/10.8.20
- Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34*, 72–107. doi:10.1006/cogp.1997.0667
- Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*, 520–527. doi:10.1016/j.tics.2007.09.009
- Pomplun, M., Reingold, E. M., & Shen, J. (2001). Peripheral and parafoveal cueing and masking on saccadic selectivity in a gaze-contingent window paradigm. *Vision Research, 41*, 2757–2769. doi:10.1016/S0042-6989(01)00145-6
- Rayner, K. (2009). Eye movements and attention during reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506. doi:10.1080/17470210902816461
- Reinagel, P., & Zador, A. M. (1999). Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems, 10*, 341–350. doi:10.1088/0954-898X/10/4/304
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron, 26*, 703–714. doi:10.1016/S0896-6273(00)81206-4
- Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7*, 17.
- Saida, S., & Ikeda, M. (1979). Useful visual field size for pattern perception. *Perception & Psychophysics, 25*, 119–125. doi:10.3758/BF03198797
- Schyns, P., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5*, 195–200. doi:10.1111/j.1467-9280.1994.tb00500.x
- Scialfa, C. T., & Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception & Psychophysics, 60*, 1067–1082. doi:10.3758/BF03211940
- Shioiri, S., & Ikeda, M. (1989). Useful resolution for picture perception as a function of eccentricity. *Perception, 18*, 347–361. doi:10.1068/p180347
- Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2014). How context information and target information guide the eyes from the first epoch of search in real-world scenes. *Journal of Vision, 14*, 7. doi:10.1167/14.2.7
- Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*, 13. doi:10.1167/11.5.13
- Torralba, A., Oliva, A., Castelhana, 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
- Treisman, A. (1993). The perception of features and objects. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control* (pp. 5–35). Oxford, UK: Clarendon Press.
- van Diepen, P. M. J. (1997). A pixel-resolution video switcher for eye contingent display changes. *Spatial Vision, 10*, 335–344. doi:10.1163/156856897X00258
- van Diepen, P. M. J., Wampers, M., & d'Ydewalle, G. (1995). *The use of coarse and fine peripheral information during scene perception*. Paper presented at the Eighth European Conference on Eye Movements, Derby, UK.
- van Diepen, P. M. J., Wampers, M., & d'Ydewalle, G. (1998). Functional division of the visual field: Moving masks and moving windows. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 337–355). Oxford, UK: Elsevier. doi:10.1016/B978-008043361-5/50016-X
- 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
- Wampers, M., & van Diepen, P. M. J. (1999). The use of coarse and fine peripheral information during the final part of fixations in scene perception. In W. Becker, H. Deubel, & T. Mergner (Eds.), *Current Oculomotor Research* (pp. 257–267). New York, NY: Plenum Press. doi:10.1007/978-1-4757-3054-8\_36
- Williams, D. E., & Reingold, E. M. (2001). Preattentive guidance of eye movements during triple conjunction search tasks: The effects of feature discriminability and saccadic amplitude. *Psychonomic Bulletin & Review, 8*, 476–488. doi:10.3758/BF03196182
- Wilson, H. R., Levi, D., Maffei, L., Rovamo, J., & De Valois, R. (1990). The perception of form: Retina to striate cortex. In L. Spillmann & J. S. Werner (Eds.), *Visual perception: The neurophysiological foundations* (pp. 231–272). New York, NY: Academic Press. doi:10.1016/B978-0-12-657675-7.50016-8
- Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception, & Psychophysics, 73*, 1650–1671. doi:10.3758/s13414-011-0153-3
- Wolfe, J. M., Vö, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*, 77–84. doi:10.1016/j.tics.2010.12.001
- Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115*, 787–835. doi:10.1037/a0013118
- Zelinsky, G. J., & Schmidt, J. (2009). An effect of referential scene constraint on search implies scene segmentation. *Visual Cognition, 17*, 1004–1028. doi:10.1080/13506280902764315

Received April 3, 2014

Revision received June 19, 2014

Accepted June 20, 2014 ■